 Human Molecular Genetics, 1998, Vol. 7, No. 3 571–578

A full genome screen for autism with evidence for linkage to a region on chromosome 7q

International Molecular Genetic Study of Autism Consortium+

Received November 14, 1997; Revised and Accepted December 19, 1998

Autism is characterized by impairments in reciprocal social interaction and communication, and restricted and stereotyped patterns of interests and activities. Developmental difficulties are apparent before 3 years of age and there is evidence for strong genetic influences most likely involving more than one susceptibility gene. A two-stage genome search for susceptibility loci in autism was performed on 87 affected sib pairs plus 12 non-sib affected relative-pairs, from a total of 99 families identified by an international consortium. Regions on six chromosomes (4, 7, 10, 16, 19 and 22) were identified which generated a multipoint maximum lod score (MLS) > 1. A region on chromosome 7q was the most significant with an MLS of 3.55 near markers D7S530 and D7S684 in the subset of 56 UK affected sib-pair families, and an MLS of 2.53 in all 87 affected sib-pair families. An area on chromosome 16p near the telomere was the next most significant, with an MLS of 1.97 in the UK families, and 1.51 in all families. These results are an important step towards identifying genes predisposing to autism; establishing their general applicability requires further study.

INTRODUCTION

Autism—the prototypical Pervasive Development Disorder (PDD)—has a population prevalence of ~4/10 000 and is characterized by impairments in reciprocal social interaction and communication, restricted and stereotyped patterns of interests and activities, and the presence of developmental abnormalities by 3 years of age (1–3). A strong genetic component in autism is indicated by an increased concordance rate in monozygotic versus dizygotic twins (4,5) and a risk to siblings of idiopathic cases which is 75 times greater than the general population prevalence [λ_s = 75: ratio of 3% sibling risk divided by the population prevalence of 0.0004 (6)]. The behavioural phenotype can extend to other PDDs (4,6) and the genetic liability is probably mediated by several loci (7). A full genome-wide search for susceptibility loci was undertaken since the neurobiological basis of the disorder is unknown and there are no strong candidate genes.

+(http://www.well.ox.ac.uk/~maestrin/iat.html)

UK. MRC Child Psychiatry Unit, Institute of Psychiatry, London: Anthony Bailey*, Amaia Hervas, Nicola Matthews, Sarah Palferman, Simon Wallace, Anne Aubin, Janine Michelotti, Catherine Wainhouse, Katerina Papanikolaou and Michael Rutter; Wellcome Trust Centre for Human Genetics, University of Oxford: Elena Maestrini, Angela Marlow, Daniel E. Weeks (University of Pittsburgh), Janine Lamb, Clyde Francks, Georgina Kearsley, Pat Scudder and Anthony P. Monaco*; Newcomen Centre, Guy's Hospital, London: Gillian Baird, Anthony Cox and Helen Cockerill; Fleming Nuffield Unit, Newcastle: Ann Le Couteur, Tom Berney, Hayley Cooper and Tom Kelly; Booth Hall Children's Hospital, Manchester: Jonathan Green, Jane Whittaker and Anne Gilchrist; Developmental Psychiatry Section, University of Cambridge Clinical School: Patrick Bolton and Anne Schönewald; South East Thames Regional Genetics Centre, Guy's Hospital, London: Michael Daker, Caroline Ogilvie, Zoe Docherty and Zandra Deans; European Centre for Collection of Animal Cell Cultures, Salisbury: Bryan Bolton and Ros Packer.

Germany. J. W. Goethe-Universitat, Frankfurt: Fritz Poustka, Dorothea Rühl, Gabriele Schmötzer and Sven Bölte; Division of Molecular Genome Analysis, Deutsches Krebsforschungszentrum: Sabine M Klauck, Anja Spieler and Annemarie Poustka.

The Netherlands. AZU, Department of Child and Adolescent Psychiatry, Utrecht: Herman Van Engeland, Chantal Kemner, Maretha De Jonge and Ineke den Hartog.

USA. Department of Psychiatry, University of Chicago: Catherine Lord, Edwin Cook, Bennett Leventhal, Fred Volkmar, David Pauls, Ami Klin and Susan Smalley.

France. Centre d'Etudes et de Recherches en Psychopathologie (CERRP), Université de Toulouse le Mirail, CHU de Toulouse, Toulouse: Eric Fombonne, Bernadette Rogé, Maite Tauber, Evelyne Arti-Vartayan and Jeanne Fremolle-Kruck.

Denmark. National Centre for Autism: Lennart Pederson, Demetrious Haracopos and Karen Brondum-Nielsen; Technical University of Denmark: Rodney Cotterill.

*To whom correspondence should be addressed. A. Bailey, MRC Child Psychiatry Unit, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, UK; Tel: +44 171 2527237; Fax: +44 171 2525107; Email: spjwajb@iop.bpmf.ac.uk; A. P. Monaco, Wellcome Trust Centre for Human Genetics, University of Oxford, Windmill Road, Oxford OX3 7BN, UK; Tel: + 44 1865 740019; Fax: +44 1865 742186; Email: anthony.monaco@well.ox.ac.uk

Table 1. Summary description of family samples

Breakdown of relative-pairs	Stage 1		Stage 2		Total	
	UK	All	UK	All	UK	All
Sibling pair families	25	36	30	49	55	85
Sibling trio families	0	0	1	2	1	2
Other relative-pair families	1	3	9	9	10	12
Total number of affected individuals	201					
Composition of relative-pairs						
Case Type 1/Case Type 1	43					
Case Type 1/Case Type 2	56					
Sex of relative-pairs						
Male/male	71					
Male/female	24					
Female/female	4					
Age of probands (mean ± SD)	10.5 ± 6.5					
	Mean ADI algorithm domain scores					
	Case type 1	Case type 2				
Social	24.2 ± 4.7	20.4 ± 5.7				
Communication	16.9 ± 4.3	14.9 ± 4.4				
Repetitive	6.7 ± 2.3	5.7 ± 2.9				
Vineland Adaptive Behaviour Composite Scores	46.2 ± 18.6	52.3 ± 20.8				

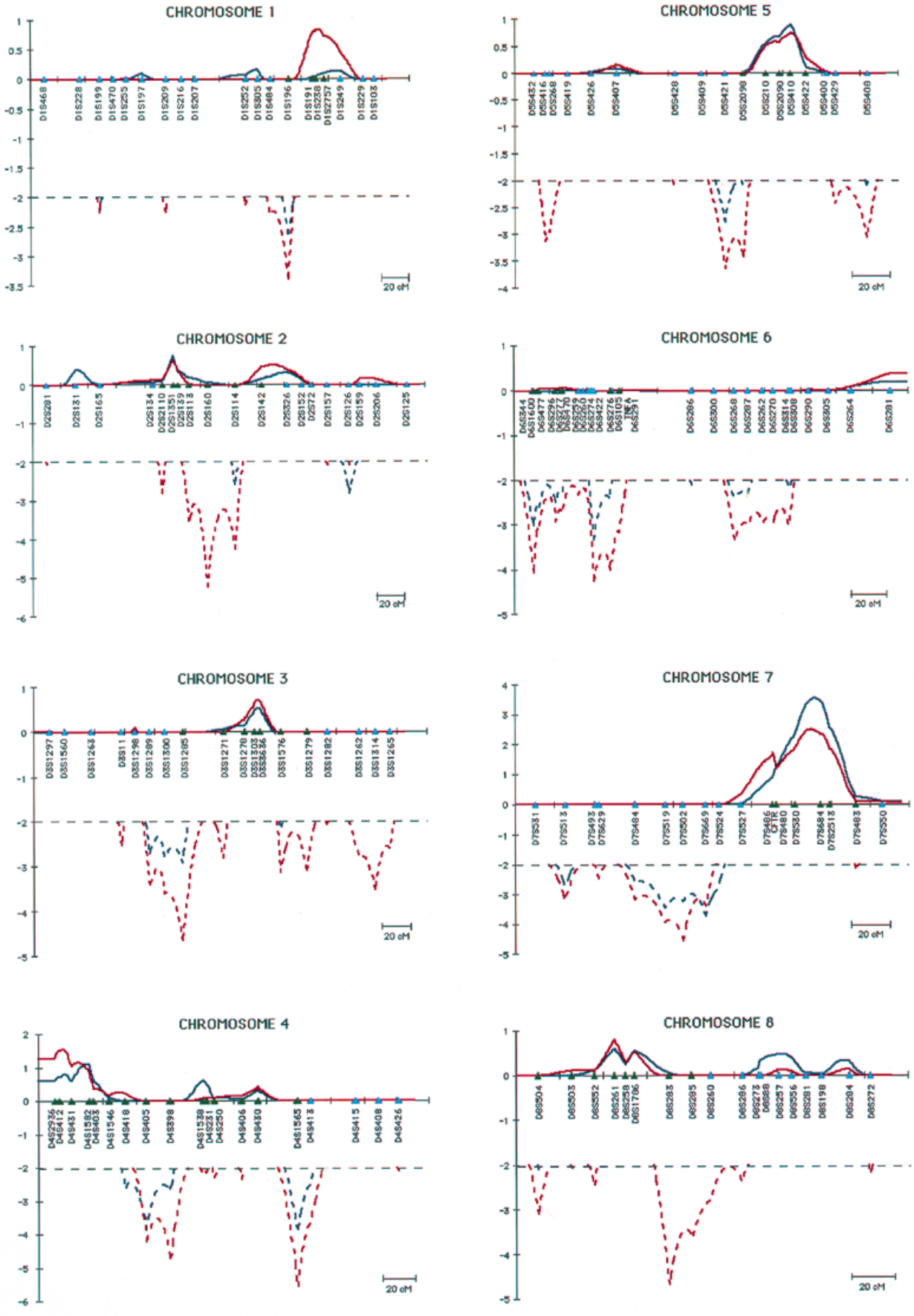
RESULTS

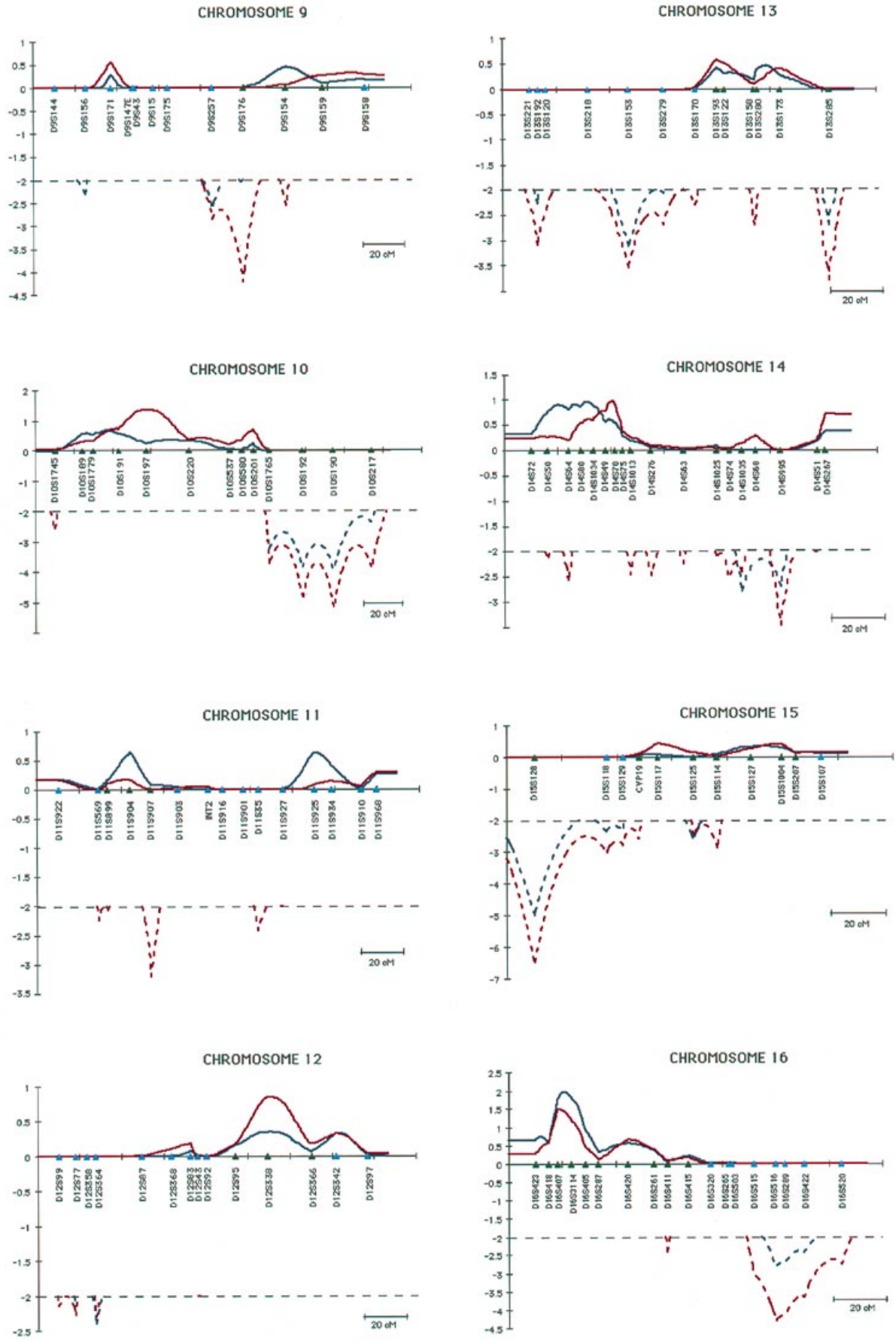
Strict criteria were applied to identify 99 families containing affected relative-pairs. At least one individual in each pair had a clinical diagnosis of autism, satisfied Autism Diagnostic Interview (ADI) algorithm criteria (8) for autism in the three behavioural domains (qualitative impairments in reciprocal social interaction; qualitative impairments in communication; restricted, repetitive and stereotyped patterns of behaviour interests and activities), showed developmental abnormalities in the first 3 years, and had a history of language delay; these individuals were designated Case Type 1. Twin and family studies of autism (4,6) indicate that the genetic liability extends to Asperger's syndrome [a disorder characterized by the same kind of abnormalities that typify autism, but in which there is no general delay in language or cognitive development (1)] and other PDDs. Because of the low base rate of autism in the population (2,3), including relative pairs in which the other proband has Asperger's syndrome or PDD is unlikely to introduce significant genetic heterogeneity. Individuals were designated as Case Type 2 if they had a clinical diagnosis of Asperger's syndrome, PDD or autism unaccompanied by language delay (even if ADI algorithm criteria for autism were met), and if they had one of these clinical diagnoses but fell 1 point below threshold on one of the behavioural domains of the ADI algorithm. Individuals fulfilling clinical and ADI algorithm criteria for autism but with apparent profound retardation were also designated Case Type 2.

Although the 99 families used for genotyping were all caucasian, 66 were from the UK, 11 from Germany, 10 from The Netherlands, five from the USA, five from France and two from Denmark. Consequently the large UK subgroup of families have been considered separately in the final analysis as they represent the largest population from a single country. Summary details of the 99 affected relative-pair families are provided in Table 1. In the UK families both cases have been karyotyped using standard methods in 62 of the 66 families and one case only in the other four. One case has been tested for Fragile X by DNA analysis in all 66 UK families. In the total sample of 99 families, at least one case has been karyotyped in 87 of the families and one case tested for Fragile X in 98 of the 99 families. No chromosomal abnormalities or cases of Fragile X were detected.

For stage 1 of the genome screen, 316 microsatellite markers were typed in 39 families, including 254 markers from the index set by Reed *et al.* (9), to which another 62 were added to fill in the larger gaps. After calculating pairwise and multipoint MLS, 38 more markers were added in regions of interest, for a total of 354 markers typed in stage 1; 62 of these markers had an MLS (pairwise or multipoint) > 0.5. In stage 2, 60 additional families were genotyped using a subset of 175 markers, that focused on the regions identified in stage 1. Although only one marker on the X chromosome (DXS996) reached an MLS > 0.5 in the stage 1 data set, markers across the whole chromosome were included in stage 2 because of the increased incidence of autism in males (2,3). Due to the small number of available families, the stage 1 and stage 2

Figure 1. Multipoint maps along each chromosome, generated by ASPEX under a model of no dominance variance. Solid lines represent maximum lod score and dashed lines represent exclusion plots for $\lambda_s = 2.5$. The results for the total data set of 87 affected sib-pair families are shown in red and the results for the subset of 56 UK affected sib-pair families are shown in blue. The position of markers used in the genome screen are shown by blue triangles if typed in stage 1 only, and by green triangles if typed in stages 1 and 2.


CHROMOSOME 1
CHROMOSOME 5
CHROMOSOME 2
CHROMOSOME 6
CHROMOSOME 3
CHROMOSOME 7
CHROMOSOME 4
CHROMOSOME 8
20 cM


CHROMOSOME 9
CHROMOSOME 13
CHROMOSOME 10
CHROMOSOME 14
CHROMOSOME 11
CHROMOSOME 15
CHROMOSOME 12
CHROMOSOME 16
20 cM

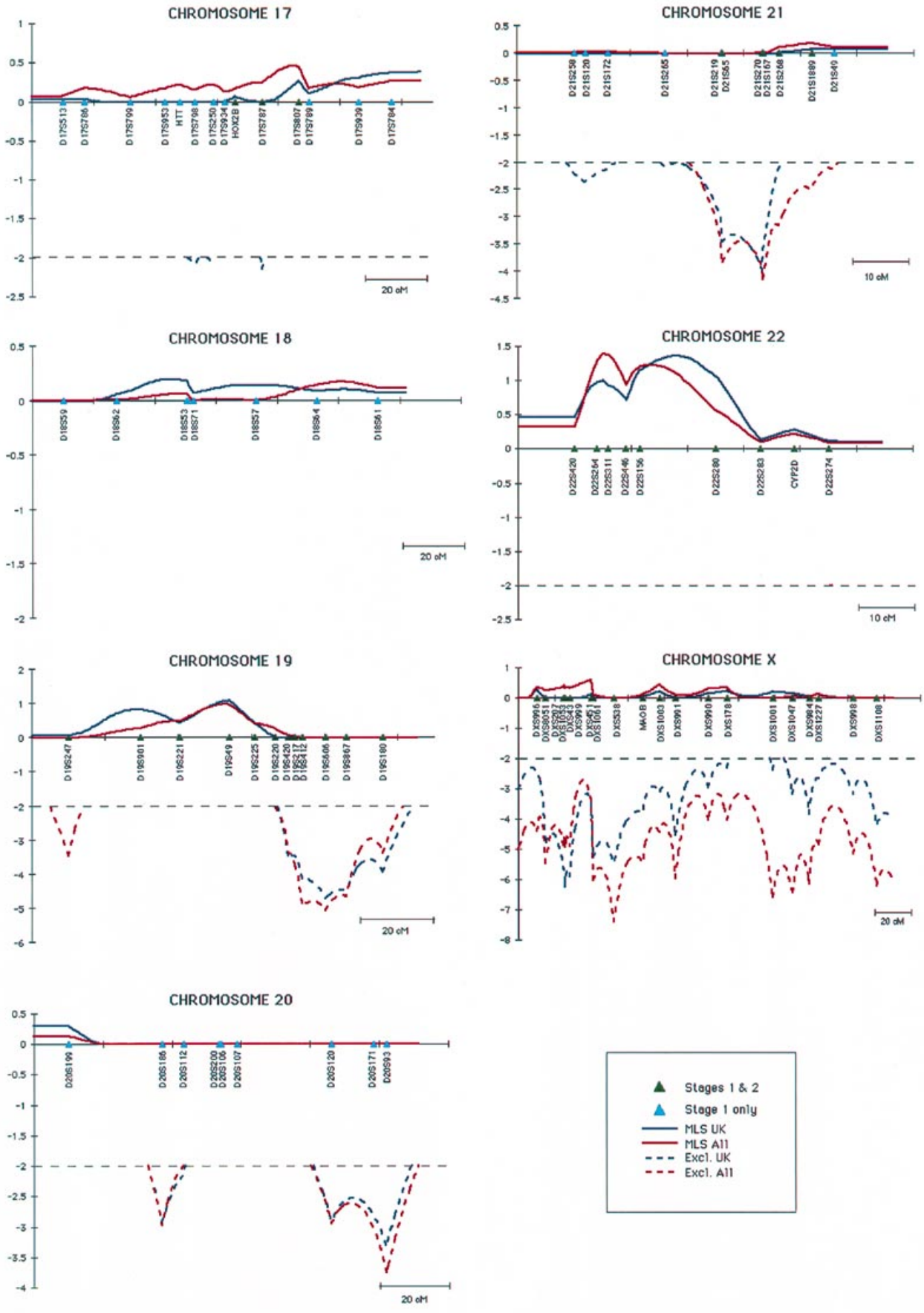

CHROMOSOME 17
CHROMOSOME 21
CHROMOSOME 18
CHROMOSOME 22
CHROMOSOME 19
CHROMOSOME X
CHROMOSOME 20
20 cM
10 cM
Stages 1 & 2
Stage 1 only
MLS UK
MLS All
Excl. UK
Excl. All

Table 2. Loci with a single-point MLS > 1 in either all familites or the UK subset determined by the SPLINK program, the lowest GH *P*-value for each region, the maximum identity-by-descent sharing and multipoint MLS generated by ASPEX

				ALL							UK		
Marker	Position cM	SPLINK MLS	SPLINK P-value	GH P-values	% Sharing	ASPEX MLS	Position cM	SPLINK MLS	SPLINK P-value	GH P-values	% Sharing	ASPEX MLS
PEAK 2	**103.0**			**0.0356**	**56.4**	**0.65**	**103.0**			**0.0361**	**58.6**	**0.76**
D2S1351	103.0	2.14	0.0015					2.36	0.0009			
D2S142	166.3	1.33	0.0109					0.32	0.1560			
D3S1303	138.4	1.06	0.0220					0.63	0.0675			
PEAK 3	**141.2**			**0.1237**	**57.6**	**0.73**	**141.2**			**0.1853**	**58.7**	**0.55**
D4S2936	0.0	1.52	0.0070					0.76	0.0470			
D4S412	3.0	1.33	0.0111					0.73	0.0510			
PEAK 4	**4.8**			**0.0036**	**60.7**	**1.55**	**21.0**			**0.0155**	**61.6**	**1.10**
CFTR	125.5	1.36	0.0101					0.87	0.0358			
D7S480	127.2	1.70	0.0045					0.99	0.0263			
D7S530	136.4	1.30	0.0121					2.87	0.0003			
PEAK 7	**144.7**			**0.0022**	**64.0**	**2.53**	**146.3**			**0.0006**	**70.1**	**3.55**
D7S684	149.6	2.26	0.0011					2.58	0.0005			
D7S2513	154.1	1.00	0.0251					1.89	0.0028			
PEAK 8	**35.0**			**0.0339**	**57.4**	**0.79**	**35.0**			**0.0420**	**58.4**	**0.61**
D8S1786	44.1	1.43	0.0087					1.86	0.0030			
D10S197	50.5	1.55	0.0065					0.45	0.1101			
PEAK 10	**51.9**			**0.0087**	**60.7**	**1.36**	**30.3**			**0.0577**	**60.5**	**0.69**
D10S201	105.9	1.25	0.0135					0.76	0.0467			
D12S338	113.3	1.05	0.0221					0.46	0.1027			
PEAK 12	**113.9**			**0.0303**	**58.8**	**0.86**	**113.9**			**0.1722**	**57.3**	**0.36**
PEAK 13	**85.0**			**0.0317**	**56.3**	**0.59**	**103.4**			**0.1084**	**58.5**	**0.46**
D13S193	85.0	1.52	0.0066					1.05	0.0218			
D14S80	20.6	2.32	0.0010					2.06	0.0019			
D14S1034	25.3	1.22	0.0146					1.39	0.0094			
PEAK 14	**32.2**			**0.0365**	**57.6**	**0.99**	**22.9**			**0.0318**	**60.0**	**0.97**
D14S70	32.9	0.99	0.0261					0.66	0.0623			
D16S407	16.7	1.28	0.0128					1.08	0.0208			
PEAK 16	**17.3**			**0.0054**	**59.4**	**1.51**	**19.3**			**0.0126**	**65.9**	**1.97**
D16S3114	21.8	1.05	0.0228					1.12	0.0190			
PEAK 19	**48.2**			**0.0324**	**59.2**	**0.99**	**49.0**			**0.0304**	**61.1**	**1.11**
D19S49	49.0	1.16	0.0164					1.45	0.0080			
D22S264	4.0	1.30	0.0121					0.71	0.0547			
PEAK 22	**5.0**			**0.0073**	**59.7**	**1.39**	**17.5**			**0.0226**	**63.7**	**1.36**
D22S280	25.0	0.92	0.0310					1.36	0.0103			

The position is the approximate relative position in cM from pter to qter. When the locations of the peaks (the highest MLS for each region) reported by GH and ASPEX differ slightly, the ASPEX position is indicated.

data were analysed together rather than treating stage 2 as a 'replication' data set.

The results of the multipoint analyses of the combined stage 1 and stage 2 data using the program ASPEX (10,11) are displayed in Figure 1. Because the ASPEX program only analyses data from sibling pair families, the Genehunter program [GH (12)] was also used since it can analyse data from both sibling-pair and other relative-pair families. The single-point results for those markers that had an MLS of 1.0 or greater and the multipoint results for each region are presented in Table 2. Based on all 87 sib-pair families, ~32% of the genome was excluded for λ_s = 2.5; the entire X chromosome was excluded at this level of λ_s, consistent with the results of Hallmayer *et al.* (13). The majority of the single-point results are consistent with the multipoint curves. However, one of the highest single-point MLS [2.14 at D2S1351 using SPLINK maximized over Holman's 'possible triangle' (14)] is much lower on the multipoint curve, most likely due to lower sharing at the flanking markers. This region gives an MLS of 0.65 using ASPEX with an additive model and also only achieves an MLS of 0.79 using MAPMAKER/SIBS (15) maximized over the 'possible triangle'. Similarly, three markers on chromosome 14 have single-point MLS at or above 1.0 (Table 2), yet the multipoint curve only achieves a value of 0.99 in this region (ASPEX additive model), while MAPMAKER/SIBS maximized over the 'possible triangle' gives an MLS of 1.48.

Using ASPEX, six chromosomes (4, 7, 10, 16, 19 and 22) with regions generating a multipoint MLS > 1 were identified in either the UK or total families (Fig. 1, Table 2). The long arm of chromosome 7 from D7S530 to D7S684 was the most significant region, with a multipoint MLS of 2.53 (GH P = 0.0022) in all families and an MLS of 3.55 (GH P = 0.00057) in the subset of UK families. Based on the estimated sharing probabilities in the interval between D7S530 and D7S684 in the UK families (z_0= 0.05, z_1= 0.50, z_2 = 0.45), this locus has an effect corresponding to a λ_s of 5.0. The next most significant region was on the short arm of chromosome 16 near markers D16S407 and D16S3114, with a multipoint MLS of 1.51 (GH P = 0.0054) in all families and 1.97 (GH P = 0.0126) in the UK families. Based on the estimated sharing probabilities in the UK families (z_0= 0.09, z_1= 0.50, z_2 = 0.41) in the interval between D16S407 and D16S3114, the region-specific λ_s is 2.8. No elevated IBD sharing is observed in the relevant regions on either chromosome 7 or 16 using the 31 non-UK families; this may be due to heterogeneity across populations or simply small sample size. There was no evidence of linkage disequilibrium in either of these regions but the markers are far apart (5–10 cM). The next most significant region

was on chromosome 4 with a multipoint MLS of 1.55 (GH P = 0.0036) in all families, an MLS of 1.1 (GH P = 0.0155) in the UK families and an MLS of 0.7 in the non-UK families.

DISCUSSION

Over 300 transcripts map to the chromosome 7q region (16) and possible candidate genes expressed in the brain include a G protein-coupled peptide receptor (GPR37), protein tyrosine phosphatase receptor type ζ polypeptide (PTPRZ1), ephrin tyrosine kinase receptor (EPHB6), muscarinic acetylcholine receptor M2, pleiotrophin (PTN), neural precursor cell expressed developmentally down-regulated 2 (NEDD2/ICH1/CASP2), glutamate receptor metabotropic 8 (GRM8), similar to potassium channel EAG, similar to synaptophysin and similar to 5′AMP-activated protein kinase γ chain. A gene for tuberous sclerosis (TSC) has been mapped telomeric to the chromosome 16p region but not ordered with respect to other markers (Stanford Human Genome Center RH map, http://www-shgc.stanford.edu). However, TSC was clinically excluded in the autistic probands in this study.

In summary, the first full genome scan in autism has revealed several interesting loci, one of which achieves an MLS of 3.55 in the largest subset of relative-pair families. Further families, including singleton cases, are currently being ascertained to replicate these findings. Fine mapping, tests for linkage disequilibrium and analysis of candidate genes in these regions are underway.

MATERIALS AND METHODS

Families

An international consortium of clinicians identified potential multiplex autism families from clinic cases, and by mailing health care professionals, special schools and members of National Autistic Societies. Initial screening excluded cases younger than 4 years, those who appeared unlikely to fulfil diagnostic criteria, cases with a past or current medical disorder of probable etiological significance and families in which both probands were apparently profoundly handicapped. Clinical assessments were then conducted on 178 potential multiplex families. The ADI (8) and the Vineland Adaptive Behaviour Scales (17) were administered to parents and an obstetric and medical history taken. Potential probands were assessed using the Autism Diagnostic Observation Schedule [ADOS (18)], or a later revision. Psychometric data are currently being obtained. A physical examination of potential cases included a careful search for phakamatoses to rule out TSC. A blood sample was taken from both cases and available first degree relatives. When possible karyotyping was performed on both cases in a family and molecular genetic testing for Fragile X on one case, previous results were also obtained. This study was approved by the ethical committees of the collaborating organizations.

Genotyping

Blood samples were taken and genomic DNA was extracted using Nucleon® kit. In addition, lymphoblastoid cell lines were generated from peripheral blood leukocytes, providing a renewable source of DNA. In 15 cases in which a blood sample could not be obtained, DNA was extracted from buccal swabs. Genotyping was undertaken using a fluorescence-based semi-automatic method (9). Polymerase chain reactions (PCR) were performed in 96-well microtitre plates, in a final volume of 15 µl containing 40 ng of genomic DNA, 10 mM Tris pH 8.3, 50 mM KCl, 1–3 mM $MgCl_2$, 200 µM dNTPs, 0.2 µM of each primer and 0.25 U Taq polymerase. Thirty-five cycles (30 s at 94°C, 30 s at 50–66°C, 30 s at 72°C) were performed in MJ Research thermocyclers. PCR products were combined into pools and typed using ABI 373A sequencing machines and the GENESCAN/GENOTYPER software (Applied Biosystems). Checking for non-mendelian inheritance of markers and conversion of allele sizes to whole numbers were performed using the GAS package (version 2, ©1993–1995, A. Young, Oxford University). Genbase (version 2.0.5, J.-M. Sebaoun and M. Lathrop) was used to store all genotypic and phenotypic data and to produce the necessary files for statistical analysis.

The genome screen consisted of 354 microsatellite markers with an average intermarker distance of 10 cM and average heterozygosity of 0.77. The order and genetic distances were taken from the Généthon map (19) and other published maps (20,21). The accuracy of the input marker map was checked by estimating intermarker genetic distances from the marker data.

Statistical analysis

In addition to straightforward error detection based on simple genotype elimination, the marker data were haplotyped using SIMWALK2 (22–24) to check for chromosomes with an excessive number of recombination events. The initial analysis of the stage 1 data used SPLINK to compute pairwise MLS scores maximized under the 'possible triangle' restrictions (14). Subsequent analyses were carried out with ASPEX (10,11) which uses information from all the marker loci on a chromosome simultaneously. Both of these programs use maximum likelihood methods to estimate marker allele frequencies from the input data. However, the results should be relatively insensitive to misspecification of marker allele frequencies as both parents were genotyped in 95% of the families in this study. ASPEX computes a multipoint MLS, maximized over λ_s, as well as an exclusion map along each chromosome. The exclusion map is a function of the assumed (fixed) value for λ_s, which for the exclusion maps presented here was taken to be 2.5. All ASPEX multipoint analyses were performed under an additive model (no dominance variance), so that if z_i is the probability of an affected sib-pair sharing i alleles identical by descent, then $z_0 = 0.25/\lambda_s$, $z_1 = 0.50$ and $z_2 = 0.50 - z_0$. For the regions on chromosome 2 and 14 where the single-point and multipoint scores under an additive model were conflicting, multipoint analyses used MAPMAKER/SIBS (15) maximized over the possible triangle. Since ASPEX and SPLINK only use sib-pairs, non-parametric Z-pair statistics were computed using Genehunter (12), which permits the inclusion of an additional 12 non-sib-pair families, each containing one extended relative-pair. In the three families with three affected individuals, all possible pairs were used in the analyses. For the Genehunter analyses, maximum likelihood estimates of marker allele frequencies as provided by SPLINK were used. In the regions of interest on chromosome 7 and 16 a total of 13 markers were tested for linkage disequilibrium using the transmission disequilibrium test (25,26), as implemented in the ASPEX program.

There has been much discussion about what is the appropriate evidence for 'significant' or 'suggestive' linkage (27–30); the suggestion of Elston (31) of forgoing such labelling of results has been followed here. The *P*-values returned by Genehunter are known to be quite conservative (12,32), however, and may understate the significance of these findings.

ACKNOWLEDGEMENTS

We thank the subjects and their families for their cooperation with this study and the professionals and Autistic societies who helped with ascertainment. We thank David Clark, Deborah Gomer and Helen Smith for administrative support. This study is supported by the UK Medical Research Council, The Wellcome Trust, Deutsche Forschungsgemeinschaft, The Janusz Korczack Foundation, NIMH, The Northern and Yorkshire Regional Research and Development Project, Danish Medical Research Council, France-Telecom, NICHD and BIOMED.

REFERENCES

1. World Health Organization (1992) *The ICD-10 Classification of Mental and Behavioral Disorders: Clinical Descriptions and Diagnostic Guidelines.* World Health Organization, Geneva.
2. Bailey, A., Phillips, W. and Rutter, M. (1996) Autism: towards an integration of clinical, genetic, neuropsychological, and neurobiological perspectives. *J. Child Psychol. Psychiatr.*, **37**, 89–126.
3. Cohen, D.J. and Volkmar, F. (eds) (1997) *Handbook of Autism and Pervasive Developmental Disorders*, 2nd edn. Wiley, New York.
4. Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E. and Rutter, M. (1995) Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol. Med.*, **25**, 63–77.
5. Steffenburg, S., Gillberg, C., Hellgren, L., Andersson, L., Gillberg, I., Jakobsson, G. and Bohman, M. (1989) A twin study of autism in Denmark, Finland, Iceland, Norway and Sweden. *J. Child Psychol. Psychiatr.*, **30**, 405–416.
6. Bolton, P., Macdonald, H., Pickles, A., Rios, P., Goode, S., Crowson, M., Bailey, A. and Rutter, M. (1994) A case-control family history study of autism. *J. Child Psychol. Psychiatr.*, **35**, 877–900.
7. Pickles, A., Bolton, P., Macdonald, H., Bailey, A., Le Couteur, A., Sim, C.-H. and Rutter, M. (1995) Latent-class analysis of recurrence risks for complex phenotypes with selection and measurement error: a twin and family history study of autism. *Am. J. Hum. Genet.*, **57**, 717–726.
8. Lord, C., Rutter, M. and 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. Aut. Dev. Dis.*, **24**, 659–685.
9. Reed, P.W., Davies, J.L., Copeman, J.B., Bennett, S.T., Palmer, S.M., Prithchard, L.E., Gough, S.C.L., Kawaguchi, Y., Cordell, H.J., C.H., Balfour, K.M., Jenkins, S.C., Powell, E.E., Vignal, A. and Todd, J.A. (1994) Chromosome-specific microsatellite sets for fluorescence-based, semi-automated genome mapping. *Nature Genet.*, **7**, 390–395.
10. Hauser, E.R., Boehnke, M., Guo, S.W. and Risch, N. (1996) Affected-sib-pair interval mapping and exclusion for complex genetic traits: sampling considerations. *Genet. Epidemiol.*, **13**, 117–137.
11. Hinds, D. and Risch, N. (1996) The ASPEX package: affected sib-pair mapping. ftp://lahmed.stanford.edu/pub/aspex.
12. Kruglyak, L., Daly, M.J., Reeve-Daly, M.P. and Lander, E.S. (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am. J. Hum. Genet.*, **58**, 1347–1363.
13. Hallmayer, J., Hebert, J.M., Spiker, D., Lotspeich, L., McMahon, W.M., Petersen, P.B., Nicholas, P., Pingree, C., Lin, A.A., Cavalli-Sforza, L.L., Risch, N. and Ciaranello, R.D. (1996) Autism and the X chromosome. *Arch. Gen. Psychiatr.*, **53**, 985–989.
14. Holmans, P. (1993) Asymptotic properties of affected-sib-pair linkage analysis. *Am. J. Hum. Genet.*, **52**, 362–374.
15. Kruglyak, L. and Lander, E.S. (1995) Complete multipoint sib-pair analysis of qualitative and quantitative traits. *Am. J. Hum. Genet.*, **57**, 439–454.
16. Schuler, G.D., Boguski, M.S., Stewart, E.A., Stein, L.D., Gyapay, G., Rice, K., White, R.E., Rodriguez-Tome, P., Aggarwal, A., Bajorek, E., Birren, B.B., Butler, A., Castle, A.B., Chiannilkulchai, N., Chu, A., Clee, C., Cowles, S., Day, P.J., Dibling, T., Drouot, N., Dunham, I., Duprat, S., East, C.,Edwards, C., Fan, J.-B., Fang, N., Fizames, C., Garrett, C., Green, L., Hadley, D., Harris, M., Harrison, P., Brady, S., Hicks, A., Holloway, E., Hui, L., Hussain, S., Luois-Dit-Sully, C., Ma, J., MacGilvery, A., Mader, C., Maratukulam, A., Matise, T.C., McKusick, K.B., Morissette, J., Mungall, A., Muselet, H.C., Nusbaum, C., Page, D.C., Peck, A., Perkins, S., Piercy, M., Qin, F., Quackenbush, J., Ranby, S., Reif, T., Rozen, S., Sanders, C., She, X., Silva, J., Slonim, D.K., Soderlund, C., Sun, W.-L., Tabar, P., Thangarajah, T., Vega-Czarny, N., Vollrath, D., Voyticky, S., Wilmer, T., Wu, X., Adams, M.D., Auffray, C., Walter, N.A.R., Branson, R., Dehejia, A., Goodfellow, P.N., Houlgatte, R., Hudson Jr, J.R., Ide, S.E., Iorio, K.R., Lee, W.Y., Seki, N., Nagase, T., Ishikawa, K., Nomura, N., Phillips, C., Polymeropoulos, M.H., Sandusky, M., Schmitt, K., Berry, R., Swanson, K., Torres, R., Venter, J.C., Sikela, J.M., Beckmann, J.S., Weissenbach, J., Myers. R.M., Cox, D.R., James, M.R., Bentley, D., Deloukas, P., Lander, E.S. and Hudson, T.J. (1996) A gene map of the human genome. *Science*, **274**, 540–546.
17. Sparrow, S.S., Balla, D. and Cicchetti, D.V. (1984) *Vineland Adaptive Behavior Scales*. American Guidance Service, Inc., Circle Pines, MN.
18. Lord, C., Rutter, M., Good, S., Heemsbergen, J., Jordan, H., Mawhood, L. and Schopler, E. (1989) Autism diagnostic observation schedule: a standardized observation of communicative and social behaviour. *J. Aut. Dev. Dis.*, **19**, 185–212.
19. Dib, C., Fauré, S., Fizames, C., Samson, D., Drouot, N., Vignal, A., Milasseau, P., Marc, S., Hazan, J., Seboun, E., Lathrop, M., Gyapay, G., Morissette, J. and Weissenbach, J. (1996) The Généthon human genetic linkage map. *Nature*, **380**, 152–154.
20. Davies, J.L., Kawaguchi, Y., Bennett, S.T., Copeman, J.B., Cordell, H.J., Pritchard, L.E., Reed, P.W., Gough, S.C.L., Jenkins, S.C., Palmer, S.M., Balfour, K.M., Rowe, B.R., Farrall, M., Barnett, A.H., Bain, S.C. and Todd, J.A. (1994) A genome-wide search for human type 1 diabetes susceptibility genes. *Nature*, **371**, 130–136.
21. Cooperative Human Linkage Center (1994) A comprehensive human linkage map with centimorgan density. *Science*, **265**, 2049–2054.
22. Sobel, E., Lange, K., O'Connell, J.R. and Weeks, D.E. (1995) In: Speed, T.P. and Waterman, M.S. (eds), *Genetic mapping and DNA sequencing: IMA Volumes in Mathematics and its Applications.* Springer-Verlag, New York.
23. Weeks, D.E., Sobel, E., O'Connell, J.R. and Lange, K. (1995) Computer programs for multilocus haplotyping of general pedigrees. *Am. J. Hum. Genet.*, **56**, 1506–1507.
24. Sobel, E. and Lange, K. (1996) Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. *Am. J. Hum. Genet.*, **58**, 1323–1337.
25. Terwilliger, J.D. and Ott, J. (1992) A haplotype-based 'haplotype relative risk' approach to detecting allelic associations. *Hum. Hered.*, **42**, 337–346.
26. Spielman, R.S., McGinnis, R.E. and Ewens, W.J. (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am. J. Hum. Genet.*, **52**, 506–516.
27. Lander, E. and Kruglyak, L. (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nature Genet.*, **11**, 241–247.
28. Sawcer, S., Jones, H.B., Judge, D., Visser, F., Compston, A., Goodfellow, P.N. and Clayton, D. (1997) Empirical genomewide significance levels established by whole genome simulations. *Genet. Epidemiol.*, **14**, 223–927.
29. Thomson, G. (1994) Identifying complex disease genes: progress and paradigms. *Nature Genet.*, **8**, 108–110.
30. Witte, J.S., Elston, R.C. and Schork, N.J. (1996) Genetic dissection of complex traits. *Nature Genet.*, **12**, 355–358.
31. Elston, R.C. (1997) 1996 William Allan Award Address. Algorithms and inferences: the challenge of multifactorial diseases. *Am. J. Hum. Genet.*, **60**, 255–262.
32. Davis, S. and Weeks, D.E. (1997) Comparison of nonparametric statistics for detection of linkage in nuclear families: single-marker evaluation. *Am. J. Hum. Genet.*, **61**, 1431–1444.