Human Molecular Genetics, 2001, Vol. 10, No. 9 973–982

Further characterization of the autism susceptibility locus *AUTS1* on chromosome 7q

International Molecular Genetic Study of Autism Consortium (IMGSAC)[+]

Received 5 January 2001; Revised and Accepted 2 March 2001

Autism is a neurodevelopmental disorder that usually arises on the basis of a complex genetic predisposition. The most significant susceptibility region in the first whole genome screen of multiplex families was on chromosome 7q, although this linkage was evident only in UK IMGSAC families. Subsequently all other genome screens of non-UK families have found some evidence of increased allele sharing in an overlapping 40 cM region of 7q. To further characterize this susceptibility locus, linkage analysis has now been completed on 170 multiplex IMGSAC families. Using a 5 cM marker grid, analysis of 125 sib pairs meeting stringent inclusion criteria resulted in a multipoint maximum LOD score (MLS) of 2.15 at *D7S477*, whereas analysis of all 153 sib pairs generated an MLS of 3.37. The 71 non-UK sib pairs now contribute to this linkage. Linkage disequilibrium mapping identified two regions of association—one lying under the peak of linkage, the other some 27 cM distal. These results are supported in part by findings in independent German and American singleton families.

INTRODUCTION

Autism (MIM 209850) is a severe neurodevelopmental disorder characterized by impairments in reciprocal social interaction and communication, and restricted and stereotyped patterns of interests and activities, with an onset before three years of age. Twin and family studies indicate a strong genetic liability to autism (1–3) and several genome screens for autism susceptibility loci have recently been completed (4–7). The International Molecular Genetic Study of Autism Consortium (IMGSAC) study of 99 multiplex families (4) pointed to the probable involvement of regions on six different chromosomes (chromosomes 4,7,10,16,19 and 22), with the region on chromosome 7q giving the most significant multipoint maximum LOD score (MLS). Subsequently other groups have also reported

[+]http://www.well.ox.ac.uk/~maestrin/iat.html

MRC Child Psychiatry Unit and Centre for Social, Genetic and Developmental Psychiatry, Institute of Psychiatry, London, UK: Sarah Palferman, Nicola Matthews, Martha Turner, Janette Moore, Amaia Hervas, Anne Aubin, Simon Wallace, Janine Michelotti, Catherine Wainhouse, Alina Paul, Elaine Thompson, Marianne Murin, Ramyani Gupta, Claire Garner, Andrew Pickles, Michael Rutter and Anthony Bailey[§].
Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK: Janine A. Lamb, Angela Marlow, Pat Scudder, Gabrielle Barnby and Anthony P. Monaco[§].
Newcomen Centre, Guy's Hospital, London, UK: Gillian Baird and Anthony Cox; Regional Genetics Centre, Division of Medical and Molecular Genetics, Guy's Hospital, London, UK: Zoe Docherty, Pamela Warburton, Elizabeth P. Green and Stephen J. Abbs.
Flemming Nuffield Unit, Newcastle, UK: Ann Le Couteur, Helen R. McConachie and Tom Berney; Neuropsychology Department, Newcastle General Hospital, Newcastle, UK: Thomas P. Kelly.
Developmental Psychiatry Section, University of Cambridge Clinical School, Cambridge, UK: Petrus J. De Vries and Patrick F. Bolton.
Booth Hall Children's Hospital, Manchester, UK: Jonathan Green, Anne Gilchrist and Jane Whittacker.
European Centre for Collection of Animal Cell Cultures, Salisbury, UK: Bryan Bolton and Ros Packer.
Dipartimento di Biologia, Universita' di Bologna, Bologna, Italy: Elena Maestrini.
AZU, Department of Child and Adolescent Psychiatry, Utrecht, The Netherlands: Herman Van Engeland, Maretha V. De Jonge and Chantal Kemner.
Deutsches Krebsforschungszentrum, Molecular Genome Analysis, Heidelberg, Germany: Sabine M. Klauck, Kim S. Beyer, Sabine Epp and Annemarie Poustka; Deutsches Krebsforschungszentrum, Biostatistics, Heidelberg, Germany: Axel Benner.
J.W. Goethe-Universität, Department of Child and Adolescent Psychiatry, Frankfurt, Germany: Fritz Poustka, Dorothea Rühl, Gabriele Schmötzer, Sven Bölte and Sabine Feineis-Matthews.
Unite de Diagnostic et Evaluation de l'autisme, Hopital la Grave, Toulouse, France: Eric Fombonne, Bernadette Rogé, Jeanne Fremolle-Kruck, Catherine Pienkowski and Marie-Thérèse Tauber.
National Centre for Autism, Virum, Denmark: Lennart Pedersen, Karen-Brondum-Nielsen, Gunna Eriksen and Demetrious Haracopos.
Biophysics, Danish Technical University, Lyngby, Denmark: Rodney M.J. Cotterill.
Department of Psychological Paediatrics, Agia Sophia Children's Hospital, Athens, Greece: John Tsiantis and Katerina Papanikolaou.
Department of Psychiatry, University of Chicago, Chicago, USA: Catherine Lord, Christina Corsello, Stephen Guter, Bennett Leventhal and Edwin Cook.
UCLA Center for Neurobehavioural Genetics, Los Angeles, USA: Susan Smalley, Julia Bailey, Amy Liu, Martha Dedricks, Lisa Chrzanowski and Jennifer Levitt.
Child Study Centre, Sterling Hall of Medicine, Yale University, New Haven, USA: David Pauls and Fred Volkmar.
Department of Human Genetics, University of Pittsburgh, Pittsburgh, USA: Daniel E. Weeks.
[§]To whom correspondence should be addressed. A.J. Bailey, Centre for Social, Genetic and Developmental Psychiatry, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, UK; Tel: +44 207 252 5756; Fax: +44 207 252 5107; Email: anthony.bailey@iop.kcl.ac.uk; A.P. Monaco, Wellcome Trust for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK; Tel: +44 1865 287502; Fax: +44 1865 287501; Email: anthony.monaco@well.ox.ac.uk

Table 1. Summary description of families meeting inclusion criteria

Breakdown of relative pairs	Original screen[a]		Additional families		Total	
Sibling pair families	77		67		144	
Sibling trio families	3		0		3	
Other extended relative-pair families	11		12		23	
Total number of affected individuals					344	
Age at interview of affected individuals (mean ± SD)					9.2 ± 6.3 years	
Case type of affected pairs	Type 1/type 1	Type 1/type 2	Type 2/type 2	Type 1/type 3	Type 2/type 3	Type 3/type 3
	54	92	14	11	5	2
Sex of affected pairs	Male/male	Male/female	Female/female			
	123	46	9			
Sex ratio of affected individuals (M:F)	4.6					
ADI and Vineland scores (Mean ± SD)	Type 1		Type 2		Type 3	
Social	24.33 ± 5.98		20.94 ± 6.31		23.22 ± 6.85	
Non verbal communication	11.21 ± 3.17		9.84 ± 3.65		10.78 ± 3.72	
Verbal communication	16.53 ± 4.59		14.77 ± 4.26		16.33 ± 4.39	
Repetitive	8.13 ± 2.76		7.10 ± 2.77		7.78 ± 2.82	
Vineland adaptive behaviour composite scores[b]	47.44 ± 16.15		54.82 ± 21.62		48.62 ± 26.90	

[a]International Molecular Genetic Study of Autism Consortium (4).
[b]Excludes data on families from Germany.

evidence of increased allele sharing on 7q in autism relative pair families (5,7,8), reviewed by Lamb *et al* (9). We report further work to define the location of the autism susceptibility gene on chromosome 7q (*AUTS1*), including addition of further families, fine mapping of the region with a high density of markers and allelic association studies.

RESULTS

Identification of families

Strict criteria were applied to identify 170 multiplex families comprising a total of 178 non-independent affected relative pairs. Summary details of these families are provided in Table 1. Affected individuals were predominantly Caucasian (97.2%). The 170 families originated from 12 sites in six countries (Denmark, 7; France, 9; Germany, 10; Netherlands, 14; UK: Cambridge, 6; Guy's Hospital, London, 10; Institute of Psychiatry, London, 67; Manchester, 4; Newcastle, 9; USA: Chicago, 22; UCLA, 4; Yale, 8). Individuals were assigned case status in a hierarchical fashion; they were designated case type 1 if they had a clinical diagnosis of autism, met autism diagnostic interview-revised (ADI-R) (10) and autism diagnostic observation schedule (ADOS) (11) or ADOS-generic (ADOS-G) (12) algorithm criteria for autism, and had a history of language delay and a performance IQ of 35 or above. Individuals were designated case type 2 if they had a clinical diagnosis of autism, atypical autism, Asperger's syndrome or pervasive developmental disorder not otherwise specified (PDDNOS) and met at least ADOS-G criteria for PDD; there was no requirement for a history of language delay or a performance IQ of 35 and individuals were allowed to fall one point below threshold on one behavioural domain of the ADI-R. Individuals were designated case type 3 if they had a clinical diagnosis of autism or another pervasive developmental disorder (PDD) and met ADI-R criteria for autism, but failed to meet ADOS-G criteria for PDD; or if they met clinical, ADI-R and ADOS-G criteria but an IQ could not be computed because they fell outside the age range of the test (i.e. low functioning individuals). In 157 of the 178 affected relative pairs, at least one individual met criteria for case type 1. In the remaining 21 affected relative pairs both individuals were designated case type 2 in 14 pairs, five pairs met criteria for case type 2/type 3, and two pairs met criteria for case type 3/type 3. The total number of affected individuals is 344. Due to tightening of diagnostic criteria, eight families included in the original publication (4) have now been excluded, either because they were non-verbal individuals administered the ADOS (for which there is no non-verbal algorithm), or because an individual would not co-operate with psychometric testing.

Genotyping and linkage analysis

Results on the 153 affected sib pairs (ASPs). Multipoint linkage analyses of the combined data from 153 ASPs from 147 nuclear families were performed using ASPEX (13,14) and GENEHUNTER-PLUS (15). Analysis with ASPEX of the 18 chromosome 7 markers originally used in the genome screen (4) generated a multipoint MLS of 1.91 (GENEHUNTER-PLUS LOD = 1.86) between markers *D7S527* and *D7S486* in the 153 ASPs; this compares with the MLS of 1.61 (GENEHUNTER-PLUS LOD = 1.77) obtained some ~20 cM distal between markers *D7S530* and *D7S684* in the 86 sib pairs from the 80 nuclear families originally published (4).

Analysis of an additional 84 markers across the region *D7S524–D7S483* further supported these results. These markers showed significant variation in marker heterozygosity (mean, 0.72; range, 0.29–0.92) and effective sample size (mean, 72; range, 23–114)—equivalent to the number of fully informative sib pairs that would carry the same information—as determined using SPLINK (16,17). Mean inter-marker distance across this region was 0.77 cM. In order to minimize the inter-marker variation in informativity, an ~5 cM grid of markers was generated; this map density has been shown to extract the maximum linkage information per marker, whilst minimizing the confounding effects of map and genotyping errors encountered at higher map resolutions (18). Seven new markers, in addition to the 18 already published (4), were selected based on the SPLINK results, to give 25 markers with mean heterozygosity of 0.79 and mean effective sample size of >90. Multipoint linkage analysis of the total sample of 153 sib pairs using this 5 cM map generated an ASPEX MLS of 3.37 (GENEHUNTER-PLUS LOD = 3.28). The maximum identity-by-descent sharing across this region for all sib pairs was 61.9% and, based on the estimated sharing probabilities, this region has a locus-specific λ_s of 1.9.

To examine parental sex-specific contributions to the linkage on chromosome 7q, analysis was performed according to the inheritance of paternally or maternally derived alleles. Comparison of the single point results revealed no significant ($P < 0.05$) parental sex specific differences in informative 0,1 allele sharing at markers across the region; the degree of fluctuation indicated a stochastic process. Analysis of the multipoint results using ASPEX showed a paternal-specific multipoint MLS of 2.26 at marker *D7S477*, compared with a maternally derived MLS of 1.08 at this locus. However, examination of these multipoint results using flanking markers to infer 0,1 sharing at this locus indicate that this result is not significantly different between the parental sexes ($P = 0.58$).

The single-point MLS results generated using SPLINK for the 153 affected sib pairs for all typed markers with an MLS >1, and all original markers in the interval *D7S524–D7S483*, are shown in Table 2. Heterozygosity, sharing probability estimates and effective sample size are also shown for each marker. The majority of single-point results are consistent with the multipoint curve. Notably, there were two markers generating an MLS >2 (*D7S2409*, MLS = 2.01, $P = 0.0020$; and *D7S480*, MLS = 2.07, $P = 0.0018$) and one marker generating an MLS >3 (*D7S477*, MLS = 3.07, $P = 0.0002$). This high single-point result for marker *D7S477* contributes to the spike in the 5 cM multipoint linkage curve due to lower sharing at flanking markers.

Results on subsets of families. To facilitate data analysis and to minimize potential sources of genetic heterogeneity, analysis was also performed on nuclear families according to case type of affected sib pairs (Fig. 1). Analysis of the 44 case type 1/type 1 affected sib pairs using the 5 cM grid of markers showed no significant evidence of linkage across this region. However, the 81 case type 1/type2 affected sib pairs generated an ASPEX MLS of 3.08 (GENEHUNTER-PLUS LOD = 2.97) at marker *D7S477*. Combined analysis of all 125 sib pairs meeting stringent inclusion criteria (case type 1/type1 and type 1/type 2 pairs) resulted in an ASPEX MLS of 2.15 (GENEHUNTER-PLUS LOD = 2.01).

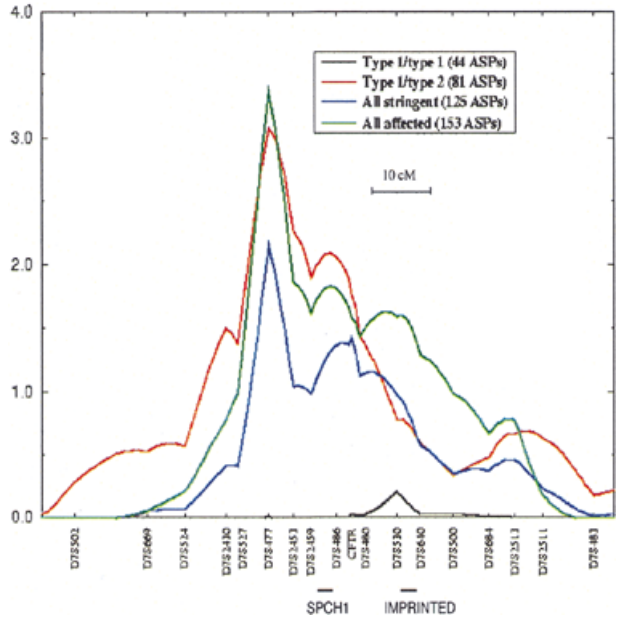

Figure 1. Multipoint MLS curves for 153 affected sib pairs generated by ASPEX under a model of no dominance variance for the 5 cM marker map. The maximum LOD scores for case type 1/type 1 sib pairs are shown in black, case type 1/type 2 pairs are shown in red, the combination of case type 1/type 1 and type 1/type 2 pairs are shown in blue and all pairs are shown in green. SPCH1 indicates the position of the locus for a specific speech and language disorder. IMPRINTED indicates the suggested imprinting cluster region containing the genes PEG1/MEST and γ2-COP.

The 82 affected sib pairs from the UK constituted the largest group from a single country, and were therefore analysed separately to represent a more homogeneous population. Using ASPEX, a multipoint MLS of 1.54 (GENEHUNTER-PLUS LOD = 1.43) was generated at marker *D7S477*, indicating that the 71 non-UK sib pairs now contribute to the linkage results, in contrast to the data reported previously (4).

Recent studies have suggested that autism and specific language impairment may have overlapping genetic aetiologies (19–21). Therefore, we tested the hypothesis that the linkage on chromosome 7q could be attributed to a gene responsible for language development. In 109 affected sib pairs both individuals had language delay (no single words by 24 months or no phrases by 33 months), whereas in only six sib pairs did neither individual have language delay, reflecting the inclusion criteria for this study. Using ASPEX, a multipoint MLS of 2.14 (GENEHUNTER-PLUS LOD = 2.07) was generated at marker *D7S477* in the sib pairs in which both individuals had language delay, reflecting the results obtained for all affected sib pairs and indicating that linkage at this marker is associated only in part with a history of language delay.

Multipoint analysis using GENEHUNTER-PLUS of the 23 extended relative-pair families containing 25 non-independent affected relative-pairs generated a LOD score of 0.54 at marker *D7S2513*.

Linkage disequilibrium

In the interests of clarity and to explain our choice of markers, linkage disequilibrium results are presented chronologically.

Table 2. Markers with a single-point MLS > 1 determined using SPLINK and for all original loci in the interval *D7S524–D7S483*

Marker	Position (Haldane cM)	SPLINK MLS	SPLINK *P*-value	Heterozygosity	z0	z1	z2	Effective sample size
D7S524	**104.86**	**0**	**0.5949**	**0.77**	**0.25**	**0.50**	**0.25**	**65.38**
D7S2409	110.57	2.01	0.0020	0.73	0.09	0.50	0.41	54.06
D7S527	**114.13**	**0.94**	**0.0299**	**0.84**	**0.17**	**0.50**	**0.33**	**80.87**
D7S651	117.98	1.37	0.0101	0.76	0.17	0.48	0.35	95.17
D7S477	119.60	3.07	0.0002	0.74	0.10	0.50	0.40	71.23
D7S523	130.26	1.23	0.0141	0.69	0.19	0.46	0.35	96.39
D7S486	**133.61**	**1.71**	**0.0042**	**0.78**	**0.17**	**0.47**	**0.36**	**94.55**
CFTR	**134.11**	**0.76**	**0.0460**	**0.69**	**0.18**	**0.50**	**0.32**	**76.37**
D7S480	**135.53**	**2.07**	**0.0018**	**0.88**	**0.15**	**0.50**	**0.35**	**110.08**
D7S2486	137.75	1.51	0.0072	0.84	0.18	0.38	0.44	39.09
D7S490	138.15	1.17	0.0165	0.79	0.17	0.50	0.33	94.68
D7S487	138.35	1.03	0.0230	0.72	0.18	0.45	0.37	50.56
M290CA1	138.85	1.65	0.0049	0.54	0.13	0.50	0.37	43.50
D7S2527	139.15	1.66	0.0049	0.69	0.18	0.45	0.37	84.69
D7S2501	140.76	1.34	0.0107	0.76	0.17	0.47	0.35	81.51
D7S530	**141.96**	**1.07**	**0.0213**	**0.77**	**0.17**	**0.50**	**0.33**	**90.77**
D7S2519	142.97	1.41	0.0090	0.82	0.17	0.49	0.34	93.34
D7S649	143.97	1.00	0.0247	0.73	0.17	0.50	0.33	82.13
D7S640	146.08	1.29	0.0122	0.87	0.17	0.50	0.33	113.93
DJS327CA2	147.65	1.09	0.0194	0.55	0.21	0.41	0.38	44.57
ALDR1	148.39	1.19	0.0155	0.70	0.16	0.50	0.34	67.68
D7S2533	148.89	1.49	0.0071	0.42	0.16	0.42	0.41	36.59
D7S500	151.93	1.17	0.0167	0.87	0.17	0.50	0.33	106.33
D7S1837	152.43	1.13	0.0177	0.68	0.21	0.42	0.37	69.49
D7S509	154.24	1.49	0.0074	0.70	0.14	0.50	0.36	58.90
D7S684	**158.08**	**0.56**	**0.0806**	**0.83**	**0.23**	**0.45**	**0.32**	**95.93**
D7S2505	158.78	1.63	0.0051	0.66	0.20	0.40	0.40	63.68
D7S1824	160.49	1.17	0.0165	0.85	0.22	0.44	0.35	111.61
D7S1518	161.50	1.23	0.0142	0.81	0.15	0.47	0.38	50.19
D7S2513	**162.62**	**0.77**	**0.0455**	**0.79**	**0.18**	**0.50**	**0.32**	**85.95**
J2A	163.12	1.14	0.0175	0.76	0.18	0.45	0.37	58.26
D7S483	**176.48**	**0**	**0.5880**	**0.70**	**0.25**	**0.50**	**0.25**	**92.98**

Original loci are shown in bold.

Although there are statistical issues inherent in testing multiple loci, this was an exploratory analysis with the intention of replicating any positive results in independent data sets; therefore the data presented here represent nominal *P* values.

Multi-allelic association tests were initially performed using the transmission disequilibrium test (TDT) with 76 microsatellite markers on chromosome 7 in 105 affected sib pairs. Two markers were identified with a nominal *P* value <0.05 (*D7S1804*, $P = 0.035$ and *D7S2533*, $P = 0.00038$), with a third marker *D7S2437* in the interval, giving a *P* value of 0.070. Measurement of allelic transmissions to 71 siblings without PDD (*D7S1804*, $P = 0.21$; *D7S2437*, $P = 0.58$; *D7S2533*, $P = 0.81$) or within a group of 41 control families (*D7S1804*, $P = 0.62$; *D7S2437*, $P = 0.41$; *D7S2533*, $P = 0.79$) showed no evidence for increased transmission of associated alleles. Inter-marker distances for *D7S1804–D7S2437* and *D7S2437–D7S2533* were 0.9 and 2.3 cM, respectively. Two markers within the first interval and three markers within the second interval, respectively, showed no evidence of association. However, the nearest flanking marker was still ~200 kb distant.

On the basis of these results, these three markers were tested for bi-allelic TDT in two independent samples of autistic

Table 3. Summary of allelic transmissions in German families

Allele	Sample A (n = 64 trios)[a]				Sample B (n = 22 trios)[b]				Sample A + B (n = 86 trios)			
	Parental allele frequency (%)	T	NT	χ^2 sum (P value)	Parental allele frequency (%)	T	NT	χ^2 sum (P value)	Parental allele frequency (%)	T	NT	χ^2 sum (P value)
D7S1804[c]												
6	18.8	25	11	10.89 (0.029)[d]	19.3	10	5	3.33 (0.30)	18.9	35	16	14.16 (0.011)[d]
Other	81.2	11	25		80.7	5	10		81.1	16	35	
D7S2437												
1	15.6	16	20	0.89 (0.62)	14.8	8	1	10.89 (0.039)[d]	15.4	24	21	0.40 (0.76)
Other	84.4	20	16		85.2	1	8		84.6	21	24	
D7S2533												
3	71.5	31	24	1.78 (0.42)	63.6	18	4	17.82 (0.0043)[e]	69.5	49	28	11.45 (0.022)[d]
Other	28.5	24	31		36.4	4	18		30.5	28	49	

T, transmitted; NT, not transmitted; χ^2 sum, the sum of the χ^2 statistics for transmission of all alleles with the corresponding empirical P value.
[a]Sample A; individuals fulfilling criteria for autism and with a history of language delay.
[b]Sample B; individuals fulfilling criteria for autism but showing no language delay.
[c]Markers were tested bi-allelically for TDT using ASPEX, calculating transmissions of associated allele in IMGSAC families compared with all other alleles.
[d]$P < 0.05$, [e]$P < 0.01$.

singleton families—from Germany and the United States—by testing the IMGSAC associated allele versus all other alleles. The allelic transmissions in the German data set of 86 singleton families calculated using ASPEX are shown in Table 3. The sample was also subdivided into sample A, consisting of individuals meeting diagnostic criteria for autism and with a history of language delay, and sample B, consisting of individuals fulfilling diagnostic criteria for autism, but showing no language delay. Markers *D7S1804* and *D7S2533* showed significant evidence for allelic association in the combined sample (P = 0.011 and P = 0.022, respectively). Although *D7S2437* showed no evidence for association in the combined sample, there was some evidence for association in sample B (P = 0.039), although the number of transmissions was very small. Allele 6 of marker *D7S1804* showed nominally significant association (P = 0.015) in the singleton families from the United States (E. Cook, personal communication). Allelic transmissions at markers *D7S2437* and *D7S2533*, however, were not significant. These results will be published in detail separately.

In order to define the region of linkage disequilibrium in the IMGSAC families, 11 novel repeat sequences flanking these three markers and anticipated to be polymorphic were identified electronically from genomic sequence in the region (http://www.genome.washington.edu/UWGC/). Physical positions of new markers relative to existing microsatellites were accurately established where possible by electronic PCR (22) of the sequence data. PCR primers were designed flanking these repeats and typed as described previously. No evidence of linkage disequilibrium was detected. Nevertheless, due to the paucity of finished sequence data in the regions directly surrounding these associated markers, these new microsatellites are still relatively far away.

Linkage disequilibrium was calculated for all markers after the addition of a further 48 IMGSAC nuclear families. The results for the combined sample of 153 affected sib pairs for markers *D7S1804*, *D7S2437* and *D7S2533* are presented in Table 4. Markers *D7S2437* and *D7S2533* show significant allelic association (P = 0.026 and P = 0.00097, respectively), although the results for *D7S1804* are no longer supported (P = 0.22). There is also evidence for increased paternal over maternal transmission of the 99 base pair allele at marker *D7S477* (P = 0.040 and P = 0.53 for paternal and maternal transmissions, respectively) (Table 5). This marker lies some 27 cM proximal to *D7S2437* under the peak of linkage. Increased paternal over maternal transmission of this allele is also seen in an expanded German patient sample of 122 singleton families (P = 0.009 and P = 0.36 for paternal and maternal transmissions, respectively).

Although not statistically significant, six other markers show a trend ($P < 0.1$) towards transmission of specific alleles across this region. However, the 10 novel microsatellites isolated from large insert genomic clones and mapping in close proximity to six candidate genes in the region showed no evidence of allelic association in the combined data set of 153 affected sib pairs (data not shown).

The distribution of linkage disequilibrium over the region was studied using the standardized disequilibrium measure D', calculated using the program ARLEQUIN (23). No significant linkage disequilibrium was detected across the region, nor between the associated alleles of markers *D7S477*, *D7S2437* and *D7S2533* ($P < 0.001$).

DISCUSSION

These fine mapping studies add to the evidence for an autism susceptibility locus *AUTS1* on chromosome 7q. With the

Table 4. Allelic transmissions for associated markers determined using ASPEX for 153 affected sib pairs

Allele[a]	Parental allele frequency (%)	Paternal			Maternal			Combined		
		T	NT	χ^2 sum (*P* value)	T	NT	χ^2 sum (*P* value)	T	NT	χ^2 sum (*P* value)
D7S1804 (145.68)[b]				0.69			0.61			25.44 (0.22)
6 (273)[c]	20.0	53	41		34	33		88	75	
3 (261)	19.1	38	43		34	39		75	85	
7 (277)	18.1	33	43		31	32		67	78	
5 (269)	12.4	18	18		33	21		54	42	
Other	30.6	55	52		84	91		141	145	
D7S2437 (146.58)				0.22			0.18			27.52 (0.026)[d]
9 (230)	31.3	52	71		51	49		108	125	
8 (228)	14.4	25	25		33	32		61	60	
10 (232)	13.1	29	23		25	23		56	48	
1 (202)	**11.8**	**27**	**14**		**31**	**18**		**58**	**32**	
Other	29.5	66	66		58	76		126	144	
D7S2533 (148.89)				0.072			0.0023[e]			32.27 (0.00097)[e]
3 (189)	**72.1**	**50**	**27**		**49**	**26**		**113**	**67**	
4 (191)	21.4	25	39		21	49		60	102	
Other	6.5	10	19		19	14		29	33	

T, transmitted; NT, not transmitted; χ^2 sum, sum of the χ^2 statistics for transmission of all alleles with the corresponding empirical *P* value.
Associated alleles are indicated in bold.
[a]Alleles with a frequency <10% are grouped as 'other'.
[b]Positions in Haldane cM from pter.
[c]Allele size in base pairs.
[d]$P < 0.05$; [e]$P < 0.01$.

Table 5. Sex-specific allelic transmissions for marker *D7S477* determined using ASPEX for IMGSAC and German families

Allele[a]	Parental allele	Paternal			Maternal			Combined		
	frequency (%)	T	NT	χ^2 sum (*P* value)	T	NT	χ^2 sum (*P* value)	T	NT	χ^2 sum (*P* value)
IMGSAC (153 affected sib pairs)										
D7S477 (119.60)[b]				17.82 (0.040)[e]			5.48 (0.53)			13.44 (0.10)
3 (95)[c]	39.1	51	53		53	47		118	114	
5 (99)	**25.5**	**55**	**26**		**31**	**30**		**95**	**65**	
4 (97)	16.5	22	30		32	28		59	63	
7 (103)	12.1	17	26		17	22		36	50	
Other	6.7	10	20		9	15		19	35	
German patients A + B (122 trios)										
5 (99)	**16.5**	**25**	**9**	$P = 0.009$[f]	**12**	**7**	$P = 0.36$	**43**	**22**	13.57 (0.012)[e]
Other[d]	83.5	9	25		7	12		22	43	

T, transmitted; NT, not transmitted; χ^2 sum, sum of the χ^2 statistics for transmission of all alleles with the corresponding empirical *P* value.
The associated allele is indicated in bold.
[a]IMGSAC alleles with a frequency <10% are grouped as 'other'.
[b]Position in Haldane cM from pter.
[c]Allele size in base pairs.
[d]Markers were tested bi-allelically for TDT using ASPEX in German patients, calculating transmissions of associated allele in IMGSAC families compared with all other alleles.
[e]$P < 0.05$; [f]$P < 0.01$.

inclusion of more families and markers the peak of linkage is proximal to the initial peak (4); the linkage region extends over some 70 cM and overlaps regions subsequently identified in a number of other studies (5–8). This study highlights the difficulties in localizing linkage signals for complex disease. Simulation models chosen to represent complex traits have indicated a large degree of variation in location estimates (regions giving maximum evidence for linkage) (24), with confounding variables such as genetic heterogeneity potentially causing a downward bias in LOD scores and subsequent loss of power to detect linkage.

In this study, we have attempted to minimize potential sources of heterogeneity by testing for linkage after splitting the sample by case type of affected individuals, by country and by language delay. The reason for the very low MLS in case type 1/type 1 pairs is uncertain, but may simply be due to small sample size. This category excludes pairs in which both individuals are very low functioning, as well as those pairs in which one individual has another PDD. The MLS of 3.08 seen in the case type 1/type 2 affected sib pairs may indicate that linkage to the autism susceptibility locus *AUTS1* on chromosome 7q underlies the relatively high functioning component of autism seen in some of these case type 2 individuals. This hypothesis is supported by analysis of the case type 1/type 2 pairs after removal of 15 sib pairs containing an individual with an IQ <35, resulting in a slight increase in ASPEX MLS from 3.08 (81 pairs) to 3.38 (66 pairs). However, the relatively modest overall sample size limits fine grained phenotypic analysis and this finding may simply have arisen by chance. Analysis of the sample by country of origin and by language delay failed to reveal any further obvious sources of heterogeneity in this study and subdivision of the sample does not substantially narrow the region of linkage interest. Simulation of genome-wide linkage studies, however, has shown that true positive linkage peaks caused by the presence of a disease gene are expected to be wider than false positive peaks caused by random fluctuations alone (25). The small degree of sharing shown by the collection of extended relative pair families is again surprising, as these families would be expected to be most informative for localization of linkage as they have an increased number of meioses between affected individuals.

Increasing marker density and a concomitant increase in map and genotype error have also been shown to reduce the power to detect linkage, with the effects of errors in assumed parameter values being generally more severe when a larger number of less informative marker loci, rather than fewer more informative marker loci, are used (26). Our efforts to minimize data error in this study, however, including genotyping of both parents in >90% of families, use of an optimum map density of markers and use of the program SIBMED, indicate that genotype error is unlikely to exert a large influence over our results.

A current central strategy for mapping disease genes is the use of linkage analysis to approximately localize a gene, with subsequent use of linkage disequilibrium mapping for more accurate localization. Our results indicate two regions of association in our autistic disorder families; results supported in part by the findings in the German and American singleton families. The excess of paternal transmissions at marker *D7S477* is of interest with regard to the presence of two imprinted genes on chromosome 7q (27,28), as significant differences in parental transmission disequilibrium may indicate parent of origin specific effects. Although these results should be interpreted with caution due to the number of tests performed, a similar finding in the German singleton sample suggests that this result merits further study. Addition of further microsatellite markers, however, failed to better define the initial region of linkage disequilibrium. Although the ethnicity of our sample is >97% Caucasian, families were drawn from six countries, hence our assumption that linkage disequilibrium is the same in all populations is by no means certain, and may explain why stronger evidence for linkage disequilibrium has not been detected. Furthermore, several recent studies have indicated that the extent of disequilibrium is highly variable across the genome, with the assumed linear relationship between physical distance and linkage disequilibrium being influenced by a large number of factors. Association analysis of several markers in even a small region may each give very different results (29–33). The lower mutation rate of single nucleotide polymorphisms (SNPs), the initiative of the SNP Consortium (34) and the advent of novel high-throughput SNP typing technologies should soon make linkage disequilibrium mapping over a region of this size a feasible approach. Additional studies with a higher density of markers are required, and collection of singleton families by the Consortium is ongoing for replication of these results.

It is also of interest to note that in a recent study (35) the locus for a specific speech and language disorder (*SPCH1*) has been mapped to an ~6 Mb region of chromosome 7q. This region lies within the linkage interval identified in this study. A number of studies have suggested that specific language impairment and autism may have overlapping genetic aetiologies (19–21) and it is not yet clear whether the results presented here represent in part the same underlying genetic effect, with the operation of two, or possibly more, distinct genes. Better elucidation of the underlying autistic phenotype will make quantitative methods a more feasible and powerful approach, minimizing the influences of genetic heterogeneity.

In summary, this study further supports the existence of the autism susceptibility locus *AUTS1* on chromosome 7q. Inclusion of further families and markers implicates a broad chromosomal region and we are currently undertaking systematic screening of positional candidate genes for aetiological variants in autistic disorder.

MATERIALS AND METHODS

Families

The identification of families and assessment methods used by the IMGSAC consortium have been described in detail previously (4). Briefly, in families passing an initial screen, parents were administered the ADI-R (10) and the Vineland adaptive behaviour scales (36). Potential probands were assessed using the ADOS (11) or ADOS-G (12). When possible, psychometric evaluation was conducted using Ravens progressive matrices (37) or the Mullens (38) as appropriate, and the British picture vocabulary scale (39) or the Peabody picture vocabulary test (40) (or an appropriate translation). A physical examination was undertaken to exclude recognizable medical causes of autism, particularly tuberous sclerosis. A blood sample for DNA extraction was taken when possible from all probands and available first-degree relatives.

In addition, lymphoblastoid cell lines were generated from peripheral blood leukocytes to ensure a renewable source of DNA. In cases in which a blood sample could not be obtained, buccal swabs were taken. A sample was available from at least one parent in all families, and from both parents in >90% of families. Karyotyping was performed—or previous results obtained—when possible on all affected individuals, and molecular genetic testing for Fragile X performed or previous results obtained—on one case per family. Karyotype abnormalities were excluded in at least one case per family in 160 families and in both cases in 147 families, and Fragile X was excluded in one individual per family in 160 out of the170 families. There is no known overlap between the families included in this study and those ascertained by other research groups. Diagnostic criteria for the German singleton family sample have been described by Klauck *et al.* (41).

Genotyping

Genomic DNA was extracted from whole-blood, lymphoblastoid cell lines or buccal swabs using Nucleon kits. In cases where only a buccal swab was available, DNA was pre-amplified by primer extension pre-amplification (PEP) according to a technique modified from Zhang *et al.* (42). This method utilizes pre-amplification of genomic DNA with a random 15-mer, resulting in a 50–100-fold increase in template DNA for subsequent amplification. In our hands this has been found to preserve allelic ratios, supporting accurate genotyping. Polymerase chain reactions (PCRs), genotyping and analysis of microsatellite markers was carried out as described previously (4). Genomic DNA from blood was used for PCR when possible to minimize the rates of microsatellite mutation. Genbase (version 2.0.5, J. -M. Sebaoun and M. Lathrop) and Discovery Manager 2.3 (GENOMICA) were used to store all genotypic and phenotypic data and to produce the necessary files for statistical analysis.

Genetic markers

A total of 102 microsatellite markers were genotyped on chromosome 7. These markers included 81 published markers, the majority of which were dinucleotide repeat polymorphisms developed by Généthon (43) and tri- and tetra-nucleotide repeats from the co-operative human linkage centre (CHLC) (44). Ten novel dinucleotide repeat polymorphisms mapping in close proximity to six candidate genes in the region were isolated by PCR screening of Research Genetics human bacterial artificial chromosome (BAC) library release IV and hybridization to a labelled dinucleotide repeat primer using standard techniques. A further 11 novel microsatellites were identified electronically from available genomic sequence (http://www.genome.washington.edu/UWGC/). The average heterozygosity of all markers was 0.73. Marker order and genetic distances were determined from multiple sources including Généthon (43), the National Human Genome Research Institute (http://genome.nhgri.nih.gov/) and the University of Washington Genome Centre (http://www.genome.washington.edu/UWGC/). In cases in which the published physical marker map order differed from that of the genetic map, the physical map was used. Inter-marker distances were approximated from physical data under the assumption of 1 Mb = 1 cM. Marker order was also verified by mapping against yeast artificial chromosome (YAC) clones in the region (kindly provided by Dr Eric Green) (45). Primer sequences and Genbank accession numbers for the genomic sequence containing the novel repeat sequences are available from the IMGSAC web site (http://www.well.ox.ac.uk/~maestrin/iat.html).

Error detection

Maximum-likelihood reconstruction of multi-marker haplotypes was performed using GENEHUNTER version 2.0 and SIMWALK2 (46,47) to check for an excessive number of recombination events indicative of genotyping errors or marker mutation. Recombinant individuals were re-scored where necessary.

In addition, the program SIBMED (SIBpair Mutation and Error Detection) version 1.0 (48) was used to identify likely genotype errors or marker mutations in the multi-locus linkage data. This program computes the posterior probability of genotyping error or marker mutation for each sibling-pair-marker combination given all available marker data, an assumed genotype error rate and a known genetic map. This has been shown to be most advantageous for dense marker maps, recovering lost evidence for linkage masked by genotype errors, without generating false evidence for linkage when none is present. In order to determine marker-specific cutoff constants to identify unusually high posterior error probabilities, we used the Monte Carlo simulation option (100 000 simulations), a conservative assumed prior error rate of 1% and a false positive rate of 0.001. Seventeen likely errors were identified for the 25 markers used in the 5 cM map. All flagged individuals were subsequently untyped at a given marker locus. In cases in which multiple pairs of sibs within a family yielded a high posterior probability at the same marker, the sibling in common to these pairs was untyped at this marker.

Statistical analysis

Genotype data were analysed initially using SPLINK (16,17) to compute single-point maximum LOD scores for each marker maximized under the 'possible triangle' restrictions, and to calculate estimated marker heterozygosity and determination of effective sample size. Linkage analysis using information from multiple marker loci simultaneously was performed using ASPEX, allowing use of 147 nuclear families containing 153 non-independent affected sib-pairs. All ASPEX multipoint analyses were performed under an additive model and assuming no dominance variance. Determinations of parental sex-specific contributions to the linkage results were calculated using the 'sex_split' option of ASPEX. Multipoint analyses of the 23 extended families containing 25 non-independent affected relative-pairs was computed using the 'all' statistic of GENEHUNTER-PLUS (15) and the linear model of the KAC program (49). This modified score statistic has a standard normal distribution asymptotically, and has been shown to give less conservative results than GENEHUNTER for data that are not perfectly informative. In the three nuclear families and one extended family with three affected individuals, all possible affected pair combinations were used in the analyses.

Multi-allelic association tests for all markers on chromosome 7 were performed using the TDT (50), implemented in

ASPEX. This version of sib_tdt calculates empirical probabilities for χ^2 statistics by permuting parental alleles whilst fixing the IBD status of siblings within a family. These results therefore give an accurate measure of association independent of linkage allowing use of multiple siblings within a nuclear family (51,52). As this was an exploratory analysis, in order to minimize type II error the issue of testing of multiple marker loci was disregarded in favour of replication of any positive results in independent data sets. Data presented here are nominal *P* values, as it was considered that Bonferroni correction was too conservative both at the marker and at the allelic level for data that are not completely independent. The distribution of linkage disequilibrium over the region was studied using the standardized disequilibrium measure D', calculated using the program ARLEQUIN (23).

Electronic database information

The URLs for this article are as follows:
ASPEX program, ftp://lahmed.stanford.edu/pub/aspex;
SIBMED program, http://www.sph.umich.edu/group/statgen/software;
ARLEQUIN, http://anthropologie.unige.ch/arlequin;
University of Washington chromosome 7 sequencing data, http://www.genome.washington.edu/UWGC/;
NHGRI chromosome 7 mapping data, http://genome.nhgri.nih.gov/;
Electronic PCR, http://www.ncbi.nlm.nih.gov/genome/sts/epcr.cgi;
IMGSAC, http://www.ox.ac.uk/~maestrin/iat.html.

ACKNOWLEDGEMENTS

This work would not have been possible without the co-operation of both the subjects and their families and the many referring professionals. We also thank the computer and bioinformatics support staff at the Wellcome Trust Centre and the MRC Child Psychiatry Unit for help and assistance. This study is funded in part by support from the UK Medical Research Council, The Wellcome Trust, BIOMED 2 (CT-97-2759), EC Fifth Framework (QLG2-CT-1999-0094), Telethon, Italy (E.1007), the Janus Korczak Foundation, Deutsche Forschungsgemeinschaft, Fondation France Télécom, Conseil Régional Midi-Pyrénées, Danish Medical Research Council, Sofiefonden, the Beatrice Surovell Haskells Fond for Child Mental Health Research of Copenhagen, the Danish Natural Science Research Council (9802210), the National Institute of Child Health and Development (5-P01-HD-35482-02) and the National Institutes of Health (MO1 RR06022 GCRC NIH, NIH K05 MH01196, K02 MH01389). A.P.M. is a Wellcome Trust Principal Research Fellow.

REFERENCES

1. 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.
2. 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. Psychiatry*, **35**, 877–900.
3. Steffenberg, S., Gillberg, C., Hellgre, L., Anderen, L., Gillberg, I., Jackobsson, G. and Bohman, M. (1989) A twin study of autism in Denmark, Finland, Iceland, Norway and Sweden. *J. Child Psychol. Psychiatry*, **30**, 405–416.
4. 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.
5. Philippe, A., Martinez, M., Guilloud-Bataille, M., Gillberg, C., Rastam, M., Sponheim, E., Coleman, M., Zappella, M., Aschauer, H., Van Maldergem, L. *et al.* (1999) Genome-wide scan for autism susceptibility genes. Paris Autism Research International Sibpair Study. *Hum. Mol. Genet.*, **8**, 805–812.
6. Risch, N., Spiker, D., Lotspeich, L., Nouri, N., Hinds, D., Hallmayer, J., Kalaydjieva, L., McCague, P., Dimiceli, S., Pitts, T. *et al.* (1999) A genomic screen of autism: evidence for a multilocus etiology. *Am. J. Hum. Genet.*, **65**, 493–507.
7. Barrett, S., Beck, J.C., Bernier, R., Bisson, E., Braun, T.A., Casavant, T.L., Childress, D., Folstein, S.E., Garcia, M., Gardiner, M.B. *et al.* (1999) An autosomal genomic screen for autism. Collaborative linkage study of autism. *Am. J. Med. Genet.*, **88**, 609–615.
8. Ashley-Koch, A., Wolpert, C.M., Menold, M.M., Zaeem, L., Basu, S., Donnelly, S.L., Ravan, S.A., Powell, C.M., Qumsiyeh, M.B., Aylsworth, A.S. *et al.* (1999) Genetic studies of autistic disorder and chromosome 7. *Genomics*, **61**, 227–236.
9. Lamb, J.A., Moore, J., Bailey, A. and Monaco, A.P. (2000) Autism: recent molecular genetic advances. *Hum. Mol. Genet.*, **9**, 861–868.
10. 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. Autism Dev. Disord.*, **24**, 659–685.
11. Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L. and Schopler, E. (1989) Autism diagnostic observation schedule: a standardized observation of communicative and social behavior. *J. Autism Dev. Disord.*, **19**, 185–212.
12. Lord, C., Risi, S., Lambrecht, L., Cook, E.H., Jr, Leventhal, B.L., DiLavore, P.C., Pickles, A. and Rutter, M. (2000) The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *J. Autism Dev. Disord.*, **30**, 205–223.
13. 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.
14. Hinds, D. and Risch, N. (1996) The ASPEX package: affected sib-pair mapping. ftp://lahmed.stanford.edu/pub/aspex.
15. 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.
16. Holmans, P. and Clayton, D. (1995) Efficiency of typing unaffected relatives in an affected-sib-pair linkage study with single-locus and multiple tightly linked markers. *Am. J. Hum. Genet.*, **57**, 1221–1232.
17. Holmans, P. (1993) Asymptotic properties of affected-sib-pair linkage analysis. *Am. J. Hum. Genet.*, **52**, 362–374.
18. Feakes, R., Sawcer, S., Chataway, J., Coraddu, F., Broadley, S., Gray, J., Jones, H.B., Clayton, D., Goodfellow, P.N. and Compston, A. (1999) Exploring the dense mapping of a region of potential linkage in complex disease: an example in multiple sclerosis. *Genet. Epidemiol.*, **17**, 51–63.
19. Warburton, P., Baird, G., Chen, W., Morris, K., Jacobs, B.W., Hodgson, S. and Docherty, Z. (2000) Support for linkage of autism and specific language impairment to 7q3 from two chromosome rearrangements involving band 7q31. *Am. J. Med. Genet.*, **96**, 228–234.
20. Tager-Flusberg, H. and Folstein, S.E. (1999) Evidence for variation in the language phenotype for autism: implications for genetic subgroups. In DeLisi, L.E. (ed.), *The VIIth World Congress of Psychiatric Genetics*. Stockton Press, Monterey, CA, pp. S15.
21. Hafeman, L. and Tomblin J.B. (1999) Autism behaviours in the siblings of children with specific language impairment. In DeLisi, L.E. (ed.) *The VIIth World Congress of Psychiatric Genetics*. Stockton Press, Monterey, CA, pp. S14.
22. Schuler, G.D. (1997) Sequence mapping by electronic PCR. *Genome Res.*, **7**, 541–550.
23. Schneider, S., Roessli, D. and Excoffier, L. (2000) *Arlequin Version 2.000: A Software for Population Genetics Data Analysis*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.
24. Roberts, S.B., MacLean, C.J., Neale, M.C., Eaves, L.J. and Kendler, K.S. (1999) Replication of linkage studies of complex traits: an examination of variation in location estimates. *Am. J. Hum. Genet.*, **65**, 876–884.
25. Terwilliger, J.D., Shannon, W.D., Lathrop, G.M., Nolan, J.P., Goldin, L.R., Chase, G.A. and Weeks, D.E. (1997) True and false positive peaks in

genomewide scans: applications of length-biased sampling to linkage mapping. *Am. J. Hum. Genet.*, **61**, 430–438.
26. Goring, H.H. and Terwilliger, J.D. (2000) Linkage analysis in the presence of errors III: marker loci and their map as nuisance parameters. *Am. J. Hum. Genet.*, **66**, 1298–1309.
27. Kobayashi, S., Kohda, T., Miyoshi, N., Kuroiwa, Y., Aisaka, K., Tsutsumi, O., Kaneko-Ishino, T. and Ishino, F. (1997) Human PEG1/MEST, an imprinted gene on chromosome 7. *Hum. Mol. Genet.*, **6**, 781–786.
28. Blagitko, N., Schulz, U., Schinzel, A.A., Ropers, H.H. and Kalscheuer, V.M. (1999) γ2-COP, a novel imprinted gene on chromosome 7q32, defines a new imprinting cluster in the human genome. *Hum. Mol. Genet.*, **8**, 2387–2396.
29. Herr, M., Dudbridge, F., Zavattari, P., Cucca, F., Guja, C., March, R., Campbell, R.D., Barnett, A.H., Bain, S.C., Todd, J.A. and Koeleman, B.P. (2000) Evaluation of fine mapping strategies for a multifactorial disease locus: systematic linkage and association analysis of IDDM1 in the HLA region on chromosome 6p21. *Hum. Mol. Genet.*, **9**, 1291–1301.
30. Nair, R.P., Stuart, P., Henseler, T., Jenisch, S., Chia, N.V., Westphal, E., Schork, N.J., Kim, J., Lim, H.W., Christophers, E. *et al.* (2000) Localization of psoriasis-susceptibility locus PSORS1 to a 60-kb interval telomeric to HLA-C. *Am. J. Hum. Genet.*, **66**, 1833–1844.
31. Moffatt, M.F., Traherne, J.A., Abecasis, G.R. and Cookson, W.O. (2000) Single nucleotide polymorphism and linkage disequilibrium within the TCR α/δ locus. *Hum. Mol. Genet.*, **9**, 1011–1019.
32. Taillon-Miller, P., Bauer-Sardina, I., Saccone, N.L., Putzel, J., Laitinen, T., Cao, A., Kere, J., Pilia, G., Rice, J.P. and Kwok, P.Y. (2000) Juxtaposed regions of extensive and minimal linkage disequilibrium in human Xq25 and Xq28. *Nature Genet.*, **25**, 324–328.
33. Martin, E.R., Lai, E.H., Gilbert, J.R., Rogala, A.R., Afshari, A.J., Riley, J., Finch, K.L., Stevens, J.F., Livak, K.J., Slotterbeck, B.D. *et al.* (2000) SNPing away at complex diseases: analysis of single-nucleotide polymorphisms around APOE in Alzheimer disease. *Am. J. Hum. Genet.*, **67**, 383–394.
34. Marshall, E. (1999) Drug firms to create public database of genetic mutations. *Science*, **284**, 406–407.
35. Lai, C.S., Fisher, S.E., Hurst, J.A., Levy, E.R., Hodgson, S., Fox, M., Jeremiah, S., Povey, S., Jamison, D.C., Green, E.D. *et al.* (2000) The SPCH1 region on human 7q31: genomic characterization of the critical interval and localization of translocations associated with speech and language disorder. *Am. J. Hum. Genet.*, **67**, 357–368.
36. Sparrow, S., Balla, D. and Cicchetti, D. (1984) *Vineland Adaptive Behaviour Scales*. American Guidance Service, Circle Pines, MN.
37. Raven, J. (1989) *Standard Progressive Matrices*. Australian Council for Educational Research, Victoria, Australia.
38. Mullen, E.M. (1995) *Mullens Scales of Early Learning*. American Guidance Service, Circle Pines, MN.
39. Dunn, L.M., Dunn, L.M., Whetton, C. and Pintilie, D. (1982) *British Picture Vocabulary Scale*. NFER-Nelson, Windsor, UK.
40. Dunn, L.M. and Dunn, L.M. (1997) *Peabody Picture Vocabulary Test*, 3rd edn. American Guidance Service, Circle Pines, MN.
41. Klauck, S.M., Munstermann, E., Bieber-Martig, B., Ruhl, D., Lisch, S., Schmotzer, G., Poustka, A. and Poustka, F. (1997) Molecular genetic analysis of the FMR-1 gene in a large collection of autistic patients. *Hum. Genet.*, **100**, 224–229.
42. Zhang, L., Cui, X., Schmitt, K., Hubert, R., Navidi, W. and Arnheim, N. (1992) Whole genome amplification from a single cell: Implications for genetic analysis. *Proc. Natl Acad. Sci. USA*, **89**, 5847–5851.
43. Dib, C., Faure, S., Fizames, C., Samson, D., Drouot, N., Vignal, A., Millasseau, P., Marc, S., Hazan, J., Seboun, E. *et al.* (1996) A comprehensive genetic map of the human genome based on 5, 264 microsatellites. *Nature*, **380**, A1–A138.
44. Murray, J.C., Buetow, K.H., Weber, J.L., Ludwigsen, S., Scherpbier Haddema, T., Manion, F., Quillen, J., Sheffield, V.C., Sunden, S., Duyk, G.M. *et al.* (1994) A comprehensive human linkage map with centimorgan density. *Science*, **265**, 2049–2054.
45. Bouffard, G.G., Idol, J.R., Braden, V.V., Iyer, L.M., Cunningham, A.F., Weintraub, L.A., Touchman, J.W., Mohr-Tidwell, R.M., Peluso, D.C., Fulton, R.S. *et al.* (1997) A physical map of human chromosome 7: an integrated YAC contig map with average STS spacing of 79 kb. *Genome Res.*, **7**, 673–692.
46. 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.
47. 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.
48. Douglas, J.A., Boehnke, M. and Lange, K. (2000) A multipoint method for detecting genotyping errors and mutations in sibling-pair linkage data. *Am. J. Hum. Genet.*, **66**, 1287–1297.
49. Kong, A. and Cox, N.J. (1997) Allele-sharing models: LOD scores and accurate linkage tests. *Am. J. Hum. Genet.*, **61**, 1179–1188.
50. 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.
51. Lazzeroni, L.C. and Lange, K. (1998) A conditional inference framework for extending the transmission/disequilibrium test. *Hum. Hered.*, **48**, 67–81.
52. Martin, E.R., Kaplan, N.L. and Weir, B.S. (1997) Tests for linkage and association in nuclear families. *Am. J. Hum. Genet.*, **61**, 439–448.