European Journal of Human Genetics (2010) 18, 1013–1020

www.nature.com/ejhg

ARTICLE

Linkage and candidate gene studies of autism spectrum disorders in European populations

Richard Holt[1], Gabrielle Barnby[1], Elena Maestrini[2], Elena Bacchelli[2], Denise Brocklebank[1], Inês Sousa[1], Erik J Mulder[3], Katri Kantojärvi[4], Irma Järvelä[4], Sabine M Klauck[5], Fritz Poustka[6], Anthony J Bailey[7] and Anthony P Monaco*[,1], the EU Autism MOLGEN Consortium[8]

Over the past decade, research on the genetic variants underlying susceptibility to autism and autism spectrum disorders (ASDs) has focused on linkage and candidate gene studies. This research has implicated various chromosomal loci and genes. Candidate gene studies have proven to be particularly intractable, with many studies failing to replicate previously reported associations. In this paper, we investigate previously implicated genomic regions for a role in ASD susceptibility, using four cohorts of European ancestry. Initially, a 384 SNP Illumina GoldenGate array was used to examine linkage at six previously implicated loci. We identify linkage approaching genome-wide suggestive levels on chromosome 2 (rs2885116, MLOD=1.89). Association analysis showed significant associations in *MKL2* with ASD (rs756472, $P=4.31\times10^{-5}$) and between *SND1* and strict autism (rs1881084, $P=7.76\times10^{-5}$) in the Finnish and Northern Dutch populations, respectively. Subsequently, we used a second 384 SNP Illumina GoldenGate array to examine the association in seven candidate genes, and evidence for association was found in *RELN* (rs362780, $P=0.00165$). Further increasing the sample size strengthened the association with *RELN* (rs362780, $P=0.001$) and produced a second significant result in *GRIK2* (rs2518261, $P=0.008$). Our results strengthen the case for a more detailed study of the role of *RELN* and *GRIK2* in autism susceptibility, as well as identifying two new potential candidate genes, *MKL2* and *SND1*.

European Journal of Human Genetics (2010) **18,** 1013–1019; doi:10.1038/ejhg.2010.69; published online 5 May 2010

Keywords: autistic disorder; linkage; association; candidate gene

INTRODUCTION

Autism, a neuropsychiatric disorder with an onset before 3 years of age, is characterised by impaired reciprocal communication and social interaction, as well as by restricted and stereotyped patterns of interests and behaviour. The definition can be further broadened to include atypical autism, Asperger's syndrome and pervasive developmental disorder not otherwise specified (PDD-NOS), to create a class of conditions collectively referred to as autism spectrum disorders (ASDs) (OMIM %209850).

ASDs affect approximately 0.6–1.2% of the general population,[1,2] with a marked excess of boys compared with girls of ~4:1.[3,4] Multiple lines of evidence have shown that autism has a large genetic component. The prevalence of ASDs is increased to 2–8% among siblings of affected individuals,[4,5] and the concordance rates increase from 0% in same-sex dizygotic twins to 36–60% in monozygotic pairs.[6,7] Thus, the heritability of autism is ~90%, making it the most heritable of the childhood onset neuropsychiatric disorders.[7]

Despite the obvious importance of genetic factors in autism development, the search for genes underlying susceptibility has met with limited success. A large number of linkage studies have been conducted and have identified possible susceptibility loci on multiple chromosomes.[8] Although there is not total concordance between the different studies, certain regions, such as those on chromosomes 2, 3, 7, 11, 16, 17 and 19, have been implicated multiple times. Candidate gene studies have been used as an alternative approach for identifying variants increasing susceptibility to autism. Between 1995 and 2008, nearly 200 genes were investigated for association with autism, with more than 80 reported with nominally positive results. However, even the most frequently associated genes, such as *RELN* (*reelin*),[9] have negative replications reported.[10] Recently, the first genome-wide association study for ASDs has been published, implicating a number of genes and genomic regions, most significantly the area between cadherin genes *CDH9* (*cadherin 9*) and *CDH10* (*cadherin 10*) on chromosome 5p14.1. However, in this study, only a single SNP reached genome-wide significance after correction for multiple testing, confirming the difficulties involved in the identification of common variants contributing to ASD susceptibility.[11]

We attempted to refine our understanding of the genetics of autism by investigating previously reported linkage and candidate gene results in our European populations. Our collaboration includes groups from the International Molecular Genetic Study of Autism Consortium (IMGSAC), Paris Autism Research International Study (PARIS) and Finland. These three groups have performed whole-genome linkage screens for autism in their respective family collections,[12–15] with the

[1]The Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford, UK; [2]Dipartimento di Biologia, Università di Bologna, Bologna, Italy; [3]Department of Psychiatry, Child and Adolescent Psychiatry, University Medical Center Groningen, Groningen, The Netherlands; [4]Department of Medical Genetics, University of Helsinki, Helsinki, Finland; [5]Division of Molecular Genome Analysis, German Cancer Research Center, Heidelberg, Germany; [6]Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, Johann Wolfgang Goethe-University, Frankfurt/Main, Germany; [7]University Department of Psychiatry, Park Hospital for Children, Oxford, UK
*Correspondence: Professor AP Monaco, The Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK. Tel: +44 1865270004; Fax: +44 1865280411; E-mail: anthony.monaco@well.ox.ac.uk
[8]See Appendix
Received 4 September 2009; revised 7 April 2010; accepted 8 April 2010; published online 5 May 2010

six highest LOD scores on chromosomes 2q, 3q, 6q, 7q, 16p and 17q. In addition, IMGSAC and PARIS are members of the Autism Genome Project (AGP) consortium, which recently conducted a genome-wide linkage study using Affymetrix (Santa Clara, CA, USA) 10k arrays, including a large number of samples from these two consortia.[2] To refine previously identified linkage peaks, we chose SNPs from the Affymetrix 10k array that tag variation within these top six loci and genotyped them in additional IMGSAC, PARIS and Finnish families. These results were analysed in combination with overlapping SNPs from the AGP study for IMGSAC and PARIS samples in a meta-analysis of linkage. Furthermore, we genotyped these SNPs in trios from Finnish and Northern Dutch isolated populations to identify regions of extended linkage disequilibrium (LD) within these loci.

Second, we attempted to identify or replicate associations of autism and ASD with seven key candidate genes. These were *NOSTRIN* (*nitric oxide synthase trafficker*),[16] *GRIK2* (*glutamate receptor, ionotropic, kainite 2*),[17,18] *RELN*,[9] *PRKCB1* (*protein kinase C, beta*),[19,20] *SLC6A4* (*solute carrier family 5 (neurotransmitter transporter, serotonin), member 4*),[21,22] *SHANK3* (*SH3 and multiple ankyrin repeat domains 3*)[2,23] and *ASMT* (*acetylserotonin O-methyltransferase*).[24]

METHODS

Samples

The individuals included in this study came from four European based collections/populations: IMGSAC, PARIS, Finland and Northern Holland. Assessment and selection criteria have been described previously for IMGSAC,[14] PARIS,[25] Finnish[12] and the Northern Dutch[22] probands. All affected individuals from the IMGSAC, PARIS and Northern Dutch cohorts had ADI-R[26] and/or ADOS[27] assessments. The Finnish probands were identified by ICD-10 and DSMIV criteria diagnosed by experienced clinicians. General summary phenotypic information, where available, is provided (Supplementary Table 1). For Illumina GoldenGate (Illumina, San Diego, CA, USA) genotyping, the samples were either genomic DNA with a concentration of 60–100 ng/μl ($\sim$26.4%) or whole-genome amplified DNA (GenomiPhi v2 (GE Healthcare, Amersham, UK)), with an estimated concentration of 200 ng/μl ($\sim$73.6%). For Sequenom iPLEX (Sequenom, San Diego, CA, USA)-based genotyping, samples were diluted to 10 ng/μl for genomic DNA, or to 1:10 ($\sim$20 ng/μl) for whole-genome-amplified DNA.

SNP selection and genotyping

Separate Illumina GoldenGate 384 SNP arrays were designed for linkage and candidate gene studies.

For the linkage array, a subset of SNPs from the Affymetrix 10k v2 SNP array located within previously identified regions of linkage on chromosomes 2q, 3q, 6q, 7q, 16p and 17q were identified. Data for these SNPs were downloaded for the HapMap Phase II (release 21) CEU samples. Those SNPs tagging variation in these regions with a minimum allele frequency (MAF) of $\geq$0.05 and $r^2 \geq 0.8$ were selected for genotyping using the Tagger program in Haploview v4.0.[28] In addition, four SNPs on the X chromosome were included to confirm the sex of new individuals. A total of 93 multiplex families were genotyped (44 IMGSAC, 16 PARIS and 33 Finnish). An additional 248 families (66 Finnish singletons, 182 Northern Dutch trios, the majority from singleton families) were also genotyped to examine the extent of LD within these regions in relatively isolated populations. In total, 1127 individuals from 341 families were genotyped on this array (Supplementary Table 2).

For the candidate gene array, SNPs within and 5 kb 5′ and 3′ of *NOSTRIN*, *GRIK2*, *RELN*, *PRKCB1*, *SLC6A4*, *SHANK3* and *ASMT* (NCBI build 35) were downloaded for the HapMap Phase II (release 21) CEU samples, and those tagging variation in these regions with a MAF $\geq$0.05 and $r^2 \geq 0.8$ using the Tagger program in Haploview v4.0[28] were chosen. For SNPs tagging more than 10 others, a second 'safety net' SNP was chosen to capture identical variation, accounting for 34 SNPs on the array. Additional SNPs of interest were also chosen from the published literature. Samples genotyped on the array consisted of trios of one randomly chosen affected individual, and both parents, where available, from multiplex families. In total, 1144 samples from 389 families (284 IMGSAC, 72 PARIS and 33 Finnish) were genotyped (Supplementary Table 3).

Genotyping for both arrays was performed using the standard protocols for GoldenGate assays (Illumina).

There is an overlap of the samples genotyped in this study with those used to replicate recent genome-wide association results.[11,29] However, in the latter studies, the SNPs chosen were to replicate individual signals, in contrast to our haplotype-tagging approach to investigate association across specific candidate genes.

Analysis

Genotypes were called using BeadStudio software (Illumina) with manual editing of clusters. Mendelian inheritance errors were identified and removed using PedCheck.[30] SNPs with poor clustering or $>$10 Mendelian errors, and samples with a genotyping success rate $<$80% or $>$10 Mendelian errors, were excluded from further analysis. All SNPs were tested for Hardy–Weinberg equilibrium using the BeadStudio software (Illumina).

Before analysing linkage data, genotypes for 262 IMGSAC and 49 PARIS families genotyped by the AGP using the Affymetrix 10k v2 SNP array were downloaded from the AGP database (http://davinci.tcag.ca/agp_freeze2/). Data from the 384 SNPs used in this study were combined with those we generated to create an extended cohort of 397 families (301 IMGSAC, 64 PARIS, 32 Finnish). Data were formatted using MEGA2 before linkage analysis using the ASPEX package (ftp://lahmed.stanford.edu/pub/aspex). Parent-of-origin analyses were performed using the sex_split option of the ASPEX package. LOD scores $\geq$1.9 and $\geq$3.3 were considered to be evidence of suggestive and significant linkage, respectively.[31]

Association analyses were performed using the transmission disequilibrium test (TDT), which is robust against population stratification. The candidate gene association data were analysed for each individual population genotyped and the results were combined in a meta-analysis. We estimated odds ratios and the respective SE from transmission frequencies.[32] Meta-analysis was performed by pooling all data using an inverse variant approach.[33] For all association analyses, we considered each gene or region examined as a separate hypothesis. Therefore, we determined a Bonferroni-corrected significance threshold by dividing the nominal significance threshold (P=0.05) by the number of SNPs tested at that locus. Reducing the LD threshold from r^2=0.8 to r^2=0.2 does not significantly reduce the number of independent SNPs, and therefore the Bonferroni correction could be applied. For each linkage region in which we examined association, we further divided the significance threshold by 2 because these loci were tested for association with both ASD and strict autism.

The number and density of SNPs on either array were insufficient to perform a comprehensive analysis for the presence of copy number variants at the loci investigated.

Replication

The five most significant SNPs from the candidate gene array were genotyped in an additional 282 IMGSAC singleton families and in 188 Northern Dutch trios, the majority of whom were also from singleton families. Sequences flanking the SNPs were obtained (http://www.Biomart.com) and iPLEX assays were designed (http://www.realSNP.com). Genotyping was performed using the Sequenom iPlex platform (Sequenom), with standard protocols. Power calculations for the replication sample were performed using the Genetic Power Calculator for discrete trait TDT (GPC) (http://pngu.mgh.harvard.edu/~purcell/gpc).[34] Parameters used for the GPC were 0.17% for disease prevalence, perfect LD between tested marker and disease allele, an additive model and a type 1 error rate of 0.05.

RESULTS

Array performance

Both Illumina GoldenGate 384 SNP genotyping arrays performed well. From the linkage array, a total of 379 SNPs and 1112 samples survived quality control, with a mean genotyping success rate of $>$99% for the retained samples (Supplementary Table 2). For the candidate gene array, after quality control, 354 SNPs and 1127 samples

remained for analysis, resulting in a mean genotyping success rate of >99% after quality control (Supplementary Table 3).

Linkage array

Linkage analysis. The combination of our data with those from the AGP study[2] resulted in a final sample set of 397 families (301 IMGSAC, 64 PARIS and 32 Finnish). Multipoint linkage analysis was performed on the sibpairs. The largest signal obtained was on chromosome 3, with the peak at rs2862479 (MLOD=1.5) (Figure 1). The Finnish cohort is a relatively genetically isolated population, suggesting that it may contain different susceptibility loci to either the IMGSAC or PARIS cohorts. Therefore, we repeated the analysis excluding the 32 Finnish families, which decreased the signal on chromosome 3 (peak MLOD=0.79, rs2862479). A signal approaching suggestive levels of linkage[31] was observed on chromosome 2 in the remaining families (peak MLOD=1.89, rs2885116) (Figure 1).

Parent-of-origin analysis. It has previously been shown that parent-of-origin effects may have an important role in ASD susceptibility.[35] Parent-of-origin analysis for our entire data set was performed, resulting in an increased peak signal on chromosome 3 for maternal alleles, with the zenith shifting 5′ from rs2862479 to rs4129157 (MLOD=1.78). The analysis was repeated excluding the Finnish subset of individuals, resulting in an increased signal on chromosome 2 spanning four SNPs (rs726032, rs726033, rs1374431 and rs2885116), each with an LOD score of 1.67 for the maternal alleles.

Association analysis in published regions of linkage. Two sets of trios (62 Finnish and 179 Northern Dutch after quality control), mainly consisting of singleton families, were genotyped on the linkage array. Analysis of the LD patterns of the SNPs for these populations showed no significant difference from that of the CEU samples in the HapMap data (data not shown). Therefore, a TDT test was carried out using these data for both strict definition autism and ASD, including an additional 33 Finnish multiplex families we had genotyped (Finnish strict N=80, ASD N=95; Northern Dutch strict N=99, ASD N=179) (Figure 2). In total, 375 SNPs in the linkage regions survived quality control (chromosome 2 N=142, chromosome 3 N=22, chromosome 6 N=55, chromosome 7 N=100, chromosome 16 N=21, chromosome 17 N=35). The Bonferroni-corrected significance thresholds for each locus were as follows: chromosome 2 $P \leq 0.000176$, chromosome 3 $P \leq 0.00114$, chromosome 6 $P \leq 0.000455$, chromosome 7 $P \leq 0.00025$, chromosome 16 $P \leq 0.00119$ and chromosome 17 $P \leq 0.00074$. Association was found on chromosome 7 with a SNP within the gene *SND1* (*Staphylococcal nuclease* and *tudor domain containing 1*) in Northern Dutch individuals with strictly defined autism (rs1881084, $P=7.76\times10^{-5}$), and to a lesser extent in the wider Northern Dutch ASD cohort (P=0.001838), but this association was not present in the Finnish population. A significant association was observed on chromosome 16 for rs756472 in the gene *MKL2* (*MKL/myocardin-like 2*) in the Finnish population, for both strict autism and ASD (strict autism, $P=2.46\times10^{-4}$; ASD, $P=4.31\times10^{-5}$). However, this association was not observed in the Northern Dutch population.

After identifying these associations, analysis was performed on the largest individual cohort of samples from IMGSAC (N=301). Although nominal associations were found on chromosomes 2, 6, 7, 16 and 17, none of these were with rs1880184 or rs756472.

Candidate gene array

A TDT was performed for each cohort genotyped on the candidate gene array for both strict autism and ASD. These results were combined for the 382 families in total by performing a pooled meta-analysis (Figure 3). In total, excluding safety net SNPs, 328 SNPs within the candidate genes survived quality control (*NOSTRIN* N=15, *GRIK2* N=105, *RELN* N=118, *PRKCB1* N=72, *SLC6A4* N=7, *SHANK3* N=9, *ASMT* N=2). The Bonferroni-corrected significance thresholds for each locus were as follows: *NOSTRIN* $P \leq 0.00333$, *GRIK2* $P \leq 0.000476$, *RELN* $P \leq 0.000424$, *PRKCB1* $P \leq 0.000694$, *SLC6A4* $P \leq 0.00714$, *SHANK3* $P \leq 0.00556$ and *ASMT* $P \leq 0.025$. The two strongest associations, obtained for both strict autism and ASD, were with rs362780 in *RELN* (strict autism P=0.00165, ASD P=0.00165) and rs2518261 in *GRIK2* (strict autism P=0.00955, ASD

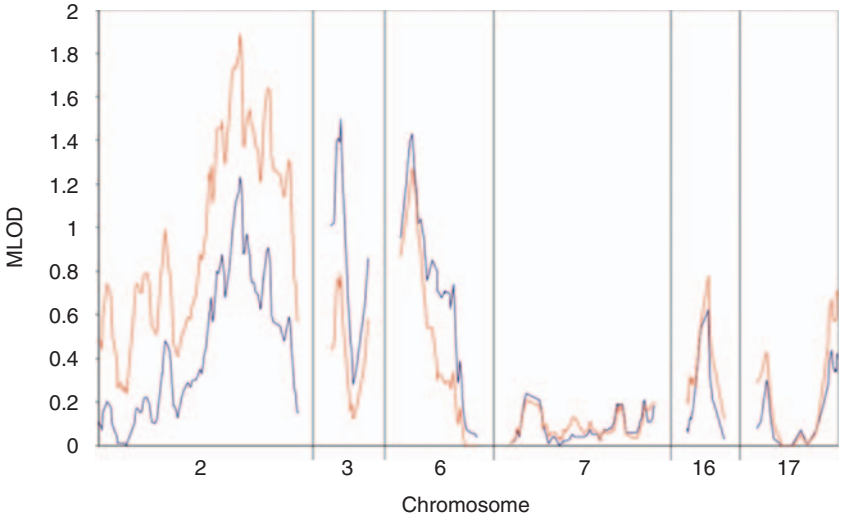

Figure 1 Results of linkage analyses of chromosomes 2, 3, 6, 7, 16 and 17. The blue line indicates results for the combined set of IMGSAC, PARIS and Finnish families, whereas the red line shows results when Finnish samples are excluded. Each locus is shown to the same scale. The regions covered are as follows (NCBI Build 36.1): chromosome 2, rs2320399 (133 458 918) → rs1020941 (194 183 515), 60 724 597 bp; chromosome 3, rs721729 (174 671 967) → rs725656 (186 008 902), 11 336 935 bp; chromosome 6, rs1590957 (89 867 130) → rs1378702 (113 107 351), 23 240 221 bp; chromosome 7, rs726820 (93 449 846) → rs273937 (137 230 419), 43 780 573 bp; chromosome 16, rs1424125 (6 002 554) → rs722075 (17 242 000), 11 239 446 bp; chromosome 17, rs719601 (25 755 541) → rs1990673 (50 405 024), 24 649 483 bp.

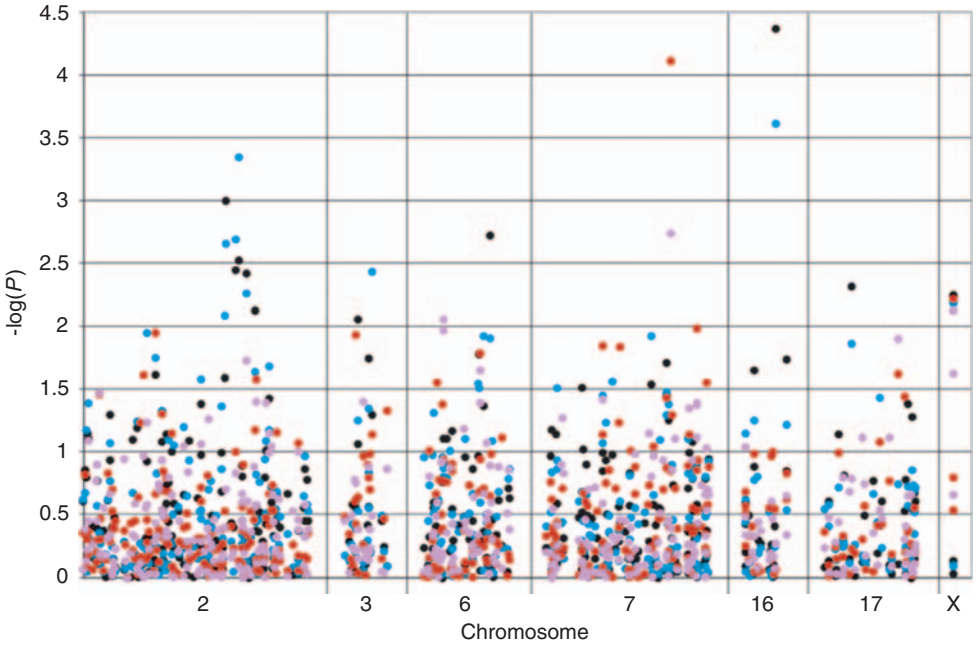

Figure 2 TDT association analysis of SNPs in previously reported regions of linkage using Finnish and Northern Dutch families. Association was tested for both strict autism and the broader phenotype of ASD in the Finnish and Northern Dutch families separately. Results are depicted as follows: Finnish strict autism=blue, Finnish ASD=black, Northern Dutch strict autism=red, Northern Dutch ASD=purple. Results are plotted as $-\log(P)$. Each locus is shown to the same scale. The regions covered are as follows (NCBI Build 36.1): chromosome 2, rs2320399 (133 458 918) → rs1020941 (194 183 515), 60 724 597 bp; chromosome 3, rs721729 (174 671 967) → rs725656 (186 008 902), 11 336 935 bp; chromosome 6, rs1590957 (89 867 130) → rs1378702 (113 107 351), 23 240 221 bp; chromosome 7, rs726820 (93 449 846) → rs273937 (137 230 419), 43 780 573 bp; chromosome 16, rs1424125 (6 002 554) → rs722075 (17 242 000), 11 239 446 bp; chromosome 17, rs719601 (25 755 541) → rs1990673 (50 405 024), 24 649 483 bp.

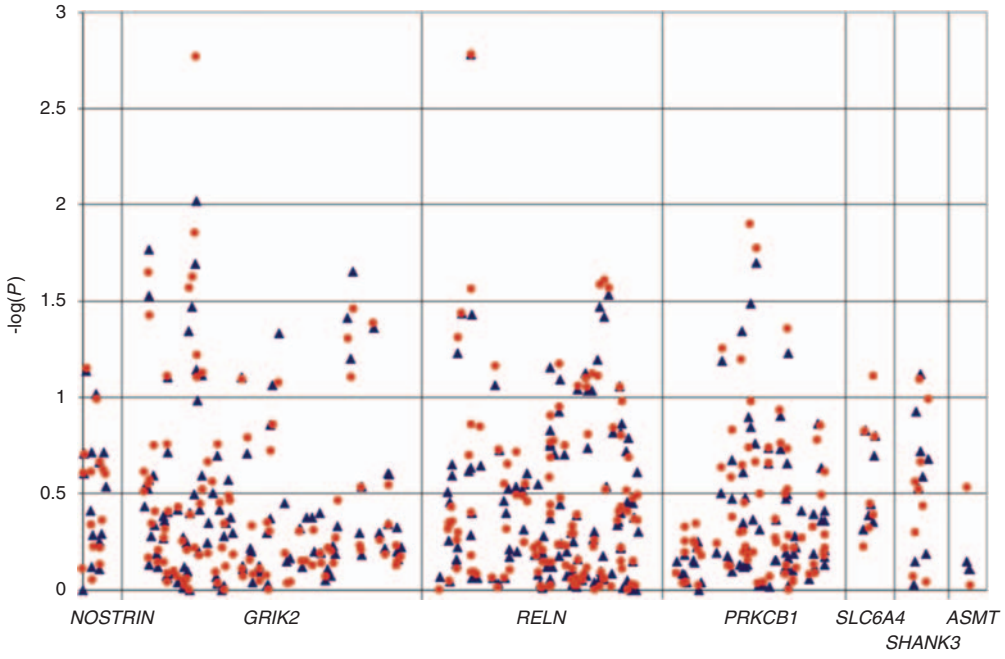

Figure 3 TDT meta-analysis of Illumina GoldenGate candidate gene data. Association with strict autism and the wider ASD phenotype was performed separately using combined data for the IMGSAC, PARIS and Finnish samples. Results for strict autism are plotted as blue triangles and for ASD as red circles. Results are plotted as $-\log(P)$. Each locus is shown to the same scale. The regions covered are as follows (NCBI Build 36.1): *NOSTRIN*, rs6433093 (169 367 190) → rs12993143 (169 427 746), 60 556 bp; *GRIK2*, rs2852512 (101 953 275) → rs2852620 (102 621 626), 668 351 bp; *RELN*, rs1978198 (102 896 685) → rs4298437 (103 413 113), 516 428 bp; *PRKCB1*, rs3760106 (23 753 297) → rs3729908 (24 138 749), 385 452 bp; *SLC6A4*, rs3813034 (25 548 930) → rs16965628 (25 579 551), 30 621 bp; *SHANK3*, rs2341009 (49 480 446) → rs756638 (49 518 559), 38 113 bp; *ASMT*, rs6588807 (1 708 581) → rs5949028 (1 715 666), 7085 bp.

Table 1 TDT results for rs2518256, rs2518261, rs362780, rs9925126 and rs11074601

Gene	Chr	SNP	Location	Position	Finnish/PARIS/ IMGSAC (N=*389*)	IMGSAC Only (N=*284*)	P-value Replication (N=*464*)	Combined data (N=*853*) OR	LCL	UCL	P-value
GRIK2	6	rs2518256	Intron 1	102 085 492	0.014	0.137	0.67	1.166	0.944	1.44	0.154
		rs2518261	Intron 1	102 088 471	0.002	0.015	0.702	0.779	0.648	0.936	0.008
RELN	7	rs362780	Intron 39	102 977 594	0.002	0.016	0.115	0.554	0.387	0.794	0.001
PRKCB1	16	rs9925126	Intron 3	23 945 617	0.013	0.061	0.162	1.081	0.923	1.267	0.334
		rs11074601	Intron 5	23 961 192	0.017	0.006	0.59	0.932	0.807	1.075	0.331

Chr, chromosome; LCL, lower confidence interval; OR, odds ratio; UCL, upper confidence interval.
SNP positions are given according to dbSNP build 129 (NCBI). Numbers of families genotyped for each analysis are given (*N*).

P=0.00170). However, neither of these associations passed the thresholds corrected for multiple testing.

The majority of trios genotyped on the candidate gene array were from the IMGSAC population (*N*=279). It was of interest that the strongest association obtained in this cohort was for rs11074601 in *PRKCB1* (*P*=0.00596). However, this SNP did not reach our corrected significance threshold.

Replication of candidate gene association. To replicate the results for the five SNPs showing the strongest association with ASD in the meta-analysis, rs2518256 (*P*=0.01399), rs2518261 (*P*=0.00170), rs362780 (*P*=0.00165), rs9925126 (*P*=0.01260) and rs11074601 (*P*=0.01693) (Table 1), additional Northern Dutch (*N*=188) and IMGSAC families (*N*=276) were genotyped using the Sequenom iPLEX platform. Power calculation analysis showed that our replication sample should give sufficient power (>78%) to replicate the results of these five SNPs. However, no significant associations were observed in the individual replication populations (data not shown) or when combined in the meta-analysis (Table 1).

A TDT meta-analysis of all available data from the original candidate gene array and the replication set was performed in a total of 853 families genotyped (IMGSAC *N*=560, Northern Dutch *N*=188, Finnish *N*=33, PARIS *N*=72). Significant associations were obtained for rs25185261 (*GRIK2* *P*=0.008) and rs362780 (*RELN* *P*=0.001) (Table 1). However, although the association observed for rs362780 increased, it failed to reach the corrected significance threshold.

DISCUSSION

Although there is considerable evidence for a strong genetic component underlying autism susceptibility, there has been slow progress towards identifying risk loci involved with a large degree of confidence. Studies have struggled to replicate positive results for linkage regions and individual candidate genes. We have investigated a set of the most commonly implicated linkage regions and candidate genes for significance in autism cohorts of European descent.

The six regions of the genome we investigated had previously been identified by members of our consortium as showing linkage with autism or ASD. Our results failed to identify loci with LOD scores reaching the levels required for evidence of linkage. A peak was found on chromosome 2 fractionally below the Lander and Kruglyak[31] cutoff for suggestive evidence of linkage. This result was obtained only after the subset of Finnish families, accounting for ~8% of our total available sample, had been removed from the analysis. The increase in signal from chromosome 2 is encouraging and we take this to strengthen support for an autism susceptibility locus on this chromosome, which has been implicated in a number of previous studies. Our results localise the peak with LOD $\geq$1.5 distal to the area of the region genotyped, between SNPs rs2161994 and rs1861896. Although not directly under the strongest point of linkage, it is interesting to note that the gene *ZNF804A* lies in the proximal region of this peak. A SNP in this gene has recently been found to be strongly associated with schizophrenia and bipolar disorder.[36] Gain-of-copy number CNVs containing this gene have also been identified in three affected individuals in a whole-genome study by the AGP,[2] and in addition, a case of autism with a translocation affecting chromosome 2q32.1 has been reported.[37,38] Our parent-of-origin analysis indicated that this signal is being driven mainly by maternal inheritance. This is in contrast to a study by Lamb *et al*,[35] which found an approximately equal contribution of alleles from both parents. In addition, our results indicate that the linkage observed on chromosome 3 is also maternal, although the top LOD score (rs4129157, LOD =1.78) does not reach suggestive levels. These results hint at a role for imprinting at these loci. Imprinting is known to be of importance in ASD, as shown by a significant number of cases being due to inheritance of maternal duplications of the 15q11–13 locus, which includes imprinted genes.[39]

Association analysis of SNPs in regions of previously reported linkage was performed in the Finnish and Northern Dutch populations. The SNPs had been chosen to tag variation within the regions examined with $r^2>0.8$. Therefore, although the SNP coverage was not dense, it should be sufficient to examine association. Our results identified two SNPs with associations surviving Bonferroni correction. The strongest association was with rs756472 on chromosome 16, a SNP located in the gene *MKL2*, in the Finnish cohort. To our knowledge, this SNP has only been investigated once previously for its role in autism, with no association found.[40] A SNP on chromosome 7, rs1881084, was also found to be associated in the Northern Dutch cohort. This SNP lies within intron 7 of *SND1*, a highly conserved transcriptional coactivator encoding one of the proteins comprising the RNA-induced silencing complex.[41] Interestingly, it has been found that the localisation of E-cadherin changes from the cell membrane to the cytoplasm of mouse cells with high levels of snd1 protein.[41] CNVs in protocadherin genes *PCDH9*[42] and *PCDH10*[43] have been implicated in autism, in addition to *PCDH8* in schizophrenia.[44] Cadherins also mediate cell–cell neural interactions and may have an important part in neural development.[41] Therefore, *SND1* may be a good candidate for further study of its role in autism development.

Moreover, the presence of the leucine-rich repeat (LRR) gene *LRRC4* (*LRR containing 4*) within intron 16 of *SND1* must also be noted. Suggestive association was found by Wang *et al*[11] in two LRR

genes, *LRRC1* (*LRR containing 1*) and *LRFN5* (*LRR and fibronectin type III domain containing 5*), and recent work in our own group has found association with two further LRR genes, *LRRTM3* (*LRR transmembrane neuronal 3*) and *LRRN3* (*LRR neuronal 3*), the latter association also being found in the Northern Dutch cohort. However, it should be noted that two SNPs located between *LRRC4* and rs1881084 failed to show any association in the Northern Dutch cohort (rs1419970 P=0.0506; rs178733 P=0.713). In addition, no association with either *SND1* or *MKL2* was found in the IMGSAC families. As both Finnish and Dutch cohorts are from relatively isolated populations, this finding may indicate the importance of particular genes or variants for autism and ASD susceptibility in different populations. Alternatively, it may represent differences in assessment of the various cohorts, or the distribution of particular phenotypes in their members (Supplementary Table 1).

In keeping with the trend of many such studies on autism, our candidate gene analysis found no associations that survived Bonferroni correction in the original set of samples genotyped. Similar to the linkage analysis, this finding may be due to heterogeneity within our populations. In addition, it may be that, although the results do not pass strict correction for multiple testing, they may represent, along with other studies, individual genes that are a part of the same network contributing to ASD susceptibility. However, we did identify several interesting nominal associations that, given the stringency of the Bonferroni method, were of sufficient interest to warrant further investigation. TDT analysis of additional European samples for the five most significant SNPs genotyped also failed to find association. However, when a meta-analysis was performed for the combined set of original and replication samples, two significant associations were observed, one with rs362780 in *RELN* and the other with rs2518261 in *GRIK2*.

RELN is located in the region of linkage for autism on chromosome 7,[14] and has been repeatedly studied as a candidate gene for autism. Evidence has been published supporting[9] and rejecting[10] the association of *RELN* with autism (reviewed by Freitag[8]). Functionally, *RELN* is considered as a good candidate gene because of its involvement in neuronal migration, and mice lacking its expression share regions of brain alteration with autistic individuals.[9] The association of this gene in our initial and extended populations adds evidence implicating *RELN* in ASD.

GRIK2 is also an interesting candidate because of its location in a region of linkage for autism on 6q21.[15] Moreover, glutamatergic neurons originate in brain regions implicated in autism, and glutamate antagonists can cause symptoms similar to those of autism.[45] Four studies have specifically looked for association between *GRIK2* and autism, with three finding positive evidence[17,18,46] and one study failing to do so.[47]

In their recent genome-wide association paper, Wang *et al*[11] report top associations ($P<0.01$) for 26 candidate genes. Their reported SNPs include 10, the closest gene to which is *GRIK2*, and of these, two SNPs lie within *GRIK2* itself (rs4839797, intron 1, P=0.003978 and rs2782908, intron 13, P=0.005692). Although we did not tag either of these SNPs in our array, it is of interest that our significant association with *GRIK2* also occurs in intron 1 of the gene. Our findings add to the cumulative case for a role for this gene in autism susceptibility.

Our data further implicate the chromosome 2 linkage region, as well as *RELN* and *GRIK2*, in autism susceptibility. Despite the difficulties in confirming these findings, further studies of these loci are warranted. We have also identified *MKL2* and *SND1* as potentially interesting new candidates for further investigation.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank the families who have participated in the study and Chris Allan at the WTCHG core genomics facility for technical support. This paper includes data generated by the Autism Genome Project (AGP) (see supplementary information for full AGP authorship list.) The AGP gratefully acknowledges the contributions of the families who participated in this study. Current support for the AGP comes mainly from grants from Autism Speaks (USA), Genome Canada (Canada), the Health Research Board (HRB; Ireland), the Hilibrand Foundation (USA) and the Medical Research Council (MRC; UK). Funding: This work was funded by EC 6th FP AUTISM MOLGEN (grant number LSHM-CT-2005–512158) and was also supported by the Wellcome Trust (grant number 075491/Z/04).

1 Baird G, Simonoff E, Pickles A *et al*: Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *Lancet* 2006; **368**: 210–215.
2 Autism Genome Project Consortium, Szatmari P, Paterson AD, Zwaigenbaum L *et al*: Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat Genet* 2007; **39**: 319–328.
3 Santangelo SL, Tsatsanis K: What is known about autism: genes, brain, and behavior. *Am J Pharmacogenomics* 2005; **5**: 71–92.
4 Fombonne E: Epidemiology of autistic disorder and other pervasive developmental disorders. *J Clin Psychiatry* 2005; **66** (Suppl): 3–8.
5 Muhle R, Trentacoste SV, Rapin I: The genetics of autism. *Pediatrics* 2004; **113**: e472–e486.
6 Folstein S, Rutter M: Infantile autism: a genetic study of 21 twin pairs. *J Child Psychol Psychiatry* 1977; **18**: 297–321.
7 Bailey A, Le Couteur A, Gottesman I *et al*: Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol Med* 1995; **25**: 63–77.
8 Freitag CM: The genetics of autistic disorders and its clinical relevance: a review of the literature. *Mol Psychiatry* 2007; **12**: 2–22.
9 Persico AM, D'Agruma L, Maiorano N *et al*: Reelin gene alleles and haplotypes as a factor predisposing to autistic disorder. *Mol Psychiatry* 2001; **6**: 150–159.
10 Bonora E, Beyer KS, Lamb JA *et al*: Analysis of reelin as a candidate gene for autism. *Mol Psychiatry* 2003; **8**: 885–892.
11 Wang K, Zhang H, Ma D *et al*: Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature* 2009; **459**: 528–533.
12 Auranen M, Vanhala R, Varilo T *et al*: A genome-wide screen for autism-spectrum disorders: evidence for a major susceptibility locus on chromosome 3q25-27. *Am J Hum Genet* 2002; **71**: 777–790.
13 International Molecular Genetic Study of Autism Consortium (IMGSAC): A full genome screen for autism with evidence for linkage to a region on chromosome 7q. *Hum Mol Genet* 1998; **7**: 571–578.
14 International Molecular Genetic Study of Autism Consortium (IMGSAC): A genome-wide screen for autism: strong evidence for linkage to chromosomes 2q, 7q and 16p. *Am J Hum Genet* 2001b; **69**: 570–581.
15 Philippe A, Martinez M, Guilloud-Bataille M *et al*: Genome-wide scan for autism susceptibility genes. Paris Autism Research International Sibpair Study. *Hum Mol Genet* 1999; **8**: 805–812.
16 Maestrini E, Pagnamenta AT, Lamb JA *et al*: High-density SNP association study and copy number variation analysis of the AUTS1 and AUTS5 loci implicate the IMMP2L-DOCK4 gene region in autism susceptibility. *Mol Psychiatry* 2009; e-pub ahead of print.
17 Shuang M, Liu J, Jia MX *et al*: Family-based association study between autism and glutamate receptor 6 gene in Chinese Han trios. *Am J Med Genet B Neuropsy Genet* 2004; **131B**: 48–50.
18 Jamain S, Betancur C, Quach H *et al*: Linkage and association of the glutamate receptor 6 gene with autism. *Mol Psychiatry* 2002; **7**: 302–310.
19 Philippi A, Roschmann E, Tores F *et al*: Haplotypes in the gene encoding protein kinase c-beta (PRKCB1) on chromosome 16 are associated with autism. *Mol Psychiatry* 2005; **10**: 950–960.
20 Lintas C, Sacco R, Garbett K *et al*: Involvement of the *PRKCB1* gene in autistic disorder: significant genetic association and reduced neocortical gene expression. *Mol Psychiatry* 2008; **14**: 705–718.
21 Cook Jr EH, Courchesne R, Lord C *et al*: Evidence of linkage between the serotonin transporter and autistic disorder. *Mol Psychiatry* 1997; **2**: 247–250.
22 Mulder EJ, Anderson GM, Kema IP *et al*: Serotonin transporter intron 2 polymorphism associated with rigid-compulsive behaviors in Dutch individuals with pervasive developmental disorder. *Am J Med Genet B Neuropsychiatr Genet* 2005; **133B**: 93–96.
23 Durand CM, Betancur C, Boeckers TM *et al*: Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. *Nat Genet* 2007; **39**: 25–27.
24 Melke J, Goubran Botros H, Chaste P *et al*: Abnormal melatonin synthesis in autism spectrum disorders. *Mol Psychiatry* 2008; **13**: 90–98.
25 Depienne C, Moreno-De-Luca D, Heron D *et al*: Screening for genomic rearrangements and methylation abnormalities of the 15q11-q13 region in the autism spectrum disorders. *Biol Psychiatry* 2009; **66**: 349–359.
26 Lord C, Rutter M, Le Couteur A: Autism diagnostic interview – revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord* 1994; **24**: 659–685.

27 Lord C, Risi S, Lambrecht L; *et al*: The Autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of Autism. *J Autism Dev Disord* 2000; **30**: 205–223.
28 Barrett JC, Fry B, Maller J, Daly MJ: Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21**: 263–265.
29 Weiss LA, Arking DE, The Gene Discovery Project of John Hopkins and the Autism Consortium: A genome-wide linkage and association scan reveals novel loci for autism. *Nature* 2009; **461**: 802–808.
30 O'Connell JR, Weeks DE: PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 1998; **63**: 259–266.
31 Lander E, Kruglyak L: Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995; **11**: 241–247.
32 Kazeem GR, Farrall M: Integrating case-control and TDT studies. *Ann Hum Genet* 2005; **69**: 329–335.
33 Normand SL: Meta-analysis: formulating, evaluating, combining, and reporting. *Stat Med* 1999; **18**: 321–359.
34 Purcell S, Cherny SS, Sham PC: Genetic power calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics Applications Note* 2003; **19**: 149–150.
35 Lamb JA, Barnby G, Bonora E *et al*: Analysis of IMGSAC autism susceptibility loci: evidence for sex limited and parent of origin specific effects. *J Med Genet* 2005; **42**: 132–137.
36 O'Donovan MC, Craddock N, Norton N *et al*: Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat Genet* 2008; **40**: 1053–1055.
37 Gallagher L, Becker K, Kearney G *et al*: Brief Report: a case of autism associated with del(2)(q32.1q32.2) or (q32.2q32.3). *J Aut Dev Dis* 2003; **33**: 105–108.
38 Conroy J, Cochrane L, Anney RJL *et al*: Fine mapping and association studies in a candidate region for autism on chromosome 2q31-q32. *Am J Med Genet B Neuropsychiatr Genet* 2008; **150B**: 535–544.
39 Veenstra-VanderWeele J, Christian SL, Cook Jr EJ: Autism as a paradigmatic complex genetic disorder. *Annu Rev Genomics Hum Genet* 2004; **5**: 379–405.
40 Barnby G, Abbott A, Sykes N *et al*: Candidate-gene screening and association analysis at the autism-susceptibility locus on chromosome 16p: evidence of association at *GRIN2A* and *ABAT*. *Am J Hum Genet* 2005; **76**: 950–966.
41 Tsuchiya N, Ochiai M, Nakashima K, Ubagai T, Sugimura T, Nakagama H: SND1, a Component of RNA-induced silencing complex, is up-regulated in human colon cancers and implicated in early stage colon carcinogenesis. *Cancer Res* 2007; **67**: 9568–9576.
42 Marshall CR, Noor A, Vincent JB *et al*: Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet* 2008; **82**: 477–488.
43 Morrow EM, Yoo SY, Flavell SW *et al*: Identifying autism loci and genes by tracing recent shared ancestry. *Science* 2008; **321**: 218–223.
44 Bray NJ, Kirov G, Owen RJ *et al*: Screening the human protocadherin 8 (*PCDH8*) gene in schizophrenia. *Genes Brain Behav* 2002; **1**: 187–191.
45 Carlsson ML: Hypothesis: is infantile autism a hypoglutamatergic disorder? Relevance of glutamate – serotonin interactions for pharmacotherapy. *J Neural Transm* 1998; **105**: 525–535.
46 Kim SA, Kim JH, Park M, Cho IH, Yoo HJ: Family-based association study between *GRIK2* polymorphisms and autism spectrum disorders in the Korean trios. *Neuro Res* 2007; **58**: 332–335.
47 Dutta S, Das S, Guhathakurta S *et al*: Glutamate receptor 6 gene (*GluR6* or *GRIK2*) polymorphisms in the Indian population: a genetic association study on autism spectrum disorder. *Cell Mol Neurobiol* 2007; **27**: 1035–1047.

APPENDIX

Members of the EU Autism MOLGEN Consortium in alphabetical order are as follows:

Reija Alen[1], Elena Bacchelli[2], Anthony Bailey[3], Gillian Baird[4], Agatino Battaglia[5], Catalina Betancur[6], Annelies de Bildt[7], Francesca Blasi[2], Sven Bölte[8], Patrick Bolton[9], Thomas Bourgeron[10], Karen Brøndum-Nielsen[11], Simona Carone[2], Pauline Chaste[10], Andreas Chiocchetti[12], Eftichia Duketis[13], Christelle Durand[10], Herman van Engeland[14], Penny Farrar[15], Sabine Feineis-Matthews[13], Bärbel Felder[12], Kostas Francis[3], Jeanne Fremolle[16], Carina Gillberg[17], Christopher Gillberg[17], Hany Goubran-Botros[10], Demetrious Haracopos[11], Evelyn Herbrecht[13], Richard Holt[15], Gemma Honeyman[3], Jasmin Honold[13], Renske Houben[14], Aislinn Hutchison[3], Roberta Igliozzi[5], Torben Isager[11], Irma Järvelä[1], Maria Johansson[17], Maretha de Jonge[14], Sabine M. Klauck[12], Anne Koivisto[1], Hanna Komu[1], Marion Leboyer[18], Ann Le Couteur[19], Justin Lowen[3], Elena Maestrini[2], Carine Mantoulan[16], Jonas Melke[10], Helen McConachie[19], Ruud Minderaa[7], Anthony Monaco[15], Erik Mulder[7], Taina Nieminen-von Wendt[1], Ilona Nummela[1], Gudrun Nygren[17], Geeta Pakalapati[12], Katerina Papanikolaou[20], Barbara Parrini[5], Lennart Pederson[11], Liz Pellicano[3], Catherine Pienkowski[16], Judith Ponsford[3], Annemarie Poustka[12], Fritz Poustka[13], Maria Rastam[17], Karola Rehnström[1], Katy Renshaw[3], Bernadette Rogé[16], Dorothea Ruehl[13], Michael Rutter[4], Susan Sarenius[1], Gabriele Schmötzer[13], Claudia Schuster[12], Henrik Anckarsater[17], Raffaella Tancredi[5], Maïté Tauber[16], John Tsiantis[20], Nora Uhlig[13], Raija Vanhala[1], Simon Wallace[3], Lennart von Wendt[1], Kerstin Wittemeyer[3], Tero Ylisaukko-oja[1].

[1]Helsinki University Central Hospital, Helsinki, Finland.
[2]University of Bologna, Alma Mater Studorium, Bologna, Italy.
[3]Department of Psychiatry, University of Oxford, Oxford, UK.
[4]Newcomen Centre, Kings College London, London, UK.
[5]Stella Maris Clinical Research Institute for Child and Adolescent Neuropsychiatry, Pisa, Italy.
[6]INSERM U952, CNRS UMR7224, UPMC Univ Paris 06, Paris, France.
[7]Stichting Universitaire en Algemene Kinder – en Jeugdpsychiatrie Noord – Nederland (Accare)/University Medical Center Groningen, University Center Child and Adolescent Psychiatry Groningen, Groningen, The Netherlands.
[8]Department of Child and Adolescent Psychiatry, Central Institute of Mental Health, Mannheim, Germany.
[9]Department of Child and Adolescent Psychiatry, Kings College London, London, UK.
[10]Institut Pasteur, Paris, France.
[11]Center for Autisme, Herlev, Denmark.
[12]Division of Molecular Genome Analysis, German Cancer Research Center (DKFZ), Heidelberg, Germany.
[13]Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, Johann Wolfgang Goethe-University, Deutschordenstr. 50, 60528, Frankfurt/Main, Germany.
[14]Department of Child and Adolescent Psychiatry, University Medical Centre Utrecht, Utrecht, The Netherlands.
[15]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK.
[16]Universite de Toulouse le Mirail., Toulouse, France.
[17]Department of Child and Adolescent Psychiatry, Goteborg University, Goteborg, Sweden.
[18]INSERM U955; AP-HP, Henri Mondor-Albert Chenevier Hospital, Department of Psychiatry; Université Paris 12, Faculty of Medicine, Créteil, France.
[19]Department of Child and Adolescent Psychiatry, Sir James Spence Institute, University of Newcastle, Royal Victoria Infirmary, Newcastle, UK.
[20]Department of Child Psychiatry, Medical School, National and Kapodistrian University of Athens, Athens, Greece.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)