EDITOR'S CHOICE

Rare familial 16q21 microdeletions under a linkage peak implicate cadherin 8 (*CDH8*) in susceptibility to autism and learning disability

Alistair T Pagnamenta,[1] Hameed Khan,[2] Susan Walker,[2] Dianne Gerrelli,[3] Kirsty Wing,[1] Maria Clara Bonaglia,[4] Roberto Giorda,[5] Tom Berney,[6] Elisa Mani,[7] Massimo Molteni,[7] Dalila Pinto,[2] Ann Le Couteur,[6] Joachim Hallmayer,[8] James S Sutcliffe,[9] Peter Szatmari,[10] Andrew D Paterson,[2] Stephen W Scherer,[2] Veronica J Vieland,[11] Anthony P Monaco[1]

► Additional figure and tables are published online only. To view these files please visit the journal online (http://jmg.bmj.com).

For numbered affiliations see end of article.

Correspondence to
Professor Anthony P Monaco, Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Headington, Oxford OX3 7BN, UK; anthony.monaco@well.ox.ac.uk

Received 19 April 2010
Revised 11 August 2010
Accepted 16 August 2010
Published Online First 23 October 2010

ABSTRACT

Background Autism spectrum disorder (ASD) is characterised by impairments in social communication and by a pattern of repetitive behaviours, with learning disability (LD) typically seen in up to 70% of cases. A recent study using the PPL statistical framework identified a novel region of genetic linkage on chromosome 16q21 that is limited to ASD families with LD.

Methods In this study, two families with autism and/or LD are described which harbour rare >1.6 Mb microdeletions located within this linkage region. The deletion breakpoints are mapped at base-pair resolution and segregation analysis is performed using a combination of 1M single nucleotide polymorphism (SNP) technology, array comparative genomic hybridisation (CGH), long-range PCR, and Sanger sequencing. The frequency of similar genomic variants in control subjects is determined through analysis of published SNP array data. Expression of *CDH8*, the only gene disrupted by these microdeletions, is assessed using reverse transcriptase PCR and in situ hybridisation analysis of 9 week human embryos.

Results The deletion of chr16: 60 025 584—61 667 839 was transmitted to three of three boys with autism and LD and none of four unaffected siblings, from their unaffected mother. In a second family, an overlapping deletion of chr16: 58 724 527—60 547 472 was transmitted to an individual with severe LD from his father with moderate LD. No copy number variations (CNVs) disrupting *CDH8* were observed in 5023 controls. Expression analysis indicates that the two *CDH8* isoforms are present in the developing human cortex.

Conclusion Rare familial 16q21 microdeletions and expression analysis implicate *CDH8* in susceptibility to autism and LD.

INTRODUCTION

Autism spectrum disorder (ASD) is a clinically heterogeneous condition characterised by impairments in social communication and by a pattern of repetitive behaviours. There is strong evidence of genetic heritability. Learning disability (LD) is typically observed in up to 70% of cases, depending on ascertainment,[1] while epilepsy is reported in over 20%.[2 3]

An accumulating body of evidence suggests that ASD may result from aberrant synaptic connections.[4] For example, rare variants involving neuroligin and neurexin genes, which encode proteins that interact across the synaptic cleft, have been implicated in autism susceptibility.[5 6] In addition, disruption of *SHANK3*, a gene that encodes a postsynaptic scaffolding protein that interacts with the neuroligins, has also been found at low frequency in some,[7 8] but not all, ASD cohorts.[9]

Linkage analysis has traditionally been used to search for genetic loci involved in autism susceptibility. The large number of loci described to date[10] in part reflects the complex genetic architecture underlying the condition. However, it is likely that subtle differences exist between clinical cohorts in terms of ascertainment strategies, inclusion/exclusion criteria, as well as the population backgrounds from which subjects are taken. In the presence of locus heterogeneity, these factors, together with simple sampling variability, can lead to very different mixtures of genetic subtypes across studies. Such differences may also act to confound replication.

To address these issues, a reanalysis of the Autism Genome Project (AGP) consortium's linkage data has recently been undertaken, using the PPL analytical framework.[11–13] This study identified a novel susceptibility locus on 16q21 coming from the low IQ ASD subgroup. We were interested to note that this linkage peak overlaps a rare deletion found in an individual with autism and learning disability, detected as part of our recent genome-wide copy number variation (CNV) scans.[6 14] Although 1.64 Mb in size, this microdeletion involves a single gene, cadherin 8 (*CDH8*). In this study, we describe the genetic characterisation of this rare microdeletion, further clinical evaluation and segregation analysis in this large nuclear family. The inheritance pattern seen, together with an absence of similar microdeletions in over 5000 control subjects and the hypothesised involvement of other cadherin genes in ASD and related neurodevelopmental disorders,[15–17] led us to conclude that this rare microdeletion may be acting as an autism predisposition factor in this family. To extend these findings, we also assessed the expression pattern of this gene in the developing human

brain, describe a second family with an overlapping microdeletion disrupting *CDH8*, and sequence additional individuals with ASD.

METHODS

Clinical details for family 3099

The Autism Diagnostic Interview-Revised (ADI-R)[18] and the Autism Diagnostic Observation Schedule (ADOS)[19] were administered when subjects were between 12–15 years of age. All three affected children met criteria for autism on the ADI and the ADOS. All three had word and phrase delay and in one affected child (3099_006) there is a history of regression. This child lost many skills, in particular language, over the course of 6 months, and took a year to regain them. Vineland Adaptive Behaviour Scales[20] scores were below 50 on all domains of socialisation, communication, and daily living skills for all three children with autism. Cognitive testing with the Ravens[21] was performed on two of the three boys and provided IQ scores in the 70–75 range. The eldest affected son (3099_006) was assessed as having intellectual disability by clinical judgement and was in treatment for language delay and learning disabilities. There was no evidence of epilepsy or associated medical problems at the time of assessment, but all three had a head circumference at the 90th to 97th centile, in spite of heights in the normal range. The mother and one unaffected sibling (3099_009) also have large head circumferences (>95th centile). The mother had three previous miscarriages (around 12 weeks gestation) and was phenotypically normal. There were no reports of neurodevelopmental disorder in her extended family. The father, who is separated from the family, was reported to have a normal developmental history and personality, but to have developed a psychiatric disorder as an adult. His first cousin is reported to have had autism. There was no other history of developmental disability, mental illness or epilepsy in the immediate family; however, the mother (3099_002) and youngest son (3099_009) have osteoarthritis.

Clinical details for family 09

The proband (09_003) was evaluated at 20 years 11 months of age as part of a study on learning disability. On the ADI-R, he had a score of 6 on social interaction, 6 on communication, and 0 on repetitive behaviours. On module 4 of the ADOS, he scored 3, 1, and 2 on the social, communication, and play sections, respectively. Therefore on neither measure does he qualify for a diagnosis of autism or ASD. However, his IQ was below 45 on the Wechsler Adult Intelligence Scale-Revised (WAIS-R) and scores on the Vineland Adaptive Behaviour Scales showed significant impairment on all scales. He was reported to have had language delay as a child. There was no history of epilepsy or other comorbid medical problems. However, peripheral palsy of the right 7 facial nerve and obesity (body mass index (BMI) of 31.9 kg/m^2) were reported at age 11 and 14 years, respectively. The father scored in the borderline range on the IQ tests (66 for verbal IQ and 77 for performance IQ). The younger brother demonstrated typical IQ scores, but was reported to have been in treatment for language delay and learning disabilities between the ages of 7 and 11 years.

CNV characterisation and segregation analysis

In order to validate the microdeletion in family 3099 and carry out segregation analysis, long-range PCR was performed using the BIO-X-ACT long DNA polymerase kit (BIOLINE, London, UK) using the manufacturer's suggested protocol. Primers GCTATCCAGTAGGAAGTGAAACA and AATGAGTATAAGAATCAAAGATGTGA were designed following visual inspection of 1M-single single nucleotide polymorphism (SNP) array data from a recent genome-wide CNV scan,[14] within BeadStudio (Illumina, San Diego, California, USA). The 3023 bp deletion-spanning amplicon was purified using exonuclease I (NEB, Ipswich, Massachusetts, USA) and SAP (USB, Cleveland, Ohio, USA) and then sequenced using BigDye v3.1 (Applied Biosystems, Foster City, California, USA). We note that this CNV was initially detected based on Affymetrix 10K SNP data and reported as a de novo event (see supplemental table 3 in Szatmari *et al*[6]). Higher resolution data from the current Illumina 1M scan indicate that it is in fact inherited from the mother.

For family 09, a combination of high resolution array based comparative genomic hybridisation (aCGH), using NA10851 (male) and NA15510 (female) as control DNAs, and quantitative PCR experiments, was first performed to help resolve the microdeletion breakpoints (data not shown). Long range PCR across the deletion was then carried out using primers CCACATCCTTTCACACATGAGAA and TAGCTGCTTTCCCACATATCAT. The 4610 bp deletion spanning amplicon was then purified and sequenced as described above.

Sanger sequencing CDH8 for 3099_006, 3099_007 and 3099_008

PCR was used to amplify all coding exons (1–11) and the 5′-UTR of *CDH8*. PCR products were purified using the ChargeSwitch PCR Clean-Up Kit (Invitrogen, Eugene, Oregon, USA) and then sequenced using the BigDye Terminator kit (v3.1), according to manufacturers' recommendations. To search for novel variants, we compared results against dbSNP version 131.[22] Primer sequences are available in supplementary table 1. Further information about thermocycling conditions is available on request.

CNV controls

We used data from 5023 control subjects of European ancestry: 2416 from the PopGen study,[14] controls from the Ontario Ottawa Heart Control study[23] or HapMap controls genotyped on the Affymetrix 6.0 array; 1287 from the SAGE control project genotyped on the Illumina 1M platform[14 24]; and 1320 from the CHOP paediatric control study genotyped on the Illumina 550 k array.[25]

RT-PCR

Reverse transcriptase PCR (RT-PCR) was performed using the OneStep-RT PCR kit (Qiagen, Crawley, UK) according to the manufacturer's protocol. Exact-match primers were designed to the first coding exon and the 3′-UTR of each of the isoforms. For the shorter isoform, primer sequences corresponding to the coding exon and 3′-UTR were as follows: CDH8.RTPCR.F1 ACCGCTCCAAAAGAGGCTGG, and CDH8.R9 GCACAGCAGGTTGTTCCAC. For the longer isoform, the same forward primer was used together with the reverse primer CDH8.RTPCR.R2 TGACTGGTGCTAAACTTGCCTC. Exact match primers GAPDH.F1 GAAGGTGAAGGTCGGAGTCA and GAPDH.R1 TGGAAGATGGTGATGGGATT were designed in the first and third coding exons of *GAPDH*, to serve as a positive control. Total RNA from fetal brain and various regions of the adult human brain was purchased (Stratagene, La Jolla, California, USA) and 100 ng of RNA template was added to each reaction. Thermal cycling conditions for products were as follows: 50°C for 30 min, 95°C for 15 min, followed by 40 cycles of 94°C for 30 s, 52°C for 30 s, and 72°C for 2 min, with a final extension at 72°C for 10 min.

In situ hybridisation

In situ hybridisation was carried out on human fetal tissue (9 weeks gestation) as described by Wilkinson.[26] Probe regions comprising the 3′-UTR region of the short *CDH8* isoform were

amplified from whole brain cDNA using primers GAAAACCCGGCCAAGTAAAT and CAGATTTCAATATTCACTTCCTACAA. Probe regions comprising the 3′-UTR region of the long *CDH8* isoform were amplified using primers TCTACTCTGTTGGTGAAAGTGACA and TGTCTGTGGTGGTCAGGTAAA. Further details are provided in the supplementary information.

RESULTS

Microdeletions involving CDH8 within a low IQ ASD linkage region

A recent study reanalysing the AGP's linkage data, using the PPL statistical framework, has identified a new susceptibility locus on chromosome 16q13-21 in the low IQ ASD subset. This reaches a maximum PPL of 95.95% at rs1476307.[13] This means there is a 95.95% chance that this region contains an autism susceptibility gene, based on the available data. The ~6 Mb region under the linkage peak, particularly in chromosome band 16q21 (distal to the peak of linkage), is relatively gene-poor. We have discovered two large inherited microdeletions within this linkage peak and overlapping the *CDH8* gene, in two independent families (figure 1).

Family 3099 is of European ancestry and comes from the International Molecular Genetic Study of Autism Consortium (IMGSAC) study cohort.[27] Subject 3099_008 was included in a recent CNV scan of 996 individuals with ASD.[14] This individual was found to carry an inherited heterozygous 1.64 Mb microdeletion involving the whole *CDH8* gene, but no other genes. This microdeletion was transmitted from the unaffected mother to the proband and his two brothers, all of whom presented with both LD and ASD. The microdeletion was not transmitted to the four unaffected siblings (figure 2A). Analysis of a combination of chromosome 16 SNPs and microsatellites indicates that the non-deleted paternal copy of *CDH8* was also shared identical-by-descent in all three affected children; however, Sanger sequencing uncovered no novel exonic variants in *CDH8* on this chromosome. We validated the microdeletion using long range PCR and then by Sanger sequencing the breakpoint junction fragment (figure 2B). The absence of any sequence similarities flanking the breakpoints suggests that this chr16:60 025 584–61 667 839 microdeletion is likely to be a rare, potentially 'private' mutation. A combination of SNPs and microsatellite data from this family, generated in previous studies,[6 27] determined that the linkage signal from this family alone reached the maximum possible for a pedigree of this size (maximum logarithm of odds (LOD)=1.7, maximising model=recessive).

Family 09 was discovered during an independent aCGH genome-wide screen of a heterogeneous cohort of 80 Italian children with generalised LD, based on Agilent 44 k technology. All subjects were negative for karyotype and telomere-FISH (fluorescent in situ hybridisation) abnormalities, as well as for fragile-X. The family comprises parents and two sons, with the deletion being transmitted from the father to both sons (figure 2A). Neither child received a diagnosis of ASD; however, the proband scored in the very low range for IQ with substantial impairment, while the brother showed normal IQ but was reported to have had a history of language delay and treatment for learning disability. The father also scored in the borderline range for IQ. The distal breakpoint of this 1.82 Mb deletion is situated within the largest (120 kb) *CDH8* intron, and all but one of the coding exons (2–11) are removed. No other genes are disrupted due to the large gene desert proximal to *CDH8* (figure 1). Analysis of chromosome 16 microsatellite markers showed that the proband and his sibling did not share their non-deleted maternal *CDH8* haplotype (data not shown). The sequence surrounding the observed chr16:58 724 527–60 547 472 deletion revealed a 7 bp tandem duplication nearby which may have occurred at the same time as the larger deletion (figure 2C).

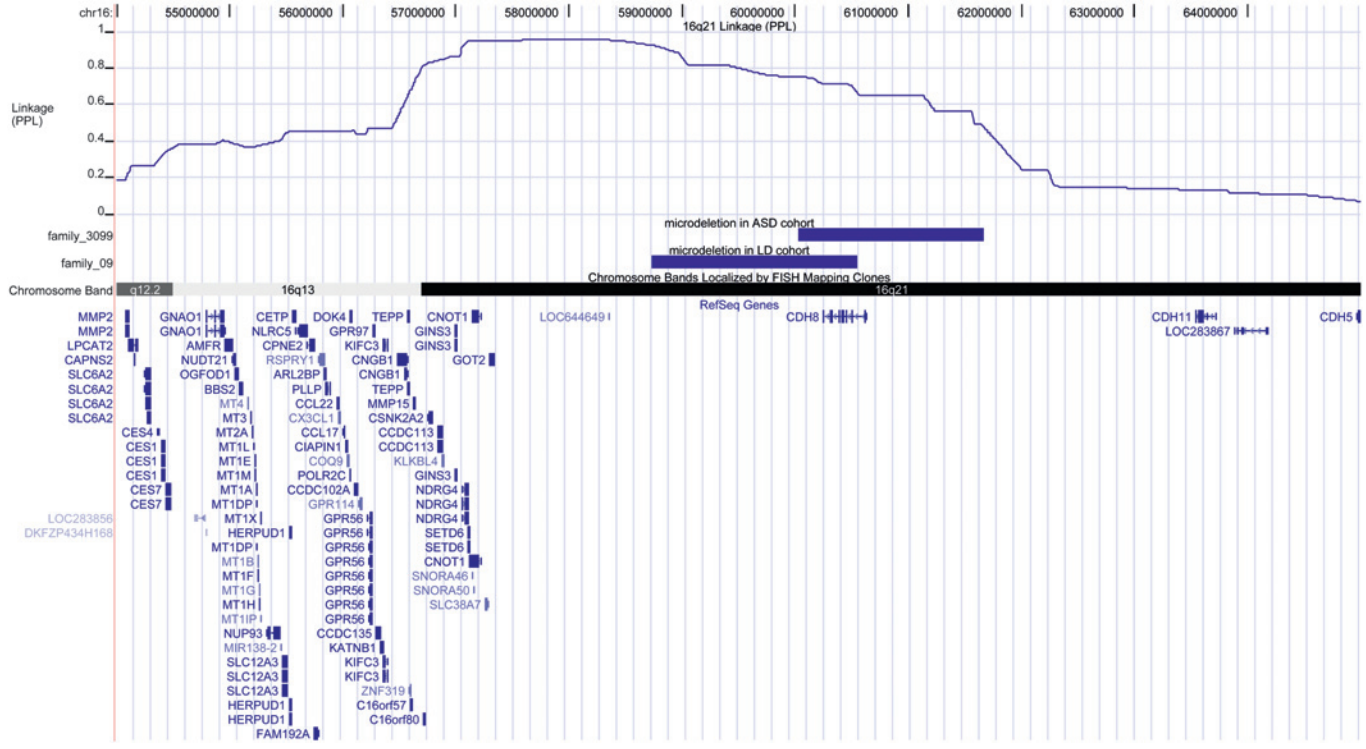

Figure 1 Schematic from the UCSC genome browser. Figure shows the position of the two inherited deletions overlapping *CDH8*, in relation to the low IQ autism spectrum disorder (ASD) linkage peak from our recent analysis.[13] The y axis indicates the PPL score. RefSeq gene coordinates are also plotted beneath the chromosome band track. The region shown corresponds to 54–65 Mb on 16q12.2-21 (NCBI build 36 coordinates).

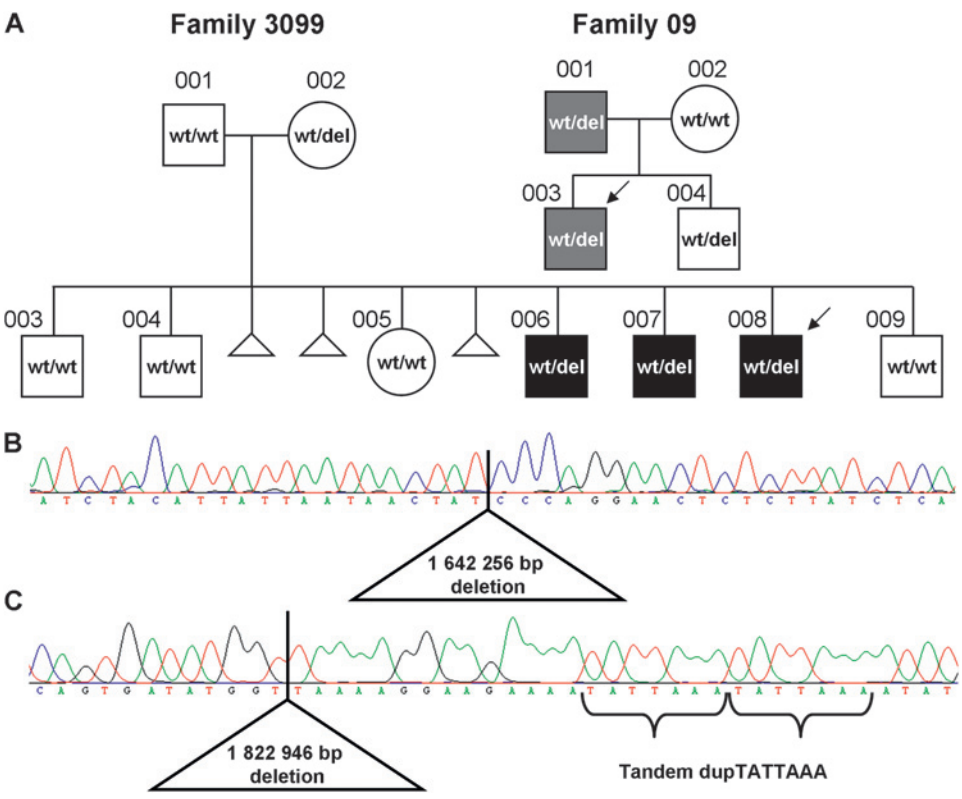

Figure 2 Deletions found in families 3099 and 09. (A) Pedigrees show the segregation pattern for these two deletions involving *CDH8*. Autism and learning disability indicated in black shading, learning disability alone indicated in grey; del, 16q21 deletion; wt, wild-type. Although individual 09_004 demonstrated typical IQ scores, he was reported to have been in treatment for language delay and learning disabilities between the ages of 7 and 11 years. (B) Electropherogram showing DNA sequence spanning the chr16:60 025 584—61 667 839 (NCBI build 36) deletion in family 3099. (C) Electropherogram showing DNA sequence spanning the chr16:58 724 527—60 547 472 deletion in family 09. The position of a 7 bp tandem duplication is indicated.

CNV analysis of controls

There were no similar CNVs disrupting the *CDH8* gene detected in 5023 control subjects from published high resolution (550 k and above) SNP array data. We note that the Database of Genomic Variants (http://projects.tcag.ca/variation/) reports a duplication of 23.7 kb in a single population control sample (NA18852), involving a single coding exon of *CDH8*.[28] This CNV was detected using PennCNV analysis and involves just four SNPs from the Illumina 550 k SNP array (supplementary figure 1). However, numerous other studies (some using higher resolution platforms[29]) have also examined this same DNA sample and do not report the CNV, suggesting that it may be a false-positive. We therefore tested this HapMap sample using a combination of qPCR primer pairs and confirmed that there is a normal copy number at this locus (see supplementary information).

Expression analysis

Although Epstein—Barr virus (EBV) transformed peripheral blood lymphocytes were available for family 3099, we were unable to amplify *CDH8* transcripts using RT-PCR (data not shown). This may be because this gene is not expressed in lymphocytes. Therefore, we could not determine how strongly the remaining copy of *CDH8* was expressed. However, using RT-PCR on a commercially available RNA panel, expression of two known isoforms of *CDH8* was confirmed in various parts of human the brain, particularly in the cortex (figures 3A,B). Both isoforms were also detected in fetal brain, although the long isoform was only just detectable (figures 3B).

A more complete, quantitative characterisation of *CDH8* expression during early brain development was also carried out using in situ hybridisation on sagittal brain sections from a 9-week-old human embryo. These data showed expression of the shorter *CDH8* isoform towards the front of the cerebral cortex (figure 3C), a similar pattern to other ASD candidate genes such as *CNTNAP2*[30] and *CDH10*.[17] The longer *CDH8* isoform demonstrated a more posterior cortical expression at this early developmental stage (figure 3D).

DISCUSSION

To better account for different ASD subtypes and potential clinical-site-specific confounders, reanalysis of existing ASD linkage data has been carried out using the PPL statistical framework. In this study we describe two 16q21 microdeletions that are present within the novel linkage peak that was identified in the low IQ ASD subset.

The rarity of CNVs at this locus means it would be difficult to gain evidence of aetiological relevance using a case—control experimental design. For example, in the AGP CNV study cohorts,[14] deletions of this region were seen in 1/996 cases (that is the proband from family 3099), and 0/1287 controls. One solution is to carry out global analysis on all rare CNVs present in a study cohort, to gather statistical support for enrichments of biological pathways. For example, analysis of rare CNVs in the AGP cohort implicated genes involved with cellular proliferation, projection, motility, and GTPase/Ras signalling.[14] Pathway analyses can also integrate other datasets such as information on mouse knockout phenotypes, as has been accomplished recently for CNVs detected in learning disability.[31] Nevertheless, in larger pedigrees linkage analysis remains an additional way of supporting disease involvement. The size of the nuclear family 3099 and segregation pattern seen for the 16q21 deletion strongly suggests that this mutation plays an aetiological role. Given the evidence for overlapping aetiology between ASD and general intellectual disabilities and the detection of this linkage region specifically in the low IQ subset of the ASD sample, family 09 represents additional corroboration of *CDH8* as a susceptibility gene for ASD and/or learning disability.

Although both deletions are large (>1.6 Mb), both disrupt a single gene in this relatively gene-sparse region of 16q21 (figure 1). This gene (*CDH8*) spans 383 kb of genomic sequence

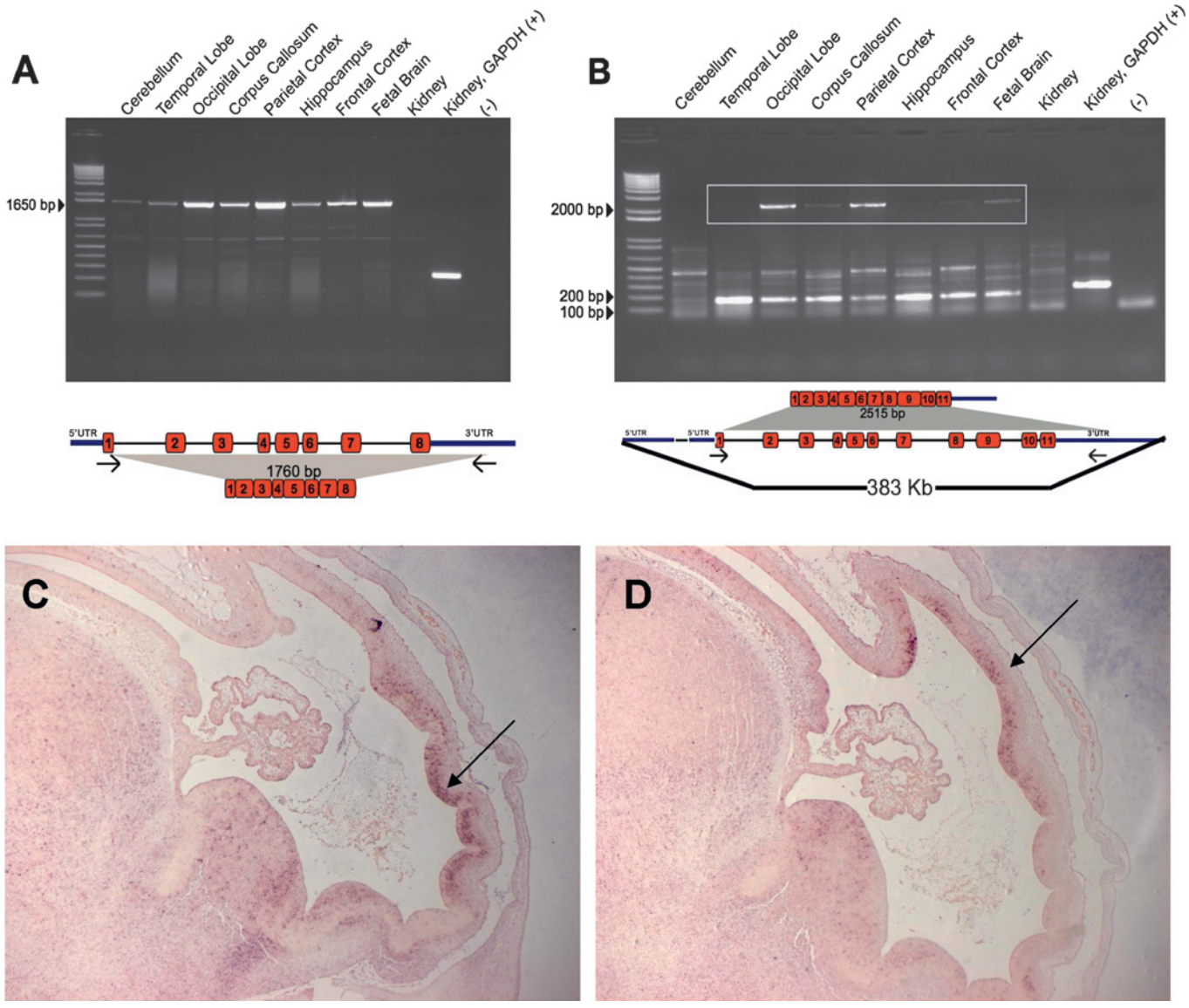

Figure 3 Expression analysis of *CDH8*. (A) Reverse transcriptase PCR (RT-PCR) analysis detected a 1760 bp amplicon corresponding to the short *CDH8* isoform. (B) RT-PCR analysis detected a 2514 bp product (boxed), corresponding to the longer *CDH8* isoform. (C) In situ hybridisation performed on sagittal sections through the head of a 9-week-old human embryo, for the short *CDH8* isoform. (D) In situ hybridization for the longer isoform. Arrows indicate cortical expression.

and encodes a classical type II cadherin. Cadherins are calcium dependant cell adhesion molecules, many of which are expressed in the brain. A recent genome-wide association study implicated common variants between *CDH9* and *CDH10* on chromosome 5, in autism susceptibility.[17] In this same study, pathway analysis implicated the whole family of cadherins (*CDH1-25*) and this enrichment was enhanced when the cadherins were grouped with the three neurexins and the five closely related *CNTNAP* genes.[17] A second study assessing both genome-wide linkage and association did not implicate the *CDH9-CDH10* locus.[32] Common variants near the *CDH7* gene have reproducibly been linked to bipolar disorder.[33] De novo deletions overlapping *CDH15*, another member of this gene family, have been detected in three individuals with ASD or 'autistic features'.[34] Alterations in *CDH15* have also been linked to LD and impaired cell–cell adhesion.[15] Meanwhile, in consanguineous kindreds, rare deletions within larger blocks of homozygosity-by-descent implicate protocadherin 10 in autism susceptibility.[16] A cadherin-rich region on 13q21 has also been implicated in specific language impairment and previous autism studies.[35 36] Finally, a de novo deletion in the affected member of a discordant monozygotic twin pair suggests *CDH12* and *CDH18* may be involved in schizophrenia.[37] In mice, knockout of the orthologous Cdh8 gene (~97% amino acid sequence identity to its human counterpart) results in abnormal synaptic transmission.[38]

Neither deletion in this study appears to be fully penetrant. Across both families, a total of 5/7 individuals with *CDH8* deletions were affected with autism and/or learning disability. This penetrance rate is similar to other recently described ASD implicated CNVs, such as microdeletion of 15q13.3.[39] Although a normal IQ was observed for the brother carrying the deletion in family 09, he did have language delay, which can be considered part of the spectrum of learning disability.

It may be that other rare mutations such as CNVs, SNPs or indels elsewhere in the genome act to modulate the penetrance and expressivity of the *CDH8* deletions in the two families described. For example, a recent study of individuals with

developmental delay and the 16p12.1 microdeletion found a higher than expected rate of large secondary CNVs, suggesting a two-hit model.[40] Although there were no other rare CNVs with obvious aetiological relevance in the families described here, the resolution of our study is such that we cannot rule out smaller CNVs or other molecular features contributing to the phenotype in a similar fashion. It may be that *CDH8* is itself just a risk factor for learning disability and this only leads to autism together with certain genetic backgrounds. The macrocephaly in family 3099 does not completely co-segregate with the deletion and so might suggest an additional risk factor interacting with the *CDH8* deletion.

A recent study has identified *DIAPH3* as a new autism susceptibility gene by virtue of rare non-synonymous variants lying in trans with a deletion.[41] The sharing of non-deleted paternal *CDH8* haplotypes in the three affected siblings in family 3099 made us consider the possibility that this 1.64 Mb deletion was also unmasking rare variants in the remaining copy of *CDH8*. Although sequence analysis did not detect any novel exonic *CDH8* variants in family 3099, we cannot exclude the possibility of mutations in non-coding regions disrupting gene regulation. Variation in the non-deleted copy of *CDH8* could potentially also explain the non-concordant phenotypes seen for the two boys with 16q21 deletions in family 09.

In situ analysis shows that the two *CDH8* isoforms have a slightly different expression pattern, suggesting that they may potentially play distinct roles in early cortical development. The more anterior expression seen for the shorter *CDH8* isoform somewhat resembles the pattern seen for *CNTNAP2* and *CDH10*, other ASD susceptibility genes for which published in situ expression data are available at 20 weeks of gestation.[17 30] However, although comparison between 9 and 20 week brain sections is difficult, we have shown that *CDH8* is expressed within the germinal zone of cortex rather than throughout the entire cortex as seen for the other two ASD candidate genes. Recent studies on *CNTNAP2* show that common ASD associated variants in this gene influence brain morphology.[42] Unfortunately, we were unable to obtain brain scans for affected individuals from our two families to assess whether *CDH8* deletions had led to abnormal cortical folding.

Although we did not detect any rare exonic *CDH8* changes in 26 individuals with ASD (data not shown), taken from families who were contributing most to the original linkage signal,[13] future studies should assess this locus for CNVs and rare sequence-level variants in larger ASD cohorts and measure the functional effects of these changes. Until additional, nonsense point mutations or de novo disruptions to the *CDH8* gene are detected in further autism cohorts, we cannot exclude the possibility that non-genic sequence motifs within this region might be acting to regulate other neighbouring genes or distant loci in trans. Nevertheless, the linkage seen at 16q21 in the low IQ ASD subgroup and the segregation pattern seen for the *CDH8* deletion in family 3099 leads us to hypothesise that disruption to this gene may influence susceptibility to autism and/or learning disability. Disruption of *CDH8* in two other individuals with learning disability, the absence of similar CNVs in controls, the expression of this gene in critical regions of the developing cortex, and the role of other cadherin genes in neurodevelopmental disorders are consistent with this interpretation.

Author affiliations
[1]The Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK
[2]The Centre for Applied Genomics and Program in Genetics and Genomic Biology, The Hospital for Sick Children and Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada
[3]Neural Development Unit, UCL Institute of Child Health, London, UK
[4]Citogenetica, Eugenio Medea Scientific Institute, Bosisio Parini, Italy
[5]Biologia Molecolare, Eugenio Medea Scientific Institute, Bosisio Parini, Italy
[6]The Institute of Health and Society, Newcastle University, Newcastle upon Tyne, UK
[7]Department of Child Psychopathology, Eugenio Medea Scientific Institute, Bosisio Parini, Italy
[8]Department of Psychiatry, Division of Child and Adolescent Psychiatry and Child Development, Stanford University School of Medicine, Stanford, California, USA
[9]Department of Molecular Physiology and Biophysics, Vanderbilt Kennedy Center, and Centers for Human Genetics Research and Molecular Neuroscience, Vanderbilt University, Nashville, Tennessee, USA
[10]Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, Ontario, Canada
[11]Battelle Center for Mathematical Medicine, The Research Institute at Nationwide Children's Hospital and The Ohio State University, Columbus, Ohio, USA

Acknowledgements The authors gratefully acknowledge the families participating in the study and the international Autism Genome Project Consortium for sharing data and ideas.*CDH8* in situ data can be accessed via the Human Developmental Biology Resource website at http://www.hdbr.org/.

Funding Funding for this work comes from the Simons Foundation, the Nancy Lurie Marks Family Foundation, the Wellcome Trust (075491/Z/04), the NIH (MH086117), a Telethon Grant (GGP06208A/B), Autism Speaks, Autistica, Genome Canada/Ontario Genomics Institute, Canadian Institutes for Health Research, The Centre for Applied Genomics and the McLaughlin Centre. Tissue for the in situ work was provided by the MRC/Wellcome Trust Human Developmental Biology Resource.

Competing interests None.

Patient consent Obtained.

Ethics approval This study was conducted with the approval of the Family 3099: Joint Ethics Committee (Newcastle & North Tyneside Health Authority/Universities of Newcastle upon Tyne/Northumbria). Family 09: The Ethics Committee at the "E. Medea" Scientific Institute. The HDBR has tissue bank ethics approval from the National Research Ethics Service.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Fombonne E.** The Changing Epidemiology of Autism. *J Appl Res Intellect Disabil* 2005;**18**:281—94.
2. **Rossi PG,** Parmeggiani A, Bach V, Santucci M, Visconti P. EEG features and epilepsy in patients with autism. *Brain Dev* 1995;**17**:169—74.
3. **Volkmar FR,** Nelson DS. Seizure disorders in autism. *J Am Acad Child Adolesc Psychiatry* 1990;**29**:127—9.
4. **Bourgeron T.** A synaptic trek to autism. *Curr Opin Neurobiol* 2009;**19**:231—4.
5. **Jamain S,** Quach H, Betancur C, Rastam M, Colineaux C, Gillberg IC, Soderstrom H, Giros B, Leboyer M, Gillberg C, Bourgeron T. Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. *Nat Genet* 2003;**34**:27—9.
6. **Szatmari P,** Paterson AD, Zwaigenbaum L, Roberts W, Brian J, Liu XQ, Vincent JB, Skaug JL, Thompson AP, Senman L, Feuk L, Qian C, Bryson SE, Jones MB, Marshall CR, Scherer SW, Vieland VJ, Bartlett C, Mangin LV, Goedken R, Segre A, Pericak-Vance MA, Cuccaro ML, Gilbert JR, Wright HH, Abramson RK, Betancur C, Bourgeron T, Gillberg C, Leboyer M, Buxbaum JD, Davis KL, Hollander E, Silverman JM, Hallmayer J, Lotspeich L, Sutcliffe JS, Haines JL, Folstein SE, Piven J, Wassink TH, Sheffield V, Geschwind DH, Bucan M, Brown WT, Cantor RM, Constantino JN, Gilliam TC, Herbert M, Lajonchere C, Ledbetter DH, Lese-Martin C, Miller J, Nelson S, Samango-Sprouse CA, Spence S, State M, Tanzi RE, Coon H, Dawson G, Devlin B, Estes A, Flodman P, Klei L, McMahon WM, Minshew N, Munson J, Korvatska E, Rodier PM, Schellenberg GD, Smith M, Spence MA, Stodgell C, Tepper PG, Wijsman EM, Yu CE, Roge B, Mantoulan C, Wittemeyer K, Poustka A, Felder B, Klauck SM, Schuster C, Poustka F, Bolte S, Feineis-Matthews S, Herbrecht E, Schmotzer G, Tsiantis J, Papanikolaou K, Maestrini E, Bacchelli E, Blasi F, Carone S, Toma C, Van Engeland H, de Jonge M, Kemner C, Koop F, Langemeijer M, Hijmans C, Staal WG, Baird G, Bolton PF, Rutter ML, Weisblatt E, Green J, Aldred C, Wilkinson JA, Pickles A, Le Couteur A, Berney T, McConachie H, Bailey AJ, Francis K, Honeyman G, Hutchinson A, Parr JR, Wallace S, Monaco AP, Barnby G, Kobayashi K, Lamb JA, Sousa I, Sykes N, Cook EH, Guter SJ, Leventhal BL, Salt J, Lord C, Corsello C, Hus V, Weeks DE, Volkmar F, Tauber M, Fombonne E, Shih A, Meyer KJ. Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat Genet* 2007;**39**:319—28.
7. **Moessner R,** Marshall CR, Sutcliffe JS, Skaug J, Pinto D, Vincent J, Zwaigenbaum L, Fernandez B, Roberts W, Szatmari P, Scherer SW. Contribution of SHANK3 mutations to autism spectrum disorder. *Am J Hum Genet* 2007;**81**:1289—97.
8. **Durand CM,** Betancur C, Boeckers TM, Bockmann J, Chaste P, Fauchereau F, Nygren G, Rastam M, Gillberg IC, Anckarsater H, Sponheim E, Goubran-Botros H, Delorme R, Chabane N, Mouren-Simeoni MC, de Mas P, Bieth E, Roge B, Heron D,

Burglen L, Gillberg C, Leboyer M, Bourgeron T. Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. *Nat Genet* 2007;**39**:25—7.
9. **Sykes NH,** Toma C, Wilson N, Volpi EV, Sousa I, Pagnamenta AT, Tancredi R, Battaglia A, Maestrini E, Bailey AJ, Monaco AP. Copy number variation and association analysis of SHANK3 as a candidate gene for autism in the IMGSAC collection. *Eur J Hum Genet* 2009;**17**:1347—53.
10. **Sousa I,** Holt R, Pagnamenta AT, Monaco AP. Unravelling the genetics of autism spectrum disorders. In: Rezaie P, Roth I, eds. *Researching the Autism Spectrum: Contemporary Perspectives*: Cambridge University Press, In press.
11. **Huang Y,** Segre A, O'Connell J, Wang H, Vieland VJ. KELVIN: A 2nd generation distributed multiprocessor linkage and linkage disequilibrium analysis program. ASHG Annual Meeting. October 2006. http://www.ashg.org/genetics/ashg06s/index.shtml.
12. **Vieland VJ.** Thermometers: something for statistical geneticists to think about. *Hum Hered* 2006;**61**:144—56.
13. **Vieland VJ,** AGC, AGP. New linkage analysis by the Autism Genome Project (AGP) reveals strong evidence of linkage to multiple loci as well as gene-gene interactions. ASHG Annual Meeting. November 2008. http://www.ashg.org/2008meeting/abstracts/fulltext Manuscript under review with Human Genetics.
14. **Pinto D,** Pagnamenta AT, Klei L, Anney R, Merico D, Regan R, Conroy J, Magalhaes TR, Correia C, Abrahams BS, Almeida J, Bacchelli E, Bader GD, Bailey AJ, Baird G, Battaglia A, Berney T, Bolshakova N, Bolte S, Bolton PF, Bourgeron T, Brennan S, Brian J, Bryson SE, Carson AR, Casallo G, Casey J, Chung BH, Cochrane L, Corsello C, Crawford EL, Crossett A, Cytrynbaum C, Dawson G, de Jonge M, Delorme R, Drmic I, Duketis E, Duque F, Estes A, Farrar P, Fernandez BA, Folstein SE, Fombonne E, Freitag CM, Gilbert J, Gillberg C, Glessner JT, Goldberg J, Green A, Green J, Guter SJ, Hakonarson H, Heron EA, Hill M, Holt R, Howe JL, Hughes G, Hus V, Igliozzi R, Kim C, Klauck SM, Kolevzon A, Korvatska O, Kustanovich V, Lajonchere CM, Lamb JA, Laskawiec M, Leboyer M, Le Couteur A, Leventhal BL, Lionel AC, Liu XQ, Lord C, Lotspeich L, Lund SC, Maestrini E, Mahoney W, Mantoulan C, Marshall CR, McConachie H, McDougle CJ, McGrath J, McMahon WM, Merikangas A, Migita O, Minshew NJ, Mirza GK, Munson J, Nelson SF, Noakes C, Noor A, Nygren G, Oliveira G, Papanikolaou K, Parr JR, Parrini B, Paton T, Pickles A, Pilorge M, Piven J, Ponting CP, Posey DJ, Poustka A, Poustka F, Prasad A, Ragoussis J, Renshaw K, Rickaby J, Roberts W, Roeder K, Roge B, Rutter ML, Bierut LJ, Rice JP, Salt J, Sansom K, Sato D, Segurado R, Sequeira AF, Senman L, Shah N, Sheffield VC, Soorya L, Sousa I, Stein O, Sykes N, Stoppioni V, Strawbridge C, Tancredi R, Tansey K, Thiruvahindrapduram B, Thompson AP, Thomson S, Tryfon A, Tsiantis J, Van Engeland H, Vincent JB, Volkmar F, Wallace S, Wang K, Wang Z, Wassink TH, Webber C, Weksberg R, Wing K, Wittemeyer K, Wood S, Wu J, Yaspan BL, Zurawiecki D, Zwaigenbaum L, Buxbaum JD, Cantor RM, Cook EH, Coon H, Cuccaro ML, Devlin B, Ennis S, Gallagher L, Geschwind DH, Gill M, Haines JL, Hallmayer J, Miller J, Monaco AP, Nurnberger JI Jr, Paterson AD, Pericak-Vance MA, Schellenberg GD, Szatmari P, Vicente AM, Vieland VJ, Wijsman EM, Scherer SW, Sutcliffe JS, Betancur C. Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* 2010;**466**:368—72.
15. **Bhalla K,** Luo Y, Buchan T, Beachem MA, Guzauskas GF, Ladd S, Bratcher SJ, Schroer RJ, Balsamo J, DuPont BR, Lilien J, Srivastava AK. Alterations in CDH15 and KIRREL3 in patients with mild to severe intellectual disability. *Am J Hum Genet* 2008;**83**:703—13.
16. **Morrow EM,** Yoo SY, Flavell SW, Kim TK, Lin Y, Hill RS, Mukaddes NM, Balkhy S, Gascon G, Hashmi A, Al-Saad S, Ware J, Joseph RM, Greenblatt R, Gleason D, Ertelt JA, Apse KA, Bodell A, Partlow JN, Barry B, Yao H, Markianos K, Ferland RJ, Greenberg ME, Walsh CA. Identifying autism loci and genes by tracing recent shared ancestry. *Science* 2008;**321**:218—23.
17. **Wang K,** Zhang H, Ma D, Bucan M, Glessner JT, Abrahams BS, Salyakina D, Imielinski M, Bradfield JP, Sleiman PM, Kim CE, Hou C, Frackelton E, Chiavacci R, Takahashi N, Sakurai T, Rappaport E, Lajonchere CM, Munson J, Estes A, Korvatska O, Piven J, Sonnenblick LI, Alvarez Retuerto AI, Herman EI, Dong H, Hutman T, Sigman M, Ozonoff S, Klin A, Owley T, Sweeney JA, Brune CW, Cantor RM, Bernier R, Gilbert JR, Cuccaro ML, McMahon WM, Miller J, State MW, Wassink TH, Coon H, Levy SE, Schultz RT, Nurnberger JI, Haines JL, Sutcliffe JS, Cook EH, Minshew NJ, Buxbaum JD, Dawson G, Grant SF, Geschwind DH, Pericak-Vance MA, Schellenberg GD, Hakonarson H. Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature* 2009;**459**:528—33.
18. **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—85.
19. **Lord C,** Risi S, Lambrecht L, Cook EH Jr, Leventhal BL, DiLavore PC, Pickles A, Rutter M. 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—23.
20. **Sparrow S,** Balla D, Cicchetti D. *The Vineland Adaptive Behavior Scales - Interview Edition, Survey Form Manual*. Circle Pines, MN: American Guidance Service, 1984.
21. **Raven J,** Raven JC, Court JH. *Raven Manual: Section 3. Standard Progressive Matrices*. Oxford, England: Oxford Psychologists Press, 1998.
22. **Sherry ST,** Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 2001;**29**:308—11.
23. **Stewart AF,** Dandona S, Chen L, Assogba O, Belanger M, Ewart G, LaRose R, Doelle H, Williams K, Wells GA, McPherson R, Roberts R. Kinesin family member 6 variant Trp719Arg does not associate with angiographically defined coronary artery disease in the Ottawa Heart Genomics Study. *J Am Coll Cardiol* 2009;**53**:1471—2.
24. **Bierut LJ,** Agrawal A, Bucholz KK, Doheny KF, Laurie C, Pugh E, Fisher S, Fox L, Howells W, Bertelsen S, Hinrichs AL, Almasy L, Breslau N, Culverhouse RC, Dick DM, Edenberg HJ, Foroud T, Grucza RA, Hatsukami D, Hesselbrock V, Johnson EO, Kramer J, Krueger RF, Kuperman S, Lynskey M, Mann K, Neuman RJ, Nothen MM, Nurnberger JI Jr, Porjesz B, Ridinger M, Saccone NL, Saccone SF, Schuckit MA, Tischfield JA, Wang JC, Rietschel M, Goate AM, Rice JP. A genome-wide association study of alcohol dependence. *Proc Natl Acad Sci U S A* 2010;**107**:5082—7.
25. **Shaikh TH,** Gai X, Perin JC, Glessner JT, Xie H, Murphy K, O'Hara R, Casalunovo T, Conlin LK, D'Arcy M, Frackelton EC, Geiger EA, Haldeman-Englert C, Imielinski M, Kim CE, Medne L, Annaiah K, Bradfield JP, Dabaghyan E, Eckert A, Onyiah CC, Ostapenko S, Otieno FG, Santa E, Shaner JL, Skraban R, Smith RM, Elia J, Goldmuntz E, Spinner NB, Zackai EH, Chiavacci RM, Grundmeier R, Rappaport EF, Grant SF, White PS, Hakonarson H. High-resolution mapping and analysis of copy number variations in the human genome: a data resource for clinical and research applications. *Genome Res* 2009;**19**:1682—90.
26. **Wilkinson DG.** *In-situ Hybridization: A Practical Approach*. New York: Oxford University Press, 1992.
27. **IMGSAC.** A genomewide screen for autism: strong evidence for linkage to chromosomes 2q, 7q, and 16p. *Am J Hum Genet* 2001;**69**:570—81.
28. **Wang K,** Li M, Hadley D, Liu R, Glessner J, Grant SF, Hakonarson H, Bucan M. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res* 2007;**17**:1665—74.
29. **Conrad DF,** Pinto D, Redon R, Feuk L, Gokcumen O, Zhang Y, Aerts J, Andrews TD, Barnes C, Campbell P, Fitzgerald T, Hu M, Ihm CH, Kristiansson K, Macarthur DG, Macdonald JR, Onyiah I, Pang AW, Robson S, Stirrups K, Valsesia A, Walter K, Wei J, Tyler-Smith C, Carter NP, Lee C, Scherer SW, Hurles ME. Origins and functional impact of copy number variation in the human genome. *Nature* 2009;**464**:704—12.
30. **Abrahams BS,** Tentler D, Perederiy JV, Oldham MC, Coppola G, Geschwind DH. Genome-wide analyses of human perisylvian cerebral cortical patterning. *Proc Natl Acad Sci U S A* 2007;**104**:17849—54.
31. **Webber C,** Hehir-Kwa JY, Nguyen DQ, de Vries BB, Veltman JA, Ponting CP. Forging links between human mental retardation-associated CNVs and mouse gene knockout models. *PLoS Genet* 2009;**5**:e1000531.
32. **Weiss LA,** Arking DE, Daly MJ, Chakravarti A. A genome-wide linkage and association scan reveals novel loci for autism. *Nature* 2009;**461**:802—8.
33. **Soronen P,** Ollila HM, Antila M, Silander K, Palo OM, Kieseppa T, Lonnqvist J, Peltonen L, Tuulio-Henriksson A, Partonen T, Paunio T. Replication of GWAS of bipolar disorder: association of SNPs near CDH7 with bipolar disorder and visual processing. *Mol Psychiatry* 2010;**5**:4—6.
34. **Willemsen MH,** Fernandez BA, Bacino CA, Gerkes E, de Brouwer AP, Pfundt R, Sikkema-Raddatz B, Scherer SW, Marshall CR, Potocki L, van Bokhoven H, Kleefstra T. Identification of ANKRD11 and ZNF778 as candidate genes for autism and variable cognitive impairment in the novel 16q24.3 microdeletion syndrome. *Eur J Hum Genet* 2010;**18**:429—35.
35. **Bartlett CW,** Flax JF, Logue MW, Smith BJ, Vieland VJ, Tallal P, Brzustowicz LM. Examination of potential overlap in autism and language loci on chromosomes 2, 7, and 13 in two independent samples ascertained for specific language impairment. *Hum Hered* 2004;**57**:10—20.
36. **Bradford Y,** Haines J, Hutcheson H, Gardiner M, Braun T, Sheffield V, Cassavant T, Huang W, Wang K, Vieland V, Folstein S, Santangelo S, Piven J. Incorporating language phenotypes strengthens evidence of linkage to autism. *Am J Med Genet* 2001;**105**:539—47.
37. **Singh SM,** Castellani C, O'Reilly R. Autism meets schizophrenia via cadherin pathway. *Schizophr Res* 2010;**116**:293—4.
38. **Suzuki SC,** Furue H, Koga K, Jiang N, Nohmi M, Shimazaki Y, Katoh-Fukui Y, Yokoyama M, Yoshimura M, Takeichi M. Cadherin-8 is required for the first relay synapses to receive functional inputs from primary sensory afferents for cold sensation. *J Neurosci* 2007;**27**:3466—76.
39. **Shinawi M,** Schaaf CP, Bhatt SS, Xia Z, Patel A, Cheung SW, Lanpher B, Nagl S, Herding HS, Nevinny-Stickel C, Immken LL, Patel GS, German JR, Beaudet AL, Stankiewicz P. A small recurrent deletion within 15q13.3 is associated with a range of neurodevelopmental phenotypes. *Nat Genet* 2009;**41**:1269—71.
40. **Girirajan S,** Rosenfeld JA, Cooper GM, Antonacci F, Siswara P, Itsara A, Vives L, Walsh T, McCarthy SE, Baker C, Mefford HC, Kidd JM, Browning SR, Browning BL, Dickel DE, Levy DL, Ballif BC, Platky K, Farber DM, Gowans GC, Wetherbee JJ, Asamoah A, Weaver DD, Mark PR, Dickerson J, Garg BP, Ellingwood SA, Smith R, Banks VC, Smith W, McDonald MT, Hoo JJ, French BN, Hudson C, Johnson JP, Ozmore JR, Moeschler JB, Surti U, Escobar LF, El-Khechen D, Gorski JL, Kussmann J, Salbert B, Lacassie Y, Biser A, McDonald-McGinn DM, Zackai EH, Deardorff MA, Shaikh TH, Haan E, Friend KL, Fichera M, Romano C, Gecz J, DeLisi LE, Sebat J, King MC, Shaffer LG, Eichler EE. A recurrent 16p12.1 microdeletion supports a two-hit model for severe developmental delay. *Nat Genet* 2009;**42**:203—9.
41. **Vorstman J,** van Daalen E, Jalali G, Schmidt E, Pasterkamp R, de Jonge M, Hennekam E, Janson E, Staal W, van der Zwaag B, Burbach J, Kahn R, Emanuel B, H vE, Ophoff R. A double hit implicates DIAPH3 as an autism risk gene. *Mol Psychiatry*. 2010; Published online doi:10.1038/mp.2010.26
42. **Tan GC,** Doke TF, Ashburner J, Wood NW, Frackowiak RS. Normal variation in fronto-occipital circuitry and cerebellar structure with an autism-associated polymorphism of CNTNAP2. *Neuroimage* 2010;**53**:1030—42.