Am. J. Hum. Genet. 76:688–695, 2005

Report

Effects of Updating Linkage Evidence across Subsets of Data: Reanalysis of the Autism Genetic Resource Exchange Data Set

Christopher W. Bartlett,[1,3] Rhinda Goedken,[1] and Veronica J. Vieland[1,2,4]

[1]Center for Statistical Genetics Research and [2]Program in Public Health Genetics, College of Public Health, and Departments of [3]Internal Medicine and [4]Psychiatry, Carver College of Medicine, University of Iowa, Iowa City

Results of autism linkage studies have been difficult to interpret across research groups, prompting the use of ever-increasing sample sizes to increase power. However, increasing sample size by pooling disparate collections for a single analysis may, in fact, not increase power in the face of genetic heterogeneity. Here, we applied the posterior probability of linkage (PPL), a method designed specifically to analyze multiple heterogeneous data sets, to the Autism Genetic Resource Exchange collection of families by analyzing six clinically defined subsets of the data and updating the PPL sequentially over the subsets. Our results indicate a substantial probability of linkage to chromosome 1, which had been previously overlooked; our findings also provide a further characterization of the possible parent-of-origin effects at the 17q11 locus that were previously described in this sample. This analysis illustrates that the way in which heterogeneity is addressed in linkage analysis can dramatically affect the overall conclusions of a linkage study.

Autism spectrum disorder (ASD [MIM 209850]) is a relatively rare pervasive developmental disorder that presents with abnormal development of language and social responses/initiation and is also characterized by stereotypic behavioral repertoires (Fombonne 1999; Folstein and Rosen-Sheidley 2001). Autism is presumed to have a genetic basis—as suggested, for example, by twin studies—and numerous groups have undertaken the search for susceptibility genes (International Molecular Genetic Study of Autism Consortium 1998; Paris Autism Research International Sibpair Study 1999; Risch et al. 1999; Collaborative Linkage Study of Autism 2001; International Molecular Genetic Study of Autism Consortium 2001; Liu et al. 2001; Alarcón et al. 2002; Shao et al. 2002*b*; Yonan et al. 2003). However, the results of these genome scans have not yielded consistent locations for autism susceptibility loci, with results typically shifting and becoming less clear as more families are added to each individual collection (Wassink et al. 2004). Across family collections, there has been some concordance for findings on 2q and 7q, but there is still no clear and convincing evidence of any specific linkage location.

Locus heterogeneity may play a substantial role in the problems encountered in autism linkage studies. Strategies for mitigating the effects of locus heterogeneity include the use of phenotypic characteristics to define more homogeneous subsets of the data and the use of statistical techniques that specifically allow for subgroup differences. Many groups have tried subsetting their family collections on the basis of phenotypic characteristics derived from the Autism Diagnostic Interview–Revised (ADI-R) (Lord et al. 1994), such as delay in acquisition of phrase speech (Bradford et al. 2001; Buxbaum et al. 2001; Shao et al. 2002*a*) or other traits (Nurmi et al. 2003; Shao et al. 2003), as well as, more recently, sex of the affected pairs (Stone et al. 2004).

Here, we reanalyze the data presented by Yonan et al. (2003), which are in the public domain as part of the Autism Genetic Resource Exchange (AGRE) (Geschwind et al. 2001). We compare the original results of Yonan et al. (2003) with what we obtain using an alternative data-analysis method that has been specifically designed to allow for heterogeneity within the sample, and we find that the results of a genome screen can be highly dependent upon the choice of data-analysis method in the (presumed) presence of locus heterogeneity.

The data used in our analyses comprise a subset of the

Received December 13, 2004; accepted for publication February 1, 2005; electronically published February 23, 2005.

Address for correspondence and reprints: Dr. Christopher W. Bartlett, Center for Statistical Genetics Research, 2188 Westlawn, Iowa City, IA 52242-1008. E-mail: christopher-bartlett@uiowa.edu

0002-9297/2005/7604-0016$15.00

Table 1

Number of Families, by Subset

Class	No. of Families in Subset		Total No. of Families
	PSD Positive	PSD Negative	
I	133	65	198
II	51	45	96
III	5	4	9
Total	189	114	303

AGRE families included by Yonan et al. (2003); some of these families have been included in additional publications as well (Liu et al. 2001; Alarcón et al. 2002; Yonan et al. 2003; Stone et al. 2004). Briefly, the sample consists of 303 multiplex families—primarily, affected sib pairs with genotyped parents—in which the children were ascertained for ASDs, including autism, Asperger syndrome, and pervasive developmental disorder (PDD). Yonan et al. (2003) used 345 families. Of these, 11 families tested positive for fragile X syndrome in March 2004, subsequent to the publication of the study by Yonan et al. (2003), and are omitted from our analyses; in addition, 31 families were trios with no linkage information and have also been omitted here. (Thus, the actual difference between the sample used by Yonan et al. [2003] and the one used in our study should be just the 11 families with fragile X syndrome that were included in their study but omitted in ours.) ADI-R data were available for all affected subjects (Lord et al. 1994).

We used all of the genotypes from AGRE that were publicly available at the time of analysis: the 408 microsatellites reported by Yonan et al. (2003), including markers at ~10-cM resolution from the Marshfield genome screening set, version 8 (see the Center for Medical Genetics Web site), as well as 73 additional markers that were used to follow up on results from previous analyses of these data (Liu et al. 2001; Alarcón et al. 2002; Yonan et al. 2003). Genotypic data were read into our Oracle database from J. A. Badner's "hypercleaned" data file, which is available on the AGRE Web site. Any discrepancies between the diagnosis in Badner's file and the AGRE diagnosis were resolved by substituting the AGRE diagnosis. Allele frequencies were estimated by allele counting in all founders.

Yonan et al. (2003) analyzed these data as a single group—that is, without explicitly considering possible group differences. They used the multipoint maximum LOD score (MLS) (Risch 1990) to analyze the data and an approximate "model-free" LOD (Göring and Terwilliger 2000) for two-point analysis. They found their maximum MLS at position 17q11 (MLS = 2.83) and identified other "suggestive" MLSs on chromosomes 5, 11, 4, and 8 (MLSs of 2.54, 2.24, 1.72, and 1.60, respectively). (Although there may be other small discrepancies between Badner's files and the data used in the original report by Yonan et al. [2003], we have verified that Badner's files produce essentially the same MLS results, with a maximum MLS of 2.7 on 17q11; in addition, the rank order of these other signals was unchanged, with the exception that the chromosome 5 MLS is only 1.7 and drops in rank from number 2 to number 4.)

For our reanalysis of these data, we divided the sample into six subsets, which are defined a priori as follows: if at least two siblings met the International Classifi-

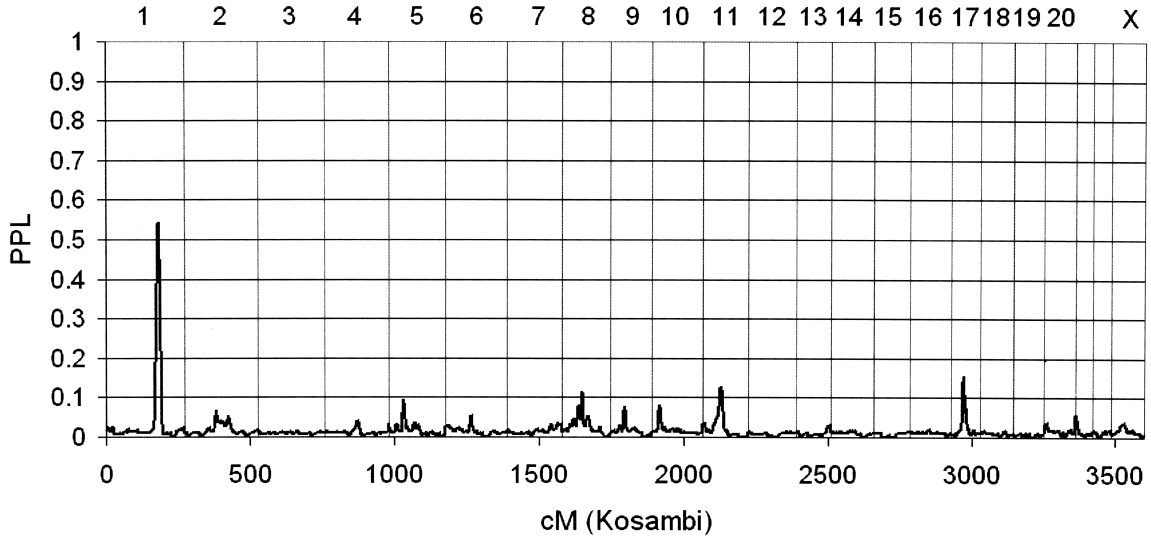

Figure 1 Summary of PPL analysis across the genome. Chromosomes are listed along the top border of the graph. PPL results shown here have been updated across the six data subsets. This graph illustrates the low signal-to-noise ratio for the peak finding.

Table 2

Results of Present Analysis, Compared with Those of Yonan et al. (2003)

	Multipoint Analysis				Two-Point Analysis			
	PPL		Yonan et al. 2003		PPL		Yonan et al. 2003	
Chromosome	%	Position (cM)	MLS	Position (cM)	%	Position (cM)	LOD	Position (cM)
1	55	183	<.5[a]	180[a]	34	170	NA	NA
17	15	45	2.8	52	8	48	1.2	48
11	13	63	2.2	45	5	54	.5	54
8	12	72	1.6	131	6	78	.8	135
5	9	53	2.5	58	12	57	1.4	57
4	4	114	1.7	94	4	107	1.7	101

Note.—Only regions with PPLs >10%—and those specifically listed in the article by Yonan et al. (2003)—are shown. By convention, PPLs >3% are rounded to the nearest whole number, whereas PPLs <3% are rounded to two significant digits. NA = data not available.

[a] MLS and position for this location were estimated from the graphs in the article by Yonan et al. (2003).

cation of Disease 10 (ICD-10) algorithm for autism, we assigned the family to class I; if only one child met criteria for ICD-10 autism (and at least one additional child met criteria for Asperger syndrome or PDD), we assigned the family to class II. The remaining families, which contained no cases of ICD-10 autism (but at least two cases of Asperger syndrome or PDD), were assigned to class III. The rationale for this division was to achieve greater *clinical* homogeneity within subgroups. Each of these classes was further broken down on the basis of whether or not at least two affected siblings (with autism, Asperger syndrome, or PDD) presented with a phrase speech delay (PSD) of >36 mo (groups with at least two affected siblings with a PSD of >36 mo are referred to as "PSD positive"; groups without at least two affected siblings with a PSD of >36 mo are referred to as "PSD negative"). This division was based on previous findings from independent groups supporting likely genetic differences between families with and without multiplex PSD (Bradford et al. 2001; Buxbaum et al. 2001; Shao et al. 2002*a*; Vieland et al. 2003). Note that all of the families we included were also included in the Yonan et al. (2003) analysis. We did not add or drop families on the basis of clinical status; we merely *classified* them with respect to clinical status in the analyses. Table 1 shows the sample-size breakdown of the AGRE families by subset. A complete list (by subset) of the families used in our analysis is available in appendix A (online only).

Table 3

Permutation Results at Linked and Unlinked Loci

	PPL (%) for	
Locus	Random Subsets [SD]	Pooled Data
1q23-24	2.1 [2.8]	1.7
Unlinked	.3 [.4]	1.2

Note.—Families were randomly permuted into six subsets, corresponding to the observed clinical subset sizes, and the PPL was sequentially updated across these random subsets. One thousand permutations were performed at each locus, and the average (SD) was calculated across permutations. PPLs for pooled data, shown for comparison, were computed by treating all six subsets as a single data set (i.e., without sequential updating across subgroups).

We then analyzed the data by use of the posterior probability of linkage (PPL), which is specifically designed to allow for differences between subgroups (Vieland 1998). The PPL is parameterized in terms of an approximating single-locus model, allowing for heterogeneity under the admixture model (Smith 1963). All parameters of this model (gene frequency, three penetrances, and the admixture parameter) are integrated out of the likelihood, independently for each subset, and the resulting marginal posterior density in the recombination fraction (two-point) or genomic location (multipoint) is sequentially updated across the subsets (Vieland et al. 2001; Vieland and Logue 2002; Logue et al. 2003; Logue and Vieland 2004). In this way, the PPL allows for heterogeneity within subsets, as well as for differences across subsets, while accumulating the total evidence for and against linkage based on all families in a mathematically rigorous way. Because the PPL does not involve maximum-likelihood estimation or maximization of linkage statistics across subsets, there is no inflation of the PPL inherent in either updating across data subsets or subsetting on the basis of genetically irrelevant factors (see table 3 for an illustration). However, we have shown that, in the presence of heterogeneity within and across subsets, sequential updating across relevant clin-

Table 4

PPLs, by Subset, for 1q23-24 and 17q11

Class and Subset	PPL (%) for Locus	
	1q23-24	17q11
I:		
PSD Positive	1.1	2.5
PSD Negative	2.2	1.8
II:		
PSD Positive	1.8	2.1
PSD Negative	76	13
III:		
PSD Positive	2.0	2.0
PSD Negative	2.0	2.1
All subgroups (sequentially updated)	55	15

ical features can improve ability to find linkage (Wang et al. 1999; Huang and Vieland 2001; Vieland et al. 2001; Bartlett et al. 2002, 2004; Logue et al. 2003). The PPL is on the probability scale, with values >2% (the prior probability of linkage) indicating evidence in favor of linkage and values <2% indicating evidence against linkage.

Figure 1 shows multipoint PPLs across the genome. Overall, 83% of the genome yielded PPLs <2%; interestingly, this includes most of chromosome 7 (72%), with the 7q34-qter interval being the largest contiguous region with PPLs not <2%. The largest PPL is 55%, located at 1q23-24. Table 2 compares our results with the MLS and LOD results of Yonan et al. (2003), in order of decreasing PPL. We note that the rank order of scores differs between the PPL and both the MLS and the LOD analyses.

This difference in rank order is not due to the use of the PPL, per se, but rather to the manner in which the PPL uses the clinically defined subsets. For example, if we pool all families into a single data set, 17q11 provides the largest PPL (5%) genomewide, just as it produced the largest MLS genomewide in Yonan et al. (2003). Additionally, at both the 1q23-24 and 17q11 loci, the "pooled" results are lower than the sequentially updated results. Thus, we draw a substantially different overall conclusion from the genome screen depending on whether we treat the families as a single homogeneous group (as the MLS and the "pooled" PPL both implicitly do) or whether we specifically allow for differences across subgroups.

As stated above, one feature of the sequential updating used by the PPL is that subsetting on genetically irrelevant characteristics has (on average) no impact on the final result, as compared with "pooled" analysis, which is conducted on the data set as a single group (i.e., there is no inherent inflationary effect of subdividing the sample). This applies even in the present case, in which the sizes of the subsets vary appreciably. To illustrate this, we randomly permuted families into six subsets, of the same sizes as the observed clinical subsets, and recomputed the sequentially updated PPL across the random subsets. This procedure was repeated 1,000 times at each of two loci: at the peak location on chromosome 1 (183 cM) and at an apparently unlinked locus (chromosome 12 at 100 cM) that was previously unnoted in the autism literature and for which the observed sequentially updated PPL was 1.5% (evidence against linkage). Table 3 shows the results. In both cases, randomly subsetting the data produces, on average, a PPL very close to the observed PPL obtained by simply pooling all the families for a single analysis. Thus, random subsetting (even into unequally sized subsets) has no inflationary impact relative to the "pooled" PPL, either at the (apparently) linked locus or at the unlinked locus. By contrast, the difference between the permutation-based PPL and the PPL obtained when we sequentially update across the *clinically defined* subsets at 1q23-24 (PPL = 55%) is striking, strongly suggesting that the predefined clinical subsetting criteria have some genetic relevance.

It is also possible to examine the PPLs individually in the separate subsets. Table 4 shows subset-specific results for 1q23-24 and 17q11. On chromosome 1, the class II PSD-negative group contributes virtually the entire linkage signal (PPL = 76% in this subset), with the class I PSD-negative group giving very slight evidence in favor of linkage (PPL = 2.2%) and the remaining subsets contributing either no information (PPL = 2%) or actually giving evidence against linkage. This pattern is consistent with (although hardly proof of) the possibility that this locus is linked in the PSD-negative, but not the PSD-positive, groups, regardless of clinical class. On 17q11, the pattern is not so clear (and, indeed, the overall PPL is considerably smaller), with multiple subgroups con-

Table 5

Sequentially Updated PPL Results over Different Subsets for 17q11

Locus (Position)	Sequentially Updated PPL (%) for		
	Pooled Data	MO-FC Pairs	MO-FC Pairs, Classes, and PSD Subsets
17q11 (44 cM)	4.2	4.9	7.4

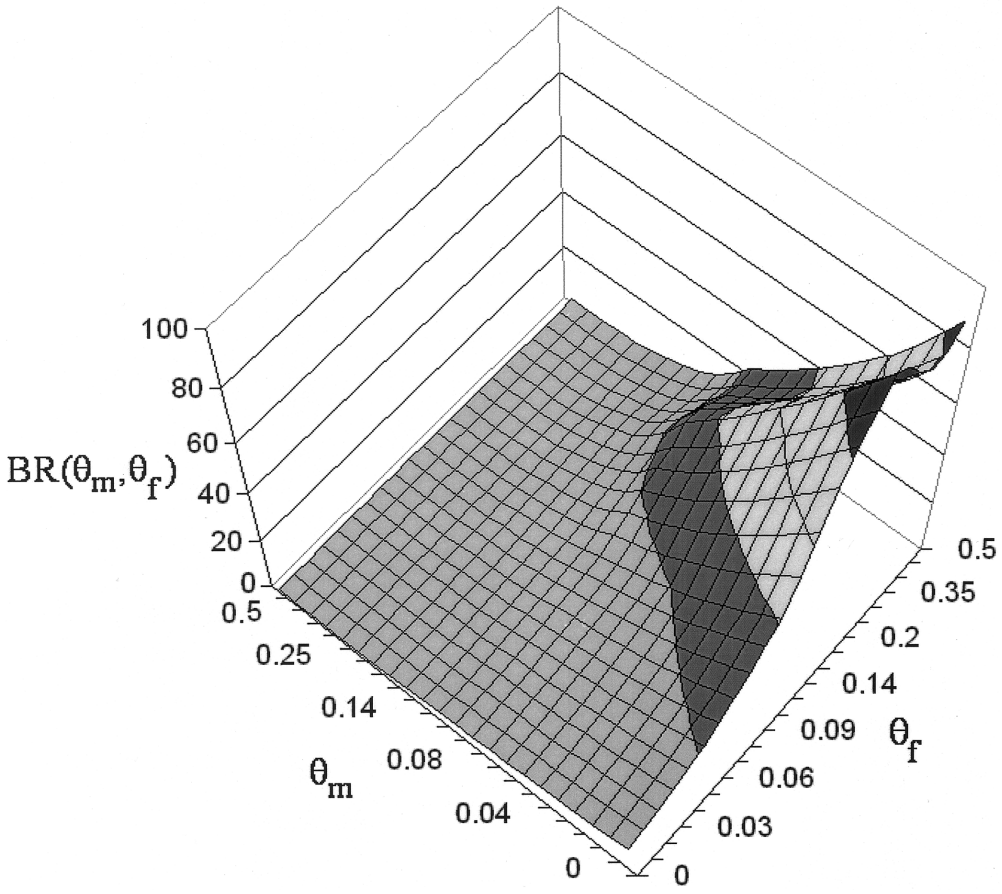

Figure 2 Plot of Bayes ratios (BRs) at *D17S1871* after sequential updating, defined as $BR(\theta_m,\theta_f) = \int 10^{HLOD(\theta_m,\theta_f,t)}\pi(\mathbf{t})d\mathbf{t}$, where θ_m and θ_f are the male and female recombination fractions, respectively; HLOD is the ordinary heterogeneity LOD score (Smith 1963; Ott 1983); **t** is the vector of trait parameters including α; and $\pi(\mathbf{t})$ is the prior on **t.** Gradations indicate changes in the BR between increments on the *Z*-axis. The posterior mode of the BR occurs at 0 for males and at 0.25 for females.

tributing very slight evidence in favor of linkage and only one group (class I, PSD negative) showing evidence against linkage. Of note, however, is the finding that the class II PSD-negative group also has the single largest PPL (13%) at this locus.

We also sought to further characterize the 17q11 locus. In a subset of these same families, Stone et al. (2004) found that considering families containing only male children with ASD resulted in an increase in the MLS (4.3, compared with 3.2 in all families), despite a sample-size reduction of 42% (n = 148, compared with 257). (Note that these 257 families represent a subset of the original sample used by Yonan et al. [2003].) This result now appears to have been replicated in an independent sample as well (Cantor-Chiu et al. 2004). Table 5 shows results of sequentially updating across male-only (MO) pairs and female-containing (FC) pairs. We found that MO-FC subsetting increased the PPL compared with "pooled" analysis but actually decreased the PPL compared with the use of both the MO-FC pairs and our original additional subsetting criteria. However, in all cases, the PPLs are not very high, making definitive interpretation moot. It is of interest that the MO pairs were relatively evenly distributed across our clinical subsets (ranging from 50% in the class III PSD-negative group to 63% in the class II PSD-positive group). (The class II PSD-negative group, which had the highest PPL, contained only 51% MO pairs.)

The 17q11 locus has also been noted to have excess observed paternal, but not maternal, sharing in the AGRE data (J. Vincent and A. Paterson, personal communication), which is consistent with observations in another autism sample (International Molecular Genetic Study of Autism Consortium 2001). We therefore repeated our original PPL analyses at this locus by use of a two-point analysis, allowing for separate male and female recombination fractions (Ott 1976); a multipoint sex-specific PPL is not yet implemented. Sex-specific differences in recombination rates can indicate imprinting or other parent-of-origin effects in nuclear families and have been

shown to be symptomatic of imprinted data even in larger pedigrees (Smalley 1993; Greenberg et al. 2000; Ludington 2000; Ludington et al. 2000; Feenstra et al. 2004). The two-point PPLs were small to moderate at both *D17S1824* (PPL = 5%) and *D17S1871* (PPL = 19%). However, we did find some suggestion of sex differences (fig. 2). In our experience, the large observed difference in male and female recombination shown in figure 2 would not be indicative of a sex-specific difference in recombination, per se; note that the female-to-male map ratio at this particular marker is in the range of 1.1 to 9.4 (using the closest flanking markers or flanking markers 1.4 cM away, respectively) from a combined linkage-physical map (Kong et al. 2004). Thus, we concur with the previous conclusions that this could indicate a parent-of-origin effect such as imprinting.

Although our primary purpose was to illustrate that differing approaches to handling heterogeneity could lead to very different genomewide results, we have also uncovered strong evidence of linkage in the interval 1q23-24 that has not been reported elsewhere for autism. This interval contains several plausible candidate genes for autism susceptibility—all highly expressed in the brain—including aldehyde dehydrogenase 9 family, member A1 (*ALDH9A1*), and regulator of G-protein signaling 4 and 5 (*RGS4* and *RGS5*). The region also holds an interesting candidate gene for schizophrenia, the carboxyl-terminal PDZ ligand of neuronal nitric oxide synthase (*CAPON*) (Brzustowicz et al. 2004). Although schizophrenia and autism are conceptualized as distinct disorders, there is evidence of increased risk of one disease in individuals with the other disease (Nylander and Gillberg 2001; Stahlberg et al. 2004; see also the study by Kay et al. [1987]).

Overall, our analyses illustrate that the way in which we allow for potential subgroup differences when analyzing genome-screen data can have a substantial impact on our conclusions. Heterogeneity can complicate detection of linkage within any given data set and can obscure findings that are based on combined analysis across data sets. Previous work had demonstrated potential loss of power—even when sample sizes are increased—if proper allowances for heterogeneity across subgroups are not made (Huang and Vieland 2001; Vieland et al. 2001). The autism analyses shown here suggest that failure to adequately allow for subgroup differences can result not only in loss of power but also in substantial changes in the rank order of findings across the genome. These observations suggest the need for caution in the analysis of data from large multisite collaborations when the disease is suspected to be heterogeneous.

Acknowledgments

We gratefully acknowledge the resources provided by the AGRE Consortium and the participating AGRE families. Public access to such data sets is an invaluable tool, not only in the search for genes but also for the development of accurate and efficient statistical methods for analyzing complex genetic data. The AGRE is a program of the Cure Autism Now Foundation. We also gratefully acknowledge support for C.W.B. from grant T32 HL07638 and support for V.J.V. and R.G. from grant R01 MH052841 (to V.J.V.). Additionally, we thank Andrew Paterson and John Vincent for helpful discussions of earlier versions.

The AGRE Consortium includes Dan Geschwind, M.D., Ph.D., University of California–Los Angeles (UCLA), Los Angeles; Maja Bucan, Ph.D., University of Pennsylvania, Philadelphia; W. Ted Brown, M.D., Ph.D., F.A.C.M.G., New York State Institute for Basic Research in Developmental Disabilities, Long Island; Joseph Buxbaum, Ph.D., Mt. Sinai School of Medicine, New York; Rita M. Cantor, Ph.D., UCLA School of Medicine, Los Angeles; John N. Constantino, M.D., Washington University School of Medicine, St. Louis; T. Conrad Gilliam, Ph.D., University of Chicago, Chicago; Clara Lajonchere, Ph.D., Cure Autism Now Foundation, Los Angeles; David H. Ledbetter, Ph.D., Emory University, Atlanta; Christa Lese-Martin, Ph.D., Emory University, Atlanta; Janet Miller, J.D., Ph.D., Cure Autism Now Foundation, Los Angeles; Stanley F. Nelson, M.D., UCLA School of Medicine, Los Angeles; Gerard D. Schellenberg, Ph.D., University of Washington, Seattle; Carol A. Samango-Sprouse, Ed.D., George Washington University, Washington, D.C.; Sarah Spence, M.D., Ph.D., UCLA, Los Angeles; and Rudolph E. Tanzi, Ph.D., Massachusetts General Hospital, Boston.

Electronic-Database Information

The URLs for data presented herein are as follows:

Autism Genetic Resource Exchange (AGRE), http://www.agre.org/ (for full diagnostic protocol for AGRE families)
Center for Medical Genetics, Marshfield Medical Research Foundation, http://research.marshfieldclinic.org/genetics/
Cure Autism Now Foundation, http://www.canfoundation.org/
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for ASD)

References

Alarcón M, Cantor RM, Liu J, Gilliam TC, the Autism Genetic Resource Exchange Consortium, Geschwind DH (2002) Evidence for a language quantitative trait locus on chromosome 7q in multiplex autism families. Am J Hum Genet 70:60–71

Bartlett CW, Flax JF, Logue MW, Smith BJ, Vieland VJ, Tallal P, Brzustowicz LM (2004) 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 57:10–20

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

Bradford Y, Haines J, Hutcheson H, Gardiner M, Braun T, Sheffield V, Cassavant T, Huang W, Wang K, Vieland V, Folstein

S, Santangelo S, Piven J (2001) Incorporating language phenotypes strengthens evidence of linkage to autism. Am J Med Genet 105:539–547

Brzustowicz LM, Simone J, Mohseni P, Hayter JE, Hodgkinson KA, Chow EW, Bassett AS (2004) Linkage disequilibrium mapping of schizophrenia susceptibility to the CAPON region of chromosome 1q22. Am J Hum Genet 74:1057–1063

Buxbaum JD, Silverman JM, Smith CJ, Kilifarski M, Reichert J, Hollander E, Lawlor BA, Fitzgerald M, Greenberg DA, Davis KL (2001) Evidence for a susceptibility gene for autism on chromosome 2 and for genetic heterogeneity. Am J Hum Genet 68:1514–1520 (erratum 69:470)

Cantor-Chiu RM, Kono N, Alarcón M, Geschwind DH (2004) Genome scan replicates linkage of autism to 17q11. Paper presented at the American Society of Human Genetics annual meeting. Toronto, October 26–30

Collaborative Linkage Study of Autism (2001) An autosomal genomic screen for autism. Am J Med Genet 105:609–615

Feenstra B, Greenberg DA, Hodge SE (2004) Using lod scores to detect sex differences in male-female recombination fractions. Hum Hered 57:100–108

Folstein SE, Rosen-Sheidley B (2001) Genetics of autism: complex aetiology for a heterogeneous disorder. Nat Rev Genet 2:943–955

Fombonne E (1999) The epidemiology of autism: a review. Psychol Med 29:769–786

Geschwind DH, Sowinski J, Lord C, Iversen P, Shestack J, Jones P, Ducat L, Spence SJ (2001) The autism genetic resource exchange: a resource for the study of autism and related neuropsychiatric conditions. Am J Hum Genet 69:463–466

Göring HH, Terwilliger JD (2000) Linkage analysis in the presence of errors IV: joint pseudomarker analysis of linkage and/or linkage disequilibrium on a mixture of pedigrees and singletons when the mode of inheritance cannot be accurately specified. Am J Hum Genet 66:1310–1327

Greenberg DA, Durner M, Keddache M, Shinnar S, Resor SR, Moshe SL, Rosenbaum D, Cohen J, Harden C, Kang H, Wallace S, Luciano D, Ballaban-Gil K, Tomasini L, Zhou G, Klotz I, Dicker E (2000) Reproducibility and complications in gene searches: linkage on chromosome 6, heterogeneity, association, and maternal inheritance in juvenile myoclonic epilepsy. Am J Hum Genet 66:508–516

Huang J, Vieland VJ (2001) 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 51:217–225

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

International Molecular Genetic Study of Autism Consortium (2001) A genomewide screen for autism: strong evidence for linkage to chromosomes 2q, 7q, and 16p. Am J Hum Genet 69:570–581

Kay SR, Fiszbein A, Opler LA (1987) The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr Bull 13:261–276

Kong X, Murphy K, Raj T, He C, White PS, Matise TC (2004) A combined linkage-physical map of the human genome. Am J Hum Genet 75:1143–1148

Liu J, Nyholt DR, Magnussen P, Parano E, Pavone P, Geschwind D, Lord C, Iversen P, Hoh J, the Autism Genetic Resource Exchange, Ott J, Gilliam TC (2001) A genomewide screen for autism susceptibility loci. Am J Hum Genet 69: 327–340

Logue MW, Vieland VJ (2004) A new method for computing the multipoint posterior probability of linkage. Hum Hered 57:90–99

Logue MW, Vieland VJ, Goedken RJ, Crowe RR (2003) Bayesian analysis of a previously published genome screen for panic disorder reveals new and compelling evidence for linkage to chromosome 7. Am J Med Genet B Neuropsychiatr Genet 121:95–99

Lord C, Rutter M, Le Couteur A (1994) Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. J Autism Dev Disord 24:659–685

Ludington EA (2000) Sex-specific recombination in linkage analysis. University of Iowa, Iowa City, pp 35–65

Ludington E, Vieland VJ, Huang J (2000) Using ordinary linkage analysis to detect imprinting in general pedigrees. Am J Hum Genet Suppl 67:1766

Nurmi EL, Dowd M, Tadevosyan-Leyfer O, Haines JL, Folstein SE, Sutcliffe JS (2003) Exploratory subsetting of autism families based on savant skills improves evidence of genetic linkage to 15q11-q13. J Am Acad Child Adolesc Psychiatry 42:856–863

Nylander L, Gillberg C (2001) Screening for autism spectrum disorders in adult psychiatric out-patients: a preliminary report. Acta Psychiatr Scand 103:428–434

Ott J (1976) A computer program for linkage analysis of general human pedigrees. Am J Hum Genet 28:528–529

Ott J (1983) Linkage analysis and family classification under heterogeneity. Ann Hum Genet 47:311–320

Paris Autism Research International Sibpair Study (1999) Genome-wide scan for autism susceptibility genes. Hum Mol Genet 8:805–812

Risch N (1990) Linkage strategies for genetically complex traits. II. The power of affected relative pairs. Am J Hum Genet 46: 229–241

Risch N, Spiker D, Lotspeich L, Nouri N, Hinds D, Hallmayer J, Kalaydjieva L, et al (1999) A genomic screen of autism: evidence for a multilocus etiology. Am J Hum Genet 65:493–507

Shao Y, Cuccaro ML, Hauser ER, Raiford KL, Menold MM, Wolpert CM, Ravan SA, Elston L, Decena K, Donnelly SL, Abramson RK, Wright HH, DeLong GR, Gilbert JR, Pericak-Vance MA (2003) Fine mapping of autistic disorder to chromosome 15q11-q13 by use of phenotypic subtypes. Am J Hum Genet 72:539–548

Shao Y, Raiford KL, Wolpert CM, Cope HA, Ravan SA, Ashley-Koch AA, Abramson RK, Wright HH, DeLong RG, Gilbert JR, Cuccaro ML, Pericak-Vance MA (2002*a*) Phenotypic homogeneity provides increased support for linkage on chromosome 2 in autistic disorder. Am J Hum Genet 70: 1058–1061

Shao Y, Wolpert CM, Raiford KL, Menold MM, Donnelly SL, Ravan SA, Bass MP, McClain C, von Wendt L, Vance JM, Abramson RH, Wright HH, Ashley-Koch A, Gilbert JR, DeLong RG, Cuccaro ML, Pericak-Vance MA (2002*b*) Ge-

nomic screen and follow-up analysis for autistic disorder. Am J Med Genet 114:99–105

Smalley SL (1993) Sex-specific recombination frequencies: a consequence of imprinting? Am J Hum Genet 52:210–212

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

Stahlberg O, Soderstrom H, Rastam M, Gillberg C (2004) Bipolar disorder, schizophrenia, and other psychotic disorders in adults with childhood onset AD/HD and/or autism spectrum disorders. J Neural Transm 111:891–902

Stone JL, Merriman B, Cantor RM, Yonan AL, Gilliam TC, Geschwind DH, Nelson SF (2004) Evidence for sex-specific risk alleles in autism spectrum disorder. Am J Hum Genet 75:1117–1123

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

Vieland VJ, Logue M (2002) HLODs, trait models, and ascertainment: implications of admixture for parameter estimation and linkage detection. Hum Hered 53:23–35

Vieland VJ, Sheffield V, Wassink T, Beck J, Goedken R, Childress D, Piven J (2003) A new genome screen for autism based on the posterior probability of linkage and incorporating language-based phenotypes finds evidence of linkage to several genomic locations, each supported by independent sources of information. Am J Hum Genet Suppl 73:174

Vieland VJ, Wang K, Huang J (2001) Power to detect linkage based on multiple sets of data in the presence of locus heterogeneity: comparative evaluation of model-based linkage methods for affected sib pair data. Hum Hered 51:199–208

Wang K, Vieland VJ, Huang J (1999) A Bayesian approach to replication of linkage studies. Genet Epidemiol 17:S749–S754

Wassink TH, Brzustowicz LM, Bartlett CW, Szatmari P (2004) The search for autism genes. Ment Retard Dev Disabil Res Rev 10:272–283

Yonan AL, Alarcón M, Cheng R, Magnusson PK, Spence SJ, Palmer AA, Grunn A, Juo SH, Terwilliger JD, Liu J, Cantor RM, Geschwind DH, Gilliam TC (2003) A genomewide screen of 345 families for autism-susceptibility loci. Am J Hum Genet 73:886–897