View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Carolina Digital Repository

Published in final edited form as:
Psychiatr Genet. 2008 April ; 18(2): 85–91. doi:10.1097/YPG.0b013e3282f9b48e.

Posterior probability of linkage analysis of autism dataset identifies linkage to chromosome 16

Thomas H. Wassink[a], **Veronica J. Vieland**[c], **Val C. Sheffield**[b], **Christopher W. Bartlett**[c], **Rhinda Goedken**[a], **Deborah Childress**[d], and **Joseph Piven**[d]

[a]Department of Psychiatry, University of Iowa, Iowa

[b]Department of Pediatrics and the Howard Hughes Medical Institute, Carver College of Medicine, University of Iowa, Iowa

[c]Department of Center for Quantitative and Computational Biology, Columbus Children's Research Institute and The Ohio State University, Ohio

[d]Department of Neurodevelopmental Disorders, Research Center and Department of Psychiatry, University of North Carolina, North Carolina, USA

Abstract

Objective—To apply phenotypic and statistical methods designed to account for heterogeneity to linkage analyses of the autism Collaborative Linkage Study of Autism (CLSA) affected sibling pair families.

Method—The CLSA contains two sets of 57 families each; Set 1 has been analyzed previously, whereas this study presents the first analyses of Set 2. The two sets were analyzed independently, and were further split based on the degree of phrase speech delay in the siblings. Linkage analysis was carried out using the posterior probability of linkage (PPL), a Bayesian statistic that provides a mathematically rigorous mechanism for combining linkage evidence across multiple samples.

Results—Two-point PPLs from Set 1 led to the follow-up genotyping of 18 markers around linkage peaks on 1q, 13p, 13q, 16q, and 17q in both sets of families. Multipoint PPLs were then calculated for the entire CLSA sample. These analyses identified four regions with at least modest evidence in support of linkage: 1q at 173 cM, PPL = 0.12; 13p at 21 cM, PPL = 0.16; 16q at 63 cM, PPL= 0.36; Xq at 40 cM, PPL = 0.11.

Conclusion—We find strengthened evidence for linkage of autism to chromosomes 1q, 13p, 16q, and Xq, and diminished evidence for linkage to 7q and 13q. The verity of these findings will be tested by continuing to update our PPL analyses with data from additional autism datasets.

Keywords

autism; genetics; heterogeneity; language delay; linkage; positional cloning

© 2008 Wolters Kluwer Health | Lippincott Williams & Wilkins.

Correspondence to Thomas H. Wassink, MD, University of Iowa Carver College of Medicine, Psychiatry Research/MEB, Iowa City, IA 52242, USA Tel: +319 335 7743; fax: + 319 353 3003; thomas-wassink@uiowa.edu.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Introduction

Autism is a developmental disorder characterized by impairments in language and social interaction and by excessively repetitive and ritualistic behaviors. Onset is before the age of three, and though the constellation and severity of symptoms may change over time, they typically persist in some form throughout life. Family and twin studies have shown that autism susceptibility is attributable primarily to genetic factors, with a heritability estimate that approaches 90% and a sibling relative risk of 25–50 (Folstein and Rutter, 1977; Bailey *et al*., 1995; Pickles *et al*., 1995; Szatmari *et al*., 1998; Risch *et al*., 1999). Linkage studies have been a primary means of searching for susceptibility genes (IMGSAC, 1998; CLSA, 1999; Philippe *et al*., 1999; Risch *et al*., 1999; Liu *et al*., 2001; Alarcon *et al*., 2002; Auranen *et al*., 2002; Shao *et al*., 2002b; Yonan *et al*., 2003). Our research group, the Collaborative Linkage Study of Autism (CLSA), carried out one of these, genotyping 350 markers in 75 autism affected sibling pair (ASP) families (CLSA, 1999). The strongest linkages that we reported were multipoint heterogeneity logarithm of the odds (LOD) scores of 3.0 at 55 cM on chromosome 13, 2.3 at 19 cM on chromosome 13, and 2.2 at 110 cM on chromosome 7 (cM distances are from the Marshfield sex-averaged genetic map). Other similar studies have also generally yielded modest evidence for linkage at multiple loci, replication across studies has been limited, and in studies with expanding samples the locations and strengths of linkages have often changed, so that it is unclear whether any single locus has yet been definitively confirmed (Wassink *et al*., 2004).

The heterogeneity of autism is likely to underlie much of this impasse. Family and twin studies suggest that autism arises from interactions of multiple genetic variants, and that different combinations of variants may be causative in different groups of people. We are likely, therefore, to be pooling under the diagnosis of autism multiple etiologies that should instead be distinguished from one another. Even with improved phenotyping, however, the power of our analyses is diminished by statistical methods that do not appropriately model heterogeneity. Traditional LOD score and model-free methods either merge multiple subgroups into one ‘pooled’ analysis or analyze the subgroups independently. In the presence of heterogeneity, true linkages from independent samples can be obscured by pooling (Huang and Vieland, 2001; Vieland *et al*., 2001), whereas treating subgroups separately constitutes a form of maximization over the phenotype that can artificially inflate evidence for linkage (Weeks *et al*., 1990).

In our initial linkage screen, therefore, we tested on a limited scale what were at the time novel phenotypic and statistical strategies designed to address heterogeneity. First we reanalyzed the CLSA linkage data from the 7q and 13q loci using language characteristics to enrich the phenotype (CLSA *et al*., 2001). Families were classified as phrase speech delay (PSD) ‘positive’ if at least two autistic children had failed to develop phrase speech by 36 months, and otherwise as PSD ‘negative’ (the average normal onset of phrase speech is 18 months; 2/3 of families were PSD positive). Analyses of the two subgroups revealed that the predominance of the 7q and 13q linkage signals arose from the PSD positive families, a pattern subsequently observed in other autism linkage studies (Buxbaum *et al*., 2001; Shao *et al*., 2002a).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

In addition, we had assessed by direct questioning whether the CLSA 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 these phenotypes into the analyses strengthened the evidence for linkage at the 7q and 13q loci in the PSD positive families. These results thus supported the hypotheses that: (i) language based classification defines more homogeneous subgroups of families with autism, and (ii) parental language phenotype is in some way genetically related to autism.

To address the analytical challenge posed by heterogeneity, we turned to the posterior probability of linkage (PPL) statistic (Vieland, 1998). The PPL is a Bayesian statistic that is essentially model-free. It is calculated on the probability scale (0, 1), so that a value of 0.60 (60%) indicates a 60% probability of linkage at a particular locus. The PPL produces significantly diminished 'noise' in comparison to traditional LOD score and nonparametric statistics and, of import for our analyses, provides a mathematically rigorous mechanism for integrating evidence for or against linkage across multiple samples and subgroups of families (Vieland *et al*., 2001; Logue *et al*., 2003; Logue and Vieland, 2004). The efficacy of the PPL has been demonstrated in both simulated and real datasets (Wang *et al*., 1999; Huang and Vieland, 2001; Vieland *et al*., 2001; Bartlett *et al*., 2002, 2004; Logue *et al*., 2003), and we recently carried out a PPL analysis of genotyping data from a large, publicly available collection of autism families – the Autism Genetics Resource Exchange (AGRE) sample – using both diagnosis and language based subgrouping (Bartlett *et al*., 2005).

In this report, we expand implementation of these phenotypic and statistical strategies to encompass the entire CLSA sample. We also present analyses of new genotyping in chromosomal regions of interest in a second set of CLSA families not previously described.

Methods

Collaborative Linkage Study of Autism families

Families for the CLSA were recruited from the Midwest and Mid-Atlantic United States through data collection sites at the University of Iowa and the University of North Carolina, as described in our original report (CLSA, 1999). All affected individuals were at least 3 years old and were required to meet the Autism Diagnostic Interview, Revised (Lord *et al*., 1994) and the Autism Diagnostic Observation Schedule (Lord *et al*., 1989) or Autism Diagnostic Observation Schedule–Generic (Lord *et al*., 2000) criteria for autism. Affecteds were excluded if they had fragile X syndrome (based on fragile X DNA testing), or any other neurological or medical condition suspected of being associated with autism. DNA was extracted from whole blood using standard techniques.

Our initial study reported on a total of 75 ASP families gathered in this way; 18 of these families were provided through a previous collaboration and are no longer available, leaving 57 families as Set 1. We have since enrolled another 57 families that are contained within Set 2, providing a total of 114 CLSA ASP families for analysis.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Language subsetting of the data

The Set 1 and 2 families were each split in two, being classified as PSD positive if at least two autistic children had failed to develop phrase speech by 36 months, and otherwise as PSD negative, creating a total of four separate subgroups for analysis. Parents were coded as affected or unaffected based on a history of developmental language problems. The numbers of families and PSD positive parents in each group are shown in Table 1. No notable differences in the proportions of these classifications or any other demographic or phenotypic characteristics between the Set 1 and 2 families were observed.

Genotyping

A total of 383 short tandem repeat polymorphism markers representing a modified version of Weber set 8 (http://research.marshfieldclinic.org/genetics/sets/Set8ScreenFrames.htm) were genotyped in the Set 1 families in our initial genome screen (CLSA, 1999). Genotyping had been carried out using standard PCR protocols with products visualized using silver stained polyacrylamide gels. Since that time, additional markers had been genotyped in regions of interest arising from the initial study. In the course of this follow-up, some of the original markers were regenotyped to confirm the data. This genotyping revealed that a number of the original markers had error rates ranging from 5–10%. Further assessment revealed that the problematic markers all came from one genotyping site. We therefore had all markers from that site – approximately 190 – regenotyped by Weber Genotyping Service using the same or nearby markers from Weber set 13. From the two-point PPL analysis of CLSA Set 1 (described below), we selected five regions for follow-up study (1q, 13p, 13q, 16q, and 17q).

For each region, we genotyped the peak marker in the Set 2 families, and additional flanking markers in the combined Set 1 and 2 families. Ten additional markers were genotyped in the chromosome 16 region, whereas one marker was genotyped on each side of the peak in the other four regions. Thus a total of 18 new markers were genotyped in both sets of families, and in addition the five peak markers were genotyped in the Set 2 families (marker information available upon request). PCR was carried out with 20 ng of genomic DNA amplified in a reaction mixture containing 1.0 μl of PCR buffer [100 mmol/l Tris-HCl (pH 8.8), 500 mmol/l KCl, 15 mmol/l $MgCl_2$, 0.01% gelatine (w/v)], 200 μmol/l each of dATP, dCTP, dGTP, and dTTP, 2.5 pM of each primer, and 0.05 units of Taq DNA polymerase, increased to a final volume of 10.0 μl with water. Samples were initially denatured at 941 for 3 min, followed by 35 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 30 s. PCR products were then electrophoresed on 6% polyacrylamide gels at 60 W for approximately 2 h and visualized with silver staining (Bassam *et al.*, 1991). Gels were read blindly with respect to sample status by two independent raters, with discrepancies resolved by regenotyping.

Error checking was carried out using PedCheck (O'Connell and Weeks, 1998) to identify non-Mendelian transmissions and GeneHunter (Kruglyak *et al.*, 1996) to identify double recombinants.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Errors flagged in this way were regenotyped. Wherever an inconsistency persisted, alleles for that marker were zeroed out in the entire family from which the inconsistency arose. Marker allele frequencies were calculated by counting alleles in the founders separately for Set 1 and 2.

Linkage analysis

The PPL has been described in detail elsewhere (Vieland, 1998). In brief, the statistic is computed from the Bayes ratio, or the ordinary single-locus LOD score allowing for locus heterogeneity under the admixture model (Smith, 1963), with the trait parameters, including the admixture parameter, integrated out. As these parameters are integrated out rather than maximized over, we incur no 'penalty' or inherent inflation of the result owing to the inclusion of trait parameters in the model, nor do we need to 'fix' the values of the trait parameters. Thus, though it is based on a single-locus approximation to the true likelihood, the method is essentially model-free. In addition, because it is based on the full likelihood, phenotypic information on all family members is readily incorporated. In this analysis, we have assigned the phenotypes affected, unaffected, and unknown to parents based on the developmental language variable described above. The PPL itself is on the probability scale, converging to 1 (as a function of effective sample size) when there is linkage, and to 0 when there is no linkage (Wang *et al*., 2000). It incorporates a small 2% prior probability of linkage (Elston and Lange, 1975; Morton, 1998), so that a PPL greater than 0.02 indicates some degree of evidence in favor of linkage, whereas a PPL < 0.02 indicates evidence against linkage.

When analyzing multiple datasets or families divided into subgroups, the PPL passes forward the posterior distribution for the recombination fraction θ (for two-point analysis), or the location parameter (for multipoint analysis) derived from one dataset as the prior distribution for analysis of the next. This process is repeated as each new set of data is incorporated into the analysis. The PPL allows for between-group differences by integrating trait parameters out of the model independently for each data set, thus enabling accurate quantification of the posterior probability based on all of the data. The final value of the PPL is independent of the order in which the data sets are fed in to the calculation (Vieland *et al*., 2001). In addition, dividing the data into subsets does not, in and of itself, affect the final value of the PPL. If data are grouped based on a genetically irrelevant feature, the PPL (on average) remains the same, as if no subsetting had been carried out. If, however, subsetting results in more homogeneous subgroups, the sequentially updated PPL will tend to be larger than a PPL carried out without subsetting.

We first report two-point results for CSLA Set 1 families, which entailed sequentially updating over PSD positive and PSD negative subsets. This serves primarily to allow comparison with our earlier results in a largely overlapping set of families. For multipoint PPL analysis, we incorporated the CLSA Set 2 PSD positive and negative families and walked down the chromosome in fixed 1 cM steps, sequentially updating the statistic across data sets at each location (Logue and Vieland, 2004).

Results

Two-point posterior probability of linkage analysis of Set 1 Collaborative Linkage Study of Autism families

Of the 383 markers analyzed in Set 1, 61 (16%) had PPLs greater than 2% (the initial prior probability of linkage), and only nine (2%) had PPLs of 10% or higher; 254 (66%) markers had PPLs less than 2%, indicating some degree of evidence against linkage. The nine markers with PPLs 10% were in six chromosomal regions: 1q, 13p, 13q, 16q, 17q, and X (Table 2). Two of these – ATA5A09N on chromosome 13p at 20 cM, and 13AL159 on 13q at 55 cM – corresponded to previously reported peak heterogeneity logarithm of the odds scores (HLODs) in this set of families (CLSA, 1999, 2001). The 1q, 16q, 17q, and X loci, however, were not suggested by our previous analyses, and the largest PPL on chromosome 7 was 2%.

Multipoint posterior probability of linkage analysis of Set 1 and 2 families

Two-point PPLs of the follow-up markers showed no evidence for linkage in the second Set families (data not shown). Figure 1 shows a graph of the multi-point PPLs across the genome that incorporates the 18 follow-up markers and the Set 2 families (recall that we genotyped Set 2 only for these 18 markers, so that the remainder of the multipoint results reflect Set 1 alone). Approximately 69% of the genome shows PPLs less than 2%, whereas only 1.6% has PPLs 10%. Four peaks stand out: 1q, 13p, 16q, and X (Table 2 and Fig. 2). Not surprisingly, these peaks correspond to regions with the highest two-point PPLs. We note, however, that at every location the multipoint are smaller than the corresponding two-point PPLs. This is not a general property of the PPL (Logue and Vieland, 2004), and we have carried out thorough error checking and regenotyping of peak markers to ensure accuracy of genotypes. The 17q multipoint PPL is just 6%, the 13q multipoint PPL is 4%, and the multipoint PPL on 16, whereas still the largest in the CLSA dataset (36% at 63 cM), is considerably smaller than the two-point PPL from the Set 1 families of 95%.

Overall, the Set 2 families contributed little to the signals in the regions of interest identified on the basis of the Set 1 families, which is consistent with the tendency of locus heterogeneity to make replication of any one locus in a second set of families improbable (Suarez *et al*., 1994; Vieland *et al*., 2001).

Discussion

We have carried out linkage analysis in a sample of autism affected sibling pair families using enriched phenotypic information. We have analyzed the resultant data using the PPL statistic, which is designed to rigorously assess evidence for or against linkage across multiple, potentially heterogeneous sets of data. The implementation of these methods reflects our efforts to account for and address the phenotypic and genetic heterogeneity of autism, with the result that previously unidentified regions of interest have emerged, whereas others have diminished.

From the original CLSA genome-wide linkage screen, the three strongest signals based on heterogeneity LOD score analyses were on 13q (HLOD = 3.0 at 55 cM), 7q (HLOD = 2.2 at

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript

110 cM), and 13p (HLOD = 2.3 at 19 cM) (CLSA *et al*., 1999, 2001). In the new PPL analysis, the 13p signal remains substantial, the 13q signal is small, and no evidence for linkage exists on 7q. The differences in these results are likely to be owing to a combination of the factors that distinguish the current from the former study: (i) genome-wide application of language based classifications, (ii) genome-wide application of the PPL, (iii) 18 of the families from the original Set 1 are not in the current set, and (iv) a large subset of the screening markers were regenotyped.

The strongest evidence for linkage in the new analyses is to chromosome 16q (Table 2, Fig. 2). This locus is substantially removed from chromosome 16 linkages identified in previous studies of autism (Philippe *et al*., 1999; Lucarelli *et al*., 2003; Barnby *et al*., 2005; Philippi *et al*., 2005), though a number of reports have described autistic individuals with chromosomal abnormalities near our peak (Hebebrand *et al*., 1994; Finelli *et al*., 2004). Candidate genes of interest in this region include *SALL1, CBLN1, NKD1, N4BP1*, and *NETO2*. The chromosome X linkage peak also appears to be novel, with only modest evidence for linkage to X having been previously reported at significant distances from our locus (Hallmayer *et al*., 1996; Liu *et al*., 2001; Shao *et al*., 2002b). Numerous genes on X have been tested as autism candidates, including *ARX, MECP2, NLGN3, NLGN4, MAOA, MAOB*, and *FMR1*. Of these, *MAOA* and *MAOB* at 43 MB are near our peak and may therefore warrant closer scrutiny in the CLSA families. The chromosome 1 locus overlaps with findings from a number of previous studies. Bartlett *et al*. reported a multipoint PPL of 55% at 183 cM in the Autism Genetics Resource Exchange sample of families, and two studies from Finland – one of autism (Auranen *et al*., 2002) and the other of Asperger's syndrome (Ylisaukko-oja *et al*., 2004) – found linkage to 1q markers at 165 and 170 cM, respectively. We also recently described, in an autistic proband, uniparental disomy of chromosome 1 with a region of isodisomy that overlaps this interval (Wassink *et al*., 2005). Lastly, the linkage signal on 13p, which was the most well-preserved from the original genome-wide screen, has not been reported in other linkage studies of autism.

Thus, we find new or strengthened evidence for linkage to specific loci on chromosomes 1q, 13p, 16q, and Xq, and diminished or absent evidence for linkage to the previously identified 7q and 13q loci. The 'significance' of these findings in more traditional LOD score and probability terms is difficult to determine. The PPL cannot be translated directly to a LOD score, and because the PPL is a Bayesian measure and thus conditioned on the observed data, it cannot be assigned a significance level, because a significance level describes the probability of the observed data given all possible outcomes (Bartlett *et al*., 2004). Given previous studies examining the properties and utility of the PPL, and the ability of the statistic to take advantage of phenotypically defined subgroups, we believe that the results we present here most accurately reflect the evidence for linkage in our data. Furthermore, based on 3600 simulated sib-pair datasets, we have observed PPLs greater than 10% only 0.4% of the time in the absence of true underlying linkage (pointwise significance), which is consistent with previous studies in extended pedigrees (Logue and Vieland, 2004). We will only come to know with certainty, however, whether the PPLs that we describe in this report are real or not by continuing to update our analyses with data from additional autism datasets. We therefore look forward to applying the PPL to alternative phenotypes, and to

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

incorporating additional samples of families into our analyses as genotyping from those samples becomes available.

Acknowledgements

The authors gratefully acknowledge grant support for C.W.B. from T32 HL07638, grant support for V.J.V. and R.G. from R01 MH052841, grant support for T.H.W. from K08-MH062123 and ROINS43550, and grant support for J.P. from R-01 MH55284.

References

Alarcon M, Cantor RM, Liu J, Gilliam TC, Geschwind DH. Evidence for a language quantitative trait locus on chromosome 7q in multiplex autism families. Am J Hum Genet. 2002; 70:60–71. [PubMed: 11741194]

Auranen M, Vanhala R, Varilo T, Ayers K, Kempas E, Ylisaukko-Oja T, et al. A genomewide screen for autism-spectrum disorders: evidence for a major susceptibility locus on chromosome 3q25–27. Am J Hum Genet. 2002; 71:777–790. [PubMed: 12192642]

Bailey A, Le CA, Gottesman I, Bolton P, Simonoff E, Yuzda E, Rutter M. Autism as a strongly genetic disorder: evidence from a British twin study. Psychol Med. 1995; 25:63–77. [PubMed: 7792363]

Barnby G, Abbott A, Sykes N, Morris A, Weeks DE, Mott R, et al. Candidate gene screening and association analysis at the autism susceptibility locus on chromosome 16p: evidence of association at GRIN2A and ABAT. Am J Hum Genet. 2005; 76:950–966. [PubMed: 15830322]

Bartlett CW, Flax JF, Logue MW, Vieland VJ, Bassett AS, Tallal P, Brzustowicz LM. A major susceptibility locus for specific language impairment is located on 13q21. Am J Hum Genet. 2002; 71:45–55. [PubMed: 12048648]

Bartlett CW, Flax JF, Logue MW, Smith BJ, Vieland VJ, Tallal P, Brzustowicz LM. 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]

Bartlett CW, Goedken RJ, Vieland VJ. Effects of updating linkage evidence across subsets of data: reanalysis of autism genetic resource exchange dataset. Am J Hum Genet. 2005; 76:688–695. [PubMed: 15729670]

Bassam BJ, Caetano-Anolles G, Gresshoff PM. Fast and sensitive silver staining of DNA in polyacrylamide gels. Anal Biochem. 1991; 196:80–83. [PubMed: 1716076]

Buxbaum JD, Silverman JM, Smith CJ, Kilifarski M, Reichert J, Hollander E, et al. Evidence for a susceptibility gene for autism on chromosome 2 and for genetic heterogeneity. Am J Hum Genet. 2001; 68:1514–1520. [PubMed: 11353400]

CLSA. An autosomal genomic screen for autism. Am J Med Genet. 1999; 88:609–615. [PubMed: 10581478]

CLSA. Incorporating language phenotypes strengthens evidence of linkage to autism. Am J Med Genet. 2001; 105:539–547.

Barrett S, Beck JC, Bernier R, Bisson E, Braun TA, Casavant TL, et al. CLSA. An autosomal genomic screen for autism. Collaborative linkage study of autism. Am J Med Genet. 1999; 88:609–615. [PubMed: 10581478]

CLSA, Bradford Y, Haines J, Hutcheson H, Gardiner M, Braun T, Sheffield V, et al. Incorporating language phenotypes strengthens evidence of linkage to autism. Am J Med Genet. 2001; 105:539–547.

Elston RC, Lange K. The prior probability of autosomal linkage. Annals of Human Genet. 1975; 38:341–350.

Finelli P, Natacci F, Bonati MT, Gottardi G, Engelen JJ, de Die-Smulders CE, et al. FISH characterisation of an identical (16)(p11.2 p12.2) tandem duplication in two unrelated patients with autistic behaviour. J Med Genet. 2004; 41:e90. [PubMed: 15235033]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Folstein S, Rutter M. Infantile autism: a genetic study of 21 twin pairs. J Child Psychol Psychiatry. 1977; 18:297–321. [PubMed: 562353]

Hallmayer J, Hebert JM, Spiker D, Lotspeich L, McMahon WM, Petersen PB, et al. Autism and the X chromosome. Multipoint sib-pair analysis. Arch Gen Psychiatry. 1996; 53:985–989. [PubMed: 8911221]

Hebebrand J, Martin M, Korner J, Roitzheim B, de BK, Werner W, Remschmidt H. Partial trisomy 16p in an adolescent with autistic disorder and Tourette's syndrome. Am J Med Genet. 1994; 54:268–270. [PubMed: 7810587]

Huang J, Vieland VJ. Comparison of 'model-free' and 'model-based' linkage statistics in the presence of locus heterogeneity: single data set and multiple data set applications. Hum Hered. 2001; 51:217–225. [PubMed: 11287743]

IMGSAC. A full genome screen for autism with evidence for linkage to a region on chromosome 7q. International Molecular Genetic Study of Autism Consortium. Hum Mol Genet. 1998; 7:571–578. [PubMed: 9546821]

Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet. 1996; 58:1347–1363. [PubMed: 8651312]

Liu J, Nyholt DR, Magnussen P, Parano E, Pavone P, Geschwind D, et al. A genomewide screen for autism susceptibility loci. Am J Hum Genet. 2001; 69:327–340. [PubMed: 11452361]

Logue MW, Vieland VJ. A new method for computing the multipoint posterior probability of linkage. Hum Hered. 2004; 57:90–99. [PubMed: 15192281]

Logue MW, Vieland VJ, Goedken RJ, Crowe RR. 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, Neuropsychiat Genet. 2003; 121B:95–99. [PubMed: 12898582]

Lord C, Rutter M, Goode S, Heemsbergen J, Jordan H, Mawhood L, Schopler E. Autism diagnostic observation schedule: a standardized observation of communicative and social behavior. J Autism Dev Disord. 1989; 19:185–212. [PubMed: 2745388]

Lord C, Rutter M, Le CA. 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]

Lord C, Risi S, Lambrecht L, Cook EH Jr, Leventhal BL, DiLavore PC, 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]

Lucarelli P, Palminiello S, Saccucci P, Bottini N, De Luca D, Elia M, et al. Association study of autistic disorder and chromosome 16p. Am J Med Genet A. 2003; 119:242–246. [PubMed: 12749074]

Morton NE. Significance levels in complex inheritance. Am J Hum Genet. 1998; 62:690–697. [PubMed: 9497238]

O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet. 1998; 63:259–266. [PubMed: 9634505]

Philippe A, Martinez M, Guilloud-Bataille M, Gillberg C, Rastam M, Sponheim E, et al. Genome-wide scan for autism susceptibility genes. Paris Autism Research International Sibpair Study. Hum Mol Genet. 1999; 8:805–812. [PubMed: 10196369]

Philippi A, Roschmann E, Tores F, Lindenbaum P, Benajou A, Germain-Leclerc L, et al. Haplotypes in the gene encoding protein kinase c-beta (PRKCB1) on chromosome 16 are associated with autism. Mol Psychiatry. 2005; 10:950–960. [PubMed: 16027742]

Pickles A, Bolton P, Macdonald H, Bailey A, Le CA, Sim CH, Rutter M. 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. 1995; 57:717–726. [PubMed: 7668301]

Risch M, Spiker D, Lotspeich L, Nouri N, Hinds D, Hallmayer J, et al. A genomic screen of autism: evidence for a multilocus etiology. Am J Hum Genet. 1999; 65:493–507. [PubMed: 10417292]

Shao Y, Raiford KL, Wolpert CM, Cope HA, Ravan SA, Ashley-Koch AA, et al. Phenotypic homogeneity provides increased support for linkage on chromosome 2 in autistic disorder. Am J Hum Genet. 2002a; 70:1058–1061. [PubMed: 11875756]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Shao Y, Wolpert CM, Raiford KL, Menold MM, Donnelly SL, Ravan SA, et al. Genomic screen and follow-up analysis for autistic disorder. Am J Med Genet. 2002b; 114:99–105. [PubMed: 11840513]

Smith CAB. Testing for heterogeneity of recombination fraction values in human genetics. Annals Hum Genet. 1963; 27:175–182.

Suarez, B.; Hampe, CL.; van Eerdewegh, P. Problems of replicating linkage claims in psychiatry. In: Gershon, ES.; Cloninger, CR., editors. Genetic approaches to mental disorders. Washington DC: American Psychiatric Press; 1994.

Szatmari P, Jones MB, Zwaigenbaum L, MacLean JE. Genetics of autism: overview and new directions. J Autism Dev Disord. 1998; 28:351–368. [PubMed: 9813773]

Vieland VJ. Bayesian linkage analysis, or: how I learned to stop worrying and love the posterior probability of linkage. Am J Hum Genet. 1998; 63:947–954. [PubMed: 9758634]

Vieland VJ, Wang K, Huang J. 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. 2001; 51:199–208. [PubMed: 11287741]

Wang K, Vieland V, Huang J. A Bayesian approach to replication of linkage findings. Genet Epidemiol. 1999; 17:S749–S754. [PubMed: 10597525]

Wang K, Huang J, Vieland VJ. The consistency of the posterior probability of linkage. Ann Hum Genet. 2000; 64:533–553. [PubMed: 11281217]

Wassink TH, Brzustowicz LM, Bartlett CW, Szatmari P. The search for autism disease genes. Ment Retard Dev Disabil Res Rev. 2004; 10:272–283. [PubMed: 15666342]

Wassink TH, Losh M, Frantz RS, Vieland VJ, Goedken RJ, Piven J, Sheffield VC. A case of autism and uniparental disomy of chromosome 1. Hum Genet. 2005; 11:200–2006. [PubMed: 15887000]

Weeks DE, Lehner T, Squires-Wheeler E, Kaufmann C, Ott J. Measuring the inflation of the lod score due to its maximization over model parameter values in human linkage analysis. Genet Epidemiol. 1990; 7:237–243. [PubMed: 2227370]

Ylisaukko-oja T, Nieminen-von Wendt T, Kempas E, Sarenius S, Varilo T, von Wendt L, et al. Genome-wide scan for loci of Asperger syndrome. Mol Psychiatry. 2004; 9:161–168. [PubMed: 14966474]

Yonan AL, Alarcon M, Cheng R, Magnusson PK, Spence SJ, Palmer AA, et al. A genomewide screen of 345 families for autism-susceptibility loci. Am J Hum Genet. 2003; 73:886–897. [PubMed: 13680528]

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

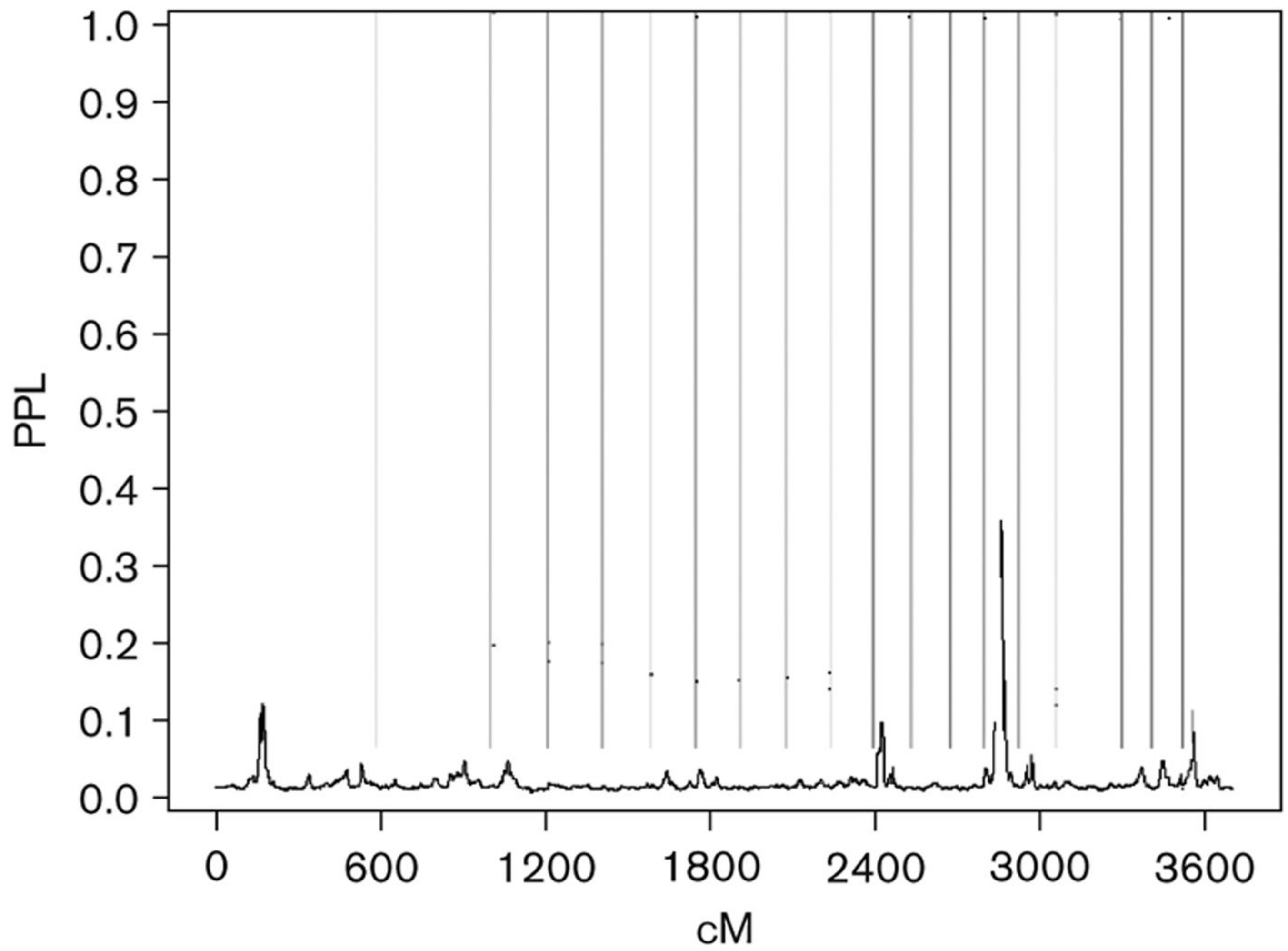

Fig. 1.
Multipoint posterior probability of linkages (PPLs) for combined Collaborative Linkage Study of Autism Set 1 and 2 families. The PPLs for this analysis were updated over the two language-based subgroups as described in Methods.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

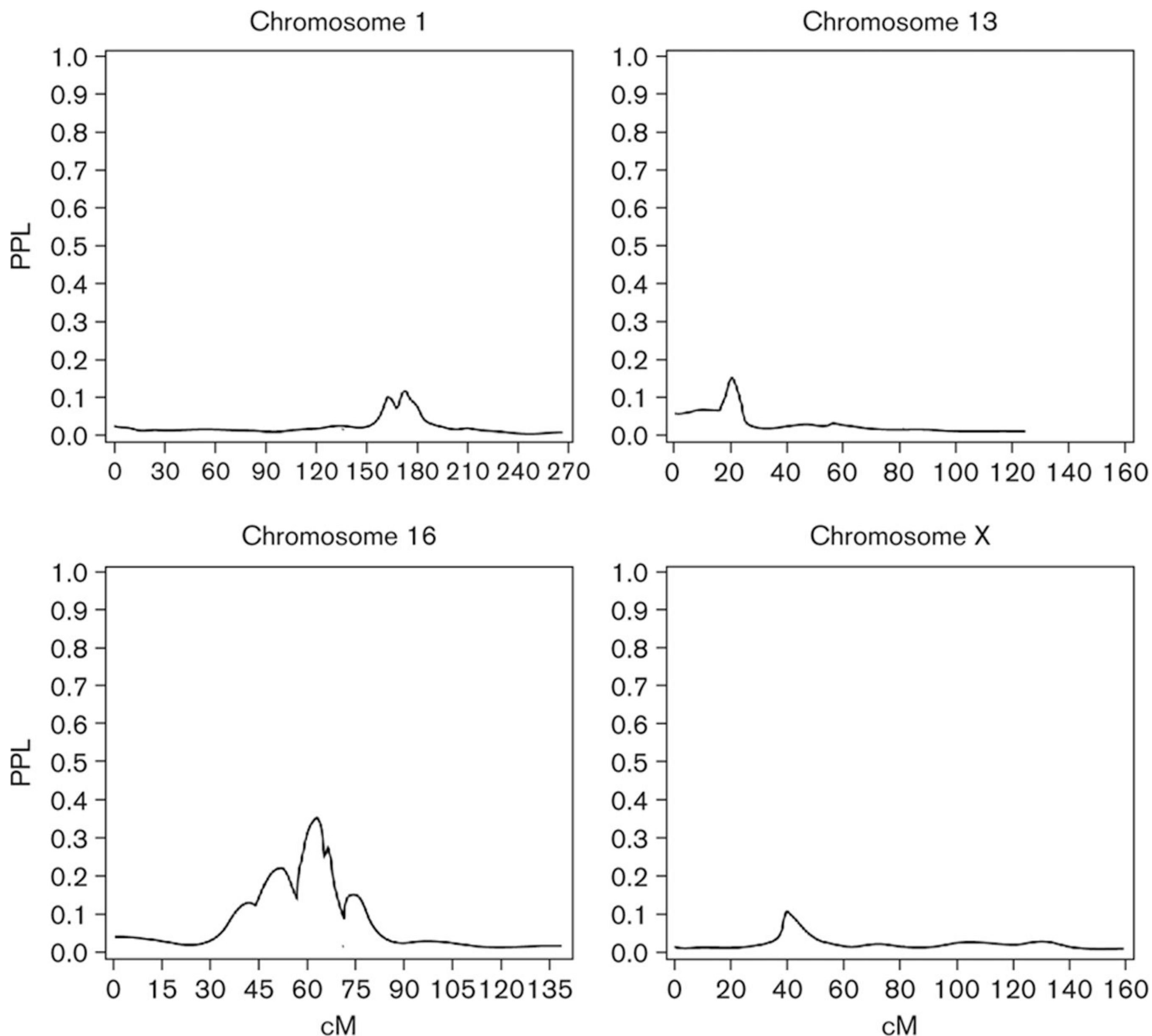

Fig. 2.

Collaborative Linkage Study of Autism multipoint posterior probability of linkages (PPLs) for selected chromosomes. These graphs present the multipoint PPL for the four strongest linkage signals. These scores are derived from PPL updating over four groups of families: Set 1 phrase speech delay (PSD) positive and PSD negative, and Set 2 PSD positive and PSD negative, and also incorporating classification of parents based on whether or not they reported a history of a developmental language impairment (see Methods). The sex-averaged genetic map provided by Marshfield was used to plot cM location.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Language characteristics of CLSA families

Samples	PSD + families *n* (%)	PSD − families *n* (%)
Set 1	40 (70)	17 (30)
Set 2	36 (63)	21 (37)
Language impaired Parents (%)	(28)	(28)

CLSA, Collaborative Linkage Study of Autism; PPL, posterior probability of linkage; PSD +, phrase speech delay positive; PSD −, phrase speech delay negative.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 2

CLSA peak posterior probability of linkages

Two-point PPL				Multipoint PPL		
Chrom	**Marker**	**cM**	**PPL**	**Chrom**	**cM**	**PPL**
1	GATA43A04	164	0.33	1	173	0.12
13	GATA23C03P	9	0.10	—	—	—
13	ATA5A09N	20	0.39	13	21	0.16
13	13AL159	56	0.29	13	56	0.04
16	GATA71H05	51	0.14	—	—	—
16	ATA55A11	64	0.95	16	63	0.36
16	GATA22F09	72	0.11	—	—	—
17	ATA78D02N	32	0.15	17	32	0.06
23	GATA124E07	40	0.10	23	40	0.11

Two-point PPLs are from the CLSA Set 1 genome-wide screen. Multipoint PPLs also include the CLSA Set 2 families, and additional markers genotyped in the two-point regions of interest. Centimorgan (cM) locations are from the Marshfield sex-averaged genetic map.

CLSA, Collaborative Linkage Study of Autism; PPL, posterior probability of linkage.