

# Evaluation of copy number variations reveals novel candidate genes in autism spectrum disorder-associated pathways

Anthony J. Griswold<sup>1,†</sup>, Deqiong Ma<sup>1,2,†</sup>, Holly N. Cukier<sup>1,†</sup>, Laura D. Nations<sup>1</sup>, Mike A. Schmidt<sup>1,2</sup>, Ren-Hua Chung<sup>1,2</sup>, James M. Jaworski<sup>1</sup>, Daria Salyakina<sup>1</sup>, Ioanna Konidari<sup>1</sup>, Patrice L. Whitehead<sup>1</sup>, Harry H. Wright<sup>3</sup>, Ruth K. Abramson<sup>3</sup>, Scott M. Williams<sup>4</sup>, Ramkumar Menon<sup>5</sup>, Eden R. Martin<sup>1,2</sup>, Jonathan L. Haines<sup>4</sup>, John R. Gilbert<sup>1,2</sup>, Michael L. Cuccaro<sup>1,2</sup> and Margaret A. Pericak-Vance<sup>1,2,\*</sup>

<sup>1</sup>John P. Hussman Institute for Human Genomics and <sup>2</sup>Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami Miller School of Medicine, Miami, FL, USA, <sup>3</sup>Department of Neuropsychiatry and Behavioral Science, School of Medicine, University of South Carolina, Columbia, SC, USA, <sup>4</sup>Center for Human Genetics Research, Vanderbilt University, Nashville, TN, USA and <sup>5</sup>Division of Maternal-Fetal Medicine Perinatal Research, Department of Obstetrics and Gynecology, The University of Texas Medical Branch at Galveston, Galveston, TX, USA

Received January 17, 2012; Revised March 9, 2012; Accepted April 20, 2012

Autism spectrum disorders (ASDs) are highly heritable, yet relatively few associated genetic loci have been replicated. Copy number variations (CNVs) have been implicated in autism; however, the majority of loci contribute to <1% of the disease population. Therefore, independent studies are important to refine associated CNV regions and discover novel susceptibility genes. In this study, a genome-wide SNP array was utilized for CNV detection by two distinct algorithms in a European ancestry case–control data set. We identify a significantly higher burden in the number and size of deletions, and disrupting more genes in ASD cases. Moreover, 18 deletions larger than 1 Mb were detected exclusively in cases, implicating novel regions at 2q22.1, 3p26.3, 4q12 and 14q23. Case-specific CNVs provided further evidence for pathways previously implicated in ASDs, revealing new candidate genes within the GABAergic signaling and neural development pathways. These include *DBI*, an allosteric binder of GABA receptors, *GABARAPL1*, the GABA receptor-associated protein, and *SLC6A11*, a postsynaptic GABA transporter. We also identified CNVs in *COBL*, deletions of which cause defects in neuronal cytoskeleton morphogenesis in model vertebrates, and *DNER*, a neuron-specific Notch ligand required for cerebellar development. Moreover, we found evidence of genetic overlap between ASDs and other neurodevelopmental and neuropsychiatric diseases. These genes include glutamate receptors (*GRID1*, *GRIK2* and *GRIK4*), synaptic regulators (*NRXN3*, *SLC6A8* and *SYN3*), transcription factor (*ZNF804A*) and RNA-binding protein *FMR1*. Taken together, these CNVs may be a few of the missing pieces of ASD heritability and lead to discovering novel etiological mechanisms.

\*To whom correspondence should be addressed at: University of Miami, Miller School of Medicine Miami, 1501 NW 10th Avenue, BRB-314 (M860), Miami, FL 33136, USA. Tel: +1 3052432308; Fax: +1 3052432396; Email: mpericak@med.miami.edu

<sup>†</sup>The authors wish it to be known that, in their opinion, the first three authors should be regarded as joint First Authors.

## INTRODUCTION

Autism spectrum disorders (ASDs) are among the most heritable of neurodevelopmental disorders (1–4). Given their strong genetic component, gene discovery in ASDs has been the focus of research for the past decade. To date, no single gene has been shown to account for a majority of ASD susceptibility. The limited success of gene discovery has been attributed to the genetic heterogeneity and phenotypic variability of ASDs. In fact, hundreds of candidate genes have been implicated. These candidates include rare variations with large effect sizes in syndromic forms of ASD, including tuberous sclerosis (*TSC1* and *TSC2*) (5,6), Rett syndrome (*MECP2*) (7) and Phelan–McDermid syndrome (*SHANK3*) (8). In addition, despite fruitful genome-wide association studies (GWAS) in many other genetic disorders (9), this approach in studying ASDs has revealed only a few associated common alleles with small effect sizes, many of which have not yet been replicated in independent studies (10–14). Therefore, studying low frequency variants with variable expression or reduced penetrance may be necessary to identify predisposing alleles.

Among aberrations of interest are copy number variations (CNVs). They represent a significant source of genetic diversity and contribute to disease susceptibility for several neurobehavioral phenotypes, including ASDs, mental retardation and schizophrenia (15–22). The accelerated discovery of CNVs coinciding with ASDs has provided evidence of their involvement in ASD etiology, although individual CNVs generally account for <1% each of attributable risk (23). Some CNVs have pleiotropic effects that associate with susceptibility to several neuropsychiatric disorders (12,20,24–32). Therefore, careful characterization of CNVs in ASD patients represents an important aspect of genetic research.

Gene-set enrichment analyses from both GWAS (33) and CNV studies (20,31) have revealed functional gene families and molecular pathways likely to contribute to the pathogenesis of ASDs. However, the gene-set enrichment statistical framework, particularly as applied to CNVs, is limited and potentially confounded (34). A joint analysis of two autism GWAS data sets implicated neuronal growth and axon guidance as well as cytoskeletal regulation and neuronal cell signaling (33). Among CNV gene-set analyses, the ubiquitin pathway and genes involved in neuronal cell adhesion have been consistently implicated in ASDs (20,31). A recent CNV study from the Autism Genome Project (AGP) identified an enrichment of CNVs disrupting GTPase/Ras signaling genes, and confirmed the involvement of genes responsible for cellular proliferation, projection and motility (20). Moreover, a recent study utilizing pathway analysis of *de novo* CNVs has implicated genes directing synapse development, axon targeting and neuron motility in autism risk (35,36). Therefore, given the heterogeneous nature of ASDs and their potential genetic causes, disruption of additional genes in these and related molecular pathways are expected to be involved in ASDs.

Further studies are required not only to refine our understanding of the roles of previously isolated genes and regions in ASDs but also to identify novel candidates within implicated molecular pathways. In this study, we used a

**Table 1.** Summary of identified autosomal CNVs

	All CNVs		Rare CNVs	
	ASD probands	Controls	ASD probands	Controls
Number of samples	813	592	773	535
Total number of CNVs	9940	8502	3209	1692
Duplications/deletions	4655/5285	5194/3308	1142/2003	835/758
Duplications/deletions > 1 Mb	34/18	86/0	17/18	40/0

combined family and case–control Caucasian data set to support previously associated ASD CNVs and identify unique, large and potentially pathogenic deletions as well as novel genes within molecular pathways previously connected to ASDs.

## RESULTS

### Global CNV burden analysis

In an attempt to recognize overall differences in CNVs between cases and controls, high confidence CNVs from 813 unrelated ASD cases and 592 control Caucasian samples (Table 1) were included in the CNV burden analysis following extensive CNV QC filtering (Table 2). To eliminate potential biases from various DNA sources and gender between cases and controls, burden analysis for deletions and duplications was performed separately and only for autosomes (see details in Materials and Methods).

We detected a significantly higher burden in the number and size of deletions in cases for both the total and rare (<1% frequency) CNV sets (Table 2) (all deletions: size:  $P = 9.90E-06$ ; number:  $P = 2.00E-05$ ; rare deletions: size:  $P = 0.04$ ; number:  $P = 9.90E-06$ ). In addition, the average number of genes disrupted by deletions per individual was significantly greater in cases when compared with controls (the number of genes disrupted: all deletions:  $P = 9.9E-06$ ; rare deletions:  $P = 9.9E-06$ ). We also observed an increase in the number of rare genic duplications (2.7 per case versus 2.2 per control,  $P = 0.01$ ), although the average size of the duplications was not different (176.6 kb in cases versus 201.1 kb controls,  $P = 0.95$ ).

### CNVs identified in previously implicated ASD regions

To identify specific CNVs which may contribute to autism risk, we first examined CNVs in eight well-established ASD associated regions (1q21, 7q11.23, 15q13, 15q11–13, 16p11.2, 17p11.2, 22q13.3 and 22q11.21). We detected and validated nine CNVs from these regions that overlap at least 50% of their length with regions previously implicated in autism risk within our data set, including reciprocal duplications and a deletion at 16p11.2. To determine whether these CNVs were concordant with disease, we determined the inheritance status in all families and segregation of variants within multiplex ASD families using the genotyping data from available family members and quantitative polymerase chain reaction (qPCR) (Table 3, Supplementary Material, Fig. S1).

**Table 2.** Burden analysis for autosomal CNVs

	All CNVs ASD probands	Controls	Burden ( <i>P</i> -value)	ASD probands	Rare CNVs Controls	Burden ( <i>P</i> -value)
Mean CNV size (kb) (SD)						
Deletions	102.7 (72.3)	87.9 (29.6)	<b>9.90E-06</b>	114.6	99	<b>0.04</b>
Duplications	195 (124.3)	214.1 (105.4)	1	176.6	201.1	0.95
Mean number of CNVs (genes)/individual(SD)						
Deletions	6.5 (4.9)	5.6 (2.3)	<b>2.00E-05</b>	2.5	1.3	<b>9.90E-06</b>
Genes deleted	6.3	3.9	<b>9.90E-06</b>	3	1.3	<b>9.90E-06</b>
Duplications	5.7 (3.6)	8.8 (3.8)	1	1.4	1.4	0.5
Genes duplicated	6.8	7.1	0.72	2.7	2.2	<b>0.01</b>

Boldface indicates significant *P*-values.

**Table 3.** CNVs in well-established autism candidate regions

Position (hg18)	Cytoband	CNV type	Family	Inheritance	Affected siblings with CNV <sup>a</sup>	Unaffected siblings with CNV <sup>b</sup>
chr1:144970329–145863421	1q21.1	Duplication	7804	<i>De novo</i>	1/2	0/0
			17331	Paternal	1/2	0/0
chr7:72355583–73778518	7q11.23	Duplication	37426	Maternal	2/2	0/0
chr15:26998850–28157296	15q13.1	Deletion	17630	Maternal	2/2	0/0
chr15:26849239–28157206	15q13.1	Deletion	7486	Maternal	1/1	0/2
chr15:21200234–26208861	15q11–13	Duplication	18054	<i>De novo</i>	1/1	0/0
			7791	Maternal	1/2	0/1
chr16:29554844–30085308	16p11.2	Duplication	7740	<i>De novo</i>	1/2	0/0
			17416	Paternal	1/1	0/0
chr16:29554844–30085308	16p11.2	Deletion	3574	<i>De novo</i>	1/1	0/0
chr17:17169258–20101517	17p11.2	Deletion	3586	<i>De novo</i>	1/1	0/0
chr22: 45159186–49562479	22q13.3	Deletion	3524	<i>De novo</i>	2/2	0/0
chr22: 19063496–19363371	22q11.21	Duplication	17164	<i>De novo</i>	1/1	0/1

<sup>a</sup>Number of affected siblings in the family with the CNV/total affected siblings.

<sup>b</sup>Number of unaffected siblings in the family with the CNV/total unaffected siblings.

Among these validated CNVs (Supplementary Material, Fig. S2) were two 5 Mb duplication at 15q11–13, one maternally inherited and the other *de novo* and a *de novo* 4 Mb deletion located at 22q13.3 which removes *SHANK3* and results in Phelan-McDermid syndrome. Other CNVs included an 1.4 Mb duplication at 7q11.23, an 1 Mb duplication at 22q11.21, two 900 kb duplications of 1q21.1, a 3 Mb deletion at 17p11.2, two 550 kb duplications and a reciprocal deletion of 16p11.2 and two 500 kb deletions at 15q13.3.

### Identification of large, novel deletions

Since no deletions >1 Mb were found in controls, we examined this subset of large and possibly pathogenic structural variations identified exclusively in cases. Eighteen such deletions ranging from 1.02 to 7.31 Mb were detected in 10 simplex families and 6 multiplex families (Table 4). Familial genotyping and CNV data as well as qPCR (Supplementary Material, Fig. S3) validation revealed that 7 of the 18 deletions were *de novo* and 10 were inherited (5 paternally, 5 maternally), while the inheritance status for one CNV could not be determined. Five of the 18 deletions overlapped with well-established ASD regions reported in Table 3. Five others have been identified in autistic patients in studies by the AGP (20,37). Independent autism CNV studies identified deletions at 5q35.3 (38), 10p15.2 (39), 10q11.22–23 (29) and Xp22.31 (40) that physically overlap at least half of a CNV

found in our data. This left four novel regions, each found in a single affected individual from a simplex family, including a paternally inherited deletion at 3p26.3 and *de novo* deletions at 2q22.1, 4q12 and 14q23 that do not overlap >20% of their length with any deletion previously cataloged in the Database of Genomic Variants (<http://projects.tcag.ca/variation/>) or the Autism CNV Database ([http://projects.tcag.ca/autism\\_500k/](http://projects.tcag.ca/autism_500k/)).

### Identification of potentially pathogenic CNVs and candidate genes from case-specific CNVs

Pathway enrichment analysis of case-specific CNVs did not reveal any gene ontology or KEGG pathways which were significantly overrepresented in our CNV set. Therefore, we evaluated our data to enrich for potentially pathogenic CNVs in ASDs. To do this, case-unique CNVs were selected. A total of 3855 CNVs met this criterion. Genes located within these CNV regions were compiled and their functions assessed by the literature review. First, CNVs in several genes previously associated with the autism risk, including deletions in *A2BP1*, *APBA2*, *GRIN2B* and *NRXN3* were identified in this way. Secondly, several novel candidate genes in molecular pathways or genetic networks implicated in autism etiology were identified, including genes in the GABAergic signaling pathway (*DBI*, *GABARAPL1* and *SLC6A11*) and genes acting in neural development and function (*COBL*, *DNER* and

**Table 4.** Large deletions (>1 Mb) found exclusively in ASD cases

Position (hg18)	Cytoband	Inheritance	Family	Affected siblings with CNV <sup>a</sup>	Unaffected siblings with CNV <sup>b</sup>
chr2:137234086–144547896	2q22.1–2	<i>De novo</i>	3771	1/1	0/0
chr3:36044–1633491	3p26.3	Paternal	18011	1/1	0/0
chr4:55697422–56874357	4q12	<i>De novo</i>	37036	1/1	0/0
chr5:117059397–118077675	5q23.1	Maternal	3522	1/2	0/0
chr5:179260834–180655211	5q35.3	<i>De novo</i>	17133	1/2 (Asperger's)	0/0
chr10:1163670–2368501	10p15.2	Paternal	7854	1/1 (PDD NOS)	0/0
chr10:48871970–51496386	10q11.22–23	Paternal	17500	1/1	1/2
chr13:22409581–23826468	13q12.12	Paternal	18011	1/1	0/0
chr13:103545187–104974554	13q33.1	Paternal	3524	1/2	0/0
chr14:63824114–65347410	14q23.2–3	<i>De novo</i>	17165	1/1	0/2
chr15:26849239–28157206 <sup>c</sup>	15q13.1	Maternal	7486	1/1 (Asperger's)	0/2
chr15:26998850–28157206 <sup>c</sup>	15q13.1	Maternal	17630	2/2	0/0
chr15:28705540–30421063 <sup>c</sup>	15q13.1–3	Maternal	17630	2/2	0/0
chr17:14040467–15423806	17p12	Maternal	3566	1/1	0/0
chr17:14085705–15416569	17p12	<i>De novo</i>	37541	1/1	0/1
chr17:17169258–20101517 <sup>c</sup>	17p11.2	<i>De novo</i>	3586	1/1	0/0
chr22:45159185–49562479 <sup>c</sup>	22q13.31–33	<i>De novo</i>	3524	2/2	0/0
chr23:6461752–8095053	Xp22.31	Maternal	7955	1/1	1/1

Non-autism diagnosis indicated in parentheses.

<sup>a</sup>Number of affected siblings in the family with the CNV/total affected siblings.

<sup>b</sup>Number of unaffected siblings in the family with the CNV/total unaffected siblings.

<sup>c</sup>Deletions in well-established autism candidate regions—reported in Table 3.

*GDNF*). In addition, our case-unique CNVs revealed several loci implicated in other neuropsychiatric disorders, including Fragile X syndrome (*FMR1*), mental retardation (*GRIK4*), obsessive compulsive disorder (*GRIK2*) and schizophrenia (*GRDI1*, *NRXN3*, *SLC6A11*, *SYN3*, *TIMP3* and *ZNF804A*). These CNVs were validated with TaqMan copy number assays and their inheritance status determined using familial genotyping data (Table 5, Supplementary Material, Fig. S4).

## DISCUSSION

In the present study, we identified several potentially pathogenic candidate regions and genes that may contribute to ASD etiology. At the global level, our results replicate a recent study which indicated an overall increase in the burden of rare genic deletions in ASD individuals (20). However, while the AGP reported an increase in rare deletions in affected individuals, our study finds that ASD cases have significantly larger deletions and more overall deletions (20). These differences may reflect data set-specific variation in factors, such as the ASD diagnostic criteria, ratios of simplex to multiplex families, regional or other sampling biases as well as array platform- and DNA source-origin biases.

Beyond identifying a higher overall burden of deletions in ASD cases compared with controls, we detected and validated CNVs at eight regions that have been previously associated with autistic traits; 1q21, 7q11.23, 15q13, 15q11–13, 16p11.2, 17p11.2, 22q13.3 and 22q11.21 (28,41–47). While many of these have been previously implicated by classical cytogenetic techniques, recent large-scale CNV studies using both genotyping and array CGH have also isolated these same CNVs as associated with autism (35,48,49). In our data, we identified two cases overlapping a classic cytogenetic aberration associated with autism, the 5 Mb maternally inherited 15q11–13 duplication. This duplication, which occurs in

~1% of all autism cases (44,50), falls between low copy repeats on chromosome 15 and duplicates many genes, including three GABA receptors. We also detected CNVs within regions known to be syndromic causes of autism; two duplications and one *de novo* deletion of 550 kb on 16p11.2 (28,30,51) and two duplications of 900 kb on 1q21.1 (41), both of which are recurrent CNVs due to their location between segmental duplications. In addition, a 3 Mb *de novo* deletion on 17p11.2, which deletes *RAI1*, has been previously implicated in both Smith-Magenis syndrome and an autistic syndrome (52). Additionally, one case carrying a complex 1.4 Mb duplication between segmental duplications at 7q11.23 partially overlaps the region deleted in Williams–Beuren syndrome in which duplications have previously been associated with autism (42,53). Finally, another individual carries the same 4 Mb deletion on 22q13.3 that causes Phelan–McDermid syndrome and contains the autism gene *SHANK3* (54,55). Moreover, several large CNVs that affect previously reported candidate genes were also identified. A 1 Mb duplication at 22q11.21 encompasses the entire *COMT* gene that has been implicated in developmental delay and autism (45,47,56,57). Two patients carry deletions of 15q13.1–3, thereby removing either the entire gene or several exons of *CHRNA7*, deletions of which have been shown to be involved in developmental delay, mental retardation and epilepsy (43,58,59). While all of these CNVs have strong evidence for being causative of ASDs, discordance within multiplex families has been previously noted for duplications and deletions on 16p11.2 (30,60) and duplications of 22q11.21 (61). Given such evidence, the lack of perfect concordance between these established CNVs and disease in some of our multiplex families (Supplementary Material, Fig. S1) suggests that there are probably additional genetic components contributing to disease in these families.

Besides the 5 large deletions discussed above (two at 15q13.1, 15q13.1–3, 17p11.2 and 22q12.31–33), 13 other

**Table 5.** Case-unique CNVs in ASD candidate genes

Position (hg 18)	Cytoband	Genes	CNV type	Inheritance	Family	Affected siblings with CNV <sup>a</sup>	Unaffected siblings with CNV <sup>b</sup>
Chr2:184744259–186437079	2q32.1	<i>FLJ44048,FSIP2,ZNF804A</i>	Duplication	Paternal	7678	2/2	0/0
chr2:119642873–119882278	2q14.2	<i>C2orf76,DBI,STEAP3</i>	Duplication	<i>De novo</i>	17113	1/1	0/0
chr2:230078358–230181261	2q36.3	<i>DNER</i>	Duplication	Paternal	17171	1/1	0/1
chr3:10823410–10943172	3p25.3	<i>SLC6A11</i>	Deletion	Paternal	7804	2/2	0/0
chr6:101260386–102979782	6q16.3	<i>ASCC3,GRIK2</i>	Deletion	Maternal	17182	2/2	0/0
chr7:51155279–51263037	7p12.1	<i>COBL</i>	Deletion	Maternal	39391	1/1	1/3 <sup>c</sup>
chr10:87754758–87839034	10q23.1	<i>GRID1</i>	Deletion	Paternal	37276	1/1	1/1
chr11:118685526–121040148	11q23.3–q24.1	<i>ARHGEF12,C1QTNF5,GRIK4,MCAM, MFRP,OAF,POU2F3,PVRL1,RNF26, SC5DL,SORL1,TBCEL,TECTATHY1, TMEM136, TRIM29,USP2</i>	Duplication	<i>De novo</i>	37341	1/1	0/1
chr12:10248041–10332152	12p13.2	<i>GABARAPL1</i>	Deletion	Paternal	37240	1/1	1/1
chr12:13753741–13788728	12p13.1	<i>GRIN2B</i>	Deletion	Maternal	7614	1/1	0/0
chr14:78373505–78506223	14q24.3–q31.1	<i>NRXN3</i>	Deletion	Maternal	7484	1/1	1/2
chr15:26849239–28157206	15q13.1–q13.2	<i>APBA2,KIAA0574,NDNL2,TJPI</i>	Deletion	Maternal	17630	2/2	0/1
chr16:6866262–7042998	16p13.2	<i>A2BP1</i>	Deletion	Paternal	7486	1/1	0/2
				Paternal	7636	2/2	0/0
				Paternal	7851	1/1	0/0
				Paternal	17855	1/1	1/1
chr22:31499975–31930759	22q12.3	<i>SYN3,TIMP3</i>	Duplication	Paternal	37350	1/1	1/1 <sup>d</sup>
				Maternal	39377	1/1	0/0
chrX:146676441–146930522	Xq27.3–q28	<i>ASFMR1,FMR1,FMR1NB</i>	Deletion	Maternal	18072	2/2	0/0

<sup>a</sup>Number of affected siblings in the family with the CNV/total affected siblings.

<sup>b</sup>Number of unaffected siblings in the family with the CNV/total unaffected siblings.

<sup>c</sup>Unaffected with CNV has speech delay.

<sup>d</sup>Unaffected with CNV has a questionable autism diagnosis.

deletions >1 Mb were observed exclusively in cases. These are of particular interest due to their possible pathogenic effect in neuropsychiatric diseases (62). Nine of these 12 have been identified as ASD candidate regions in previous CNV studies (20,29,37–40). The detection of deletions in these regions (5q23.1, 5q35.3, 10p15.2, 10q11.22-23, 13q12.12, 13q33.1, two at 17p12, and Xp22.31) lend further support to their harboring ASD pathogenic variations. Moreover, the mechanism of formation of these deletions, with the exception of 5q23.1, is predicted to be non-allelic homologous recombination due to the presence of flanking segmental duplications. Interestingly, four large deletions (2q22.1, 3p26.3, 4q12 and 14q23) are not flanked by known genomic instability elements and represent potentially novel contributors to autism etiology. A *de novo* deletion on 2q22.1 includes *HNMT*, a gene which regulates histamine metabolism in the brain and has been previously implicated in anxiety and alcoholism (63). We also identify a paternal deletion on 3p26.3 that deletes the neural adhesion molecules *CHLI* and *CNTN6*, which is near a previously reported autism-associated CNV (64). Furthermore, genes involved in circadian rhythms have been implicated in autism (65) and a new candidate, *CLOCK*, was removed in a *de novo* deletion on 4q12. Finally, pleiotropic effects of these novel, large deletions were found in an autistic patient with spherocytosis carrying a *de novo* 14q23 deletion that removes several neurodevelopmental genes, including *CHURC1*, *MTHFD1* and *PLEKHG3* as well as the red blood cell structural gene *SPTB* (66). As such, these substantial deletions point toward dysfunction of

neuron cell structure and adhesion as well as circadian regulation as contributors to autism.

In an attempt to find potentially pathogenic genes contributing to autism risk, we first used a pathway enrichment strategy to identify functional molecular networks which were overly represented by CNVs in cases. However, we did not identify a statistical overrepresentation of any category. While large-scale CNV studies report enrichment in their data sets (36,49), other studies have suggested that since the causes of autism are quite diverse, no singular mechanism or pathway is likely to explain the entire spectrum of risk (35). Therefore, we chose to search for novel candidate genes or gene clusters in rare, case-unique CNVs. This analysis revealed several genes previously associated with autism risk, including *A2BP1* (67), *APBA2* (68), *GRIN2B* (69,70), *NRXN3* (71) and *SHANK3* (72,73). Of particular interest among these genes is the splice factor *A2BP1* which has not only been associated with autism but regulates the expression of many genes misregulated in post-mortem autism brains (74). Novel members of pathways previously implicated in autism include three genes in the GABAergic signaling pathway (75–78). Due to the prior evidence of the importance of GABAergic genes in autism, it is possible that duplication of *DBI*, an allosteric binder of GABRA receptors (79), deletion of the GABRA receptor-associated protein *GABARAPL1* (80) or deletion of the first eight exons of *SLC6A11*(*GAT-3*), a postsynaptic GABA transporter (81), could affect signaling and lead to deficits in neuronal function. The role of each of these genes in ASDs warrants further investigation.

Beyond the GABAergic system, rare CNVs disrupting genes which play a role in neural development have been implicated in autism etiology (20). We identified case-unique CNVs that traverse genes affecting these processes. Our findings further implicate genes involved in neuronal cytoskeletal formation as we find a CNV which deletes exons 2–5 of *COBL*, knockouts of which show severe defects in neural cell cytoskeleton morphogenesis (82). In addition, our CNV analyses implicated signaling factors involved in neural development. We detected a duplication of exons 2–6 of *DNER*, a neuron-specific Notch ligand required for cerebellar development (83,84), containing a marker previously associated with autism (33), and deletion of the entire *GDNF* gene, a growth factor which fosters neuronal cell survival (85). The discovery of case-unique CNVs within these novel genes in previously associated molecular networks supports the importance of these pathways in autism as well as implicating new potential genetic risk factors, making them attractive candidates for follow-up.

Our data also identify several genes with neuronal functions that have been previously connected to other neuropsychiatric diseases. Such genetic overlap has been noted and previously discussed (86). We found CNVs in genes associated with schizophrenia, including two glutamate receptors, the first deleting the entire *GRIK2* gene (87,88) and the second being an intronic deletion in *GRID1* (89), a duplication of the first five exons of the synaptic protein *SYN3* (90), a deletion of two exons of the neuronal adhesion molecule *NRXN3* and duplication of the zinc finger *ZNF804A* (91). We also found CNVs in genes involved in other neurological disorders. The neurotransmitter transporter *SLC6A8* has been implicated in mental retardation (92) and we identified a deletion of its first eight exons. Expansions in the 5' untranslated region of *FMR1* are known to cause both Fragile X syndrome and Fragile X-associated tremor/ataxia syndrome (93) and we found a deletion of this entire gene. The overlap of these genes with other neuropsychiatric diseases suggests an etiological connection between these disorders and ASDs and points to the disruption of these genes as having broad neurological phenotypes.

In conclusion, we found there to be a significantly higher burden in the number and size of deletions carried by ASD individuals when compared with controls. Among the CNVs identified were several that overlapped with well-established autism-associated regions and candidate genes, further supporting their roles in autism etiology. Moreover, we isolated four large, novel deletions on 2q22.1, 3p26.3, 4q12 and 14q23 that include new genes and regions linked to ASDs. These findings provide further evidence for the involvement of previously molecular pathways, including the GABAergic signaling and neurodevelopmental pathways, while revealing several novel genes that may lead to new etiological mechanisms.

## MATERIALS AND METHODS

### Ascertainment of ASD families

This study is part of the Collaborative Autism Project (CAP), for which a total of 1084 simplex and multiplex autism families were ascertained with a range of self-reported ethnicities

through clinical groups at the John P. Hussman Institute for Human Genomics at the University of Miami (HIHG, Miami, FL, USA), the University of South Carolina (Columbia, SC, USA) and the Center for Human Genetics Research at Vanderbilt University (Nashville, TN, USA). Participating families were enrolled via support groups, advertisements and clinical and educational settings. All participants and families were ascertained using a standard protocol approved by the appropriate Institutional Review Boards. Written informed consent was obtained from parents as well as from minors who were able to give informed consent; for individuals unable to give consent due to age or developmental problems, assent was obtained whenever possible. Core inclusion criteria were as follows: (i) chronological age between 3 and 21 years of age, (ii) a presumptive clinical diagnosis of autism, (iii) expert clinical determination of an autism diagnosis using DSM-IV criteria supported by the Autism Diagnostic Interview-Revised (ADI-R) in the majority of cases and all available clinical information (94), (iv) minimal developmental level of 18 months as determined by the Vineland Adaptive Behavior Scale (VABS) (95) or the VABS-II (96) or IQ equivalent > 35. These minimal developmental levels assure that the ADI-R results are valid and reduce the likelihood of including individuals with severe mental retardation only. We excluded participants with severe sensory problems (e.g. visual impairment or hearing loss), significant motor impairments (e.g. failure to sit by 12 months or walk by 24 months) or who had a previously identified metabolic, genetic or progressive neurological disorders.

### Ascertainment of pediatric controls

Eight hundred ninety-six healthy children, including 664 self-reported Caucasians, were recruited as pediatric controls from two centers. Of these, 565 individuals aged 4 to 21 years old were sampled through the John P. Hussman Institute for Human Genomics (Miami, FL, USA). Participants were screened for eligibility by a series of preliminary questions to determine whether the individual either had been diagnosed with or had a parent or sibling with a developmental, behavioral, neurological or other disability or physical conditions. If none of those conditions were present, parents of young children or the participants were informed and signed the informed consent and completed the Social Communication Questionnaire (97) to screen for potential ASDs. The remaining 331 controls were part of a preterm birth study. Cord blood was collected at the Centennial Medical Center (Nashville, TN, USA) from women aged from 18 to 40 years old with term pregnancies (>37 weeks gestation), live, singleton births.

### Genotyping

Genomic DNA was purified from whole blood, cell line and saliva samples using Puregene chemistry on the Qiagen Autopure LS according to standard automated Qiagen protocols (Qiagen, Valencia, CA, USA). DNA samples were quantitated via the ND-8000 spectrophotometer and DNA quality was evaluated by gel electrophoresis on a 0.8% agarose gel. The concentration for all qualified samples was normalized

to 50 ng/ $\mu$ l and samples were arrayed in Matrix 0.5 ml 2D bar-coded tubes in racks of 96. Sample identity was confirmed by genotyping eight SNPs using TaqMan allelic discrimination assays (Applied Biosystems, Foster City, CA, USA) and assessing for concordance with historical data. All recruited participants, available siblings and parents and recruited controls were genotyped using either Illumina's Human 1M-v1\_C Beadchip or the 1M-DuoV3 Beadchip. The samples were processed according to Illumina Procedures for the Infinium II<sup>®</sup> assay (Illumina Inc., San Diego, CA, USA). The protocol was automated using the Tecan EVO-1 to further enhance efficiency and consistency (Tecan Group Ltd, Männedorf, Switzerland). A single quality-control DNA sample was tested during each run to ensure reproducibility of results between runs. Data were extracted with the Illumina<sup>®</sup> Beadstudio software from data files created in the Illumina BeadArray reader. Samples and markers with call rates below 95% were excluded from analysis and a GenCall cutoff score of 0.15 was used for all Infinium II<sup>®</sup> products.

### Genotyping and CNV quality control (QC)

Out of 1084 autistic families and 896 pediatric controls, samples with a genotype call rate >97% and only cases without excessive Mendelian inconsistencies within the pedigree (>2% among all tested SNPs) were processed further for CNV QC. Following this family-based sample QC, 1063 families, encompassing 1301 ASD cases, remained for CNV calling. The signal intensity QC used the parameters generated in the CNV detection program PennCNV (98). Samples with <0.30 standard deviation (SD) of the normalized intensity log R ratio (LRR) and a correlation of LRR to wave model ranged between  $-0.2 < X < 0.4$  were included. The remaining samples were examined further and excluded if the total number of CNV calls per individuals exceeded three standard deviations from the mean number of CNV calls. Thus, only samples with a total CNV call count <142 were included. A total number of 1174 cases, including 184 affected siblings from multiplex families, and 771 controls fit the above criteria suggesting that overall data quality would not remove a significant portion of our sample.

Next, unrelated cases and controls were selected for population stratification checking. To maintain independence of this data set, only one ASD case per family was analyzed. Therefore, 184 affected siblings were removed from the initial CNV analysis, but later used to examine concordance in multiplex families. In addition, two samples were taken out due to overlap with the previously published AGP data set (20). One trisomy X female and one duplicated Xp male were also removed. These steps left 986 cases and 771 controls for population stratification analysis. We performed clustering with principle component analysis using the EIGENSTRAT program (99). In order to recognize independent ethnicity clusters in our data, we combined it with HapMap genotyping data from 10 ethnicities. A total of 813 cases and 592 controls clustered with CEU, and remained for the final case-control CNV analysis. These included 260 cases from multiplex families, 483 from simplex families and 62 from families with siblings having autistic-like traits without a confirmed ASD diagnosis and 8 from families with MZ twins. Of 813 cases,

649 had both parents survive all QC steps to be used in the family-based trio function within PennCNV to determine whether CNVs were inherited or *de novo*.

### High confidence CNV identification

The QuantiSNP (100) and PennCNV (98) algorithms were both utilized for CNV identification. Intensity wave artifacts across the genome, which interfere with accurate inference of CNVs, were corrected by both programs. The Bayesian factor (BF) scores provided by QuantiSNP were calculated to reflect the accuracy of a CNV call. A BF score of 15 was used as a cutoff to include only reliable copy number alterations. Finally, high confidence CNV calls were defined as those called by both algorithms in a person and overlapped for at least 20% of their length.

### CNV filtering

PLINK (1.0.7 version) (101) was employed to remove CNVs that overlapped for at least 50% of their length with segmental duplication regions (downloaded from the UCSC genome browser, <http://genome.ucsc.edu>), peri-centromeric regions (recommended by PennCNV), pseudoautosomal regions on the X chromosome and CNVs containing >70% GC content. CNVs were only selected for analysis that spanned at least five SNPs and had a length of >30 kb as these filtering criteria in a previous large-scale study achieved a validation rate of >90% (49). Rare CNVs were defined as CNVs with a frequency of <1% in the entire case-control data set. In addition, case-unique CNVs that did not overlap more than 50% of their length with any control CNVs identified in the CHOP CNV database (<http://cnv.chop.edu/>) were used to compile a list of potential candidate genes for validation.

### Bias checking

Since the total number of CNVs per individual could be influenced by gender, genotyping platform or DNA source, several comparisons were made to evaluate these potential biases in cases and controls. In our data set, there is a distinct male to female bias in cases (707/106) when compared with controls (289/303). While females have significantly more CNVs identified on the X chromosome, no difference was found on autosomes ( $P = 0.903$  for controls;  $P = 0.795$  for cases). We also examined the potential difference between two genotyping arrays used because all control samples were genotyped on Illumina's 1M-duo-v3 while 634 out of 813 cases were genotyped on Illumina's 1M-v1-C chip. Data from cases indicate that more duplications were identified using the 1M-duo-v3 chip ( $P < 0.001$ ). Moreover, potential biases arising from different DNA sources [blood (438/813 cases versus 299/592 controls), saliva (2/813 cases versus 293/592 controls), cell lines (49/813 cases) and DNA-origin source unknown (324 cases)] were assessed. Consistent with the findings from other studies, a potential difference was present in the number of duplications across the samples from various DNA sources [cases genotyped on 1M-v1-C: one-way ANOVA  $P = 0.002$  (blood, cell-line and unknown); controls genotyped on 1M-duo-v3:  $t$ -test  $P = 0.001$  (blood, saliva)].

## CNV validation

We selected a subset of candidate genes and regions for validation by real-time PCR. These included four distinct categories; (i) regions or genes previously reported to carry ASD specific CNVs, (ii) novel deletions over 1 Mb, (iii) CNVs containing genes with a known neuronal function, and (iv) CNVs that overlapped genes involved in other neuropsychiatric and neurodevelopmental disorders. For each gene or region of interest, the individual determined to carry the CNV as well as all available family members were selected for validation using Applied Biosystems' TaqMan Copy Number Assays. An assay in every CNV region was selected using the GeneAssist software available at <http://www5.appliedbiosystems.com/tools/cnv/> (102). Assays were run in quadruplicate, multiplexed 10  $\mu$ l reactions containing 1X TaqMan Universal PCR Master Mix, VIC-labeled RNaseP reference assay and the FAM-labeled test assay. Reactions were performed in 384-well plates on the ABI 7900HT Real-Time PCR System and data were collected with Applied Biosystems SDSv2.3. Data were analyzed with a manual Ct threshold of 0.2 and an automatic baseline and copy numbers were determined using the CopyCaller software with the RNaseP level as the calibrator for copy number.

## CNV analysis

Analysis was carried out in two data sets: (i) the entire CNV set and (ii) the rare CNV set (frequency < 1%). The genome-wide CNV burden was assessed in deletions and duplication separately due to the potential bias identified from genotyping array differences on duplications. In addition, the burden analysis was carried out in autosomal CNVs only due to a skewed male to female ratio in cases when compared with controls. Significance of the mutational burden comparisons was determined via permutation (100 000 time, one-sided test) using PLINK (version 1.0.7). Pathway enrichment for CNV sets was performed using WebGestalt (<http://bioinfo.vanderbilt.edu/webgestalt/>). The CNV lists were uploaded using the Illumina Omni-1 Quad SNP gene list as the background gene list. Enrichment was performed using a hypergeometric test requiring at least two genes in a pathway with multiple test correction proposed by Benjamini and Hochberg (103).

## SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

## ACKNOWLEDGEMENTS

We thank all participants and their family members in this study and personnel at the John P. Hussman Institute for Human Genomics (HIHG) including staff at the HIHG Biorepository and HIHG autism ascertainment team as well as the members of the Vanderbilt Center for Human Genetics Research. A subset of the participants was ascertained while Dr Pericak-Vance was a faculty member at Duke University.

## FUNDING

This work was supported by the National Institutes of Health (R01MH080647, P01NS026630); and a gift from the Hussman Foundation.

## REFERENCES

- Ritvo, E.R., Freeman, B.J., Mason-Brothers, A., Mo, A. and Ritvo, A.M. (1985) Concordance for the syndrome of autism in 40 pairs of afflicted twins. *Am. J. Psychiatry*, **142**, 74–77.
- Steffenburg, S., Gillberg, C., Hellgren, L., Andersson, L., Gillberg, I.C., Jakobsson, G. and Bohman, M. (1989) A twin study of autism in Denmark, Finland, Iceland, Norway, and Sweden. *J. Child Psychol. Psychiatry*, **30**, 405–416.
- Folstein, S.E. and Piven, J. (1991) Etiology of autism: genetic influences. *Pediatrics*, **87**, 767–773.
- 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.
- Jones, A.C., Daniells, C.E., Snell, R.G., Tachataki, M., Idziaszczyk, S.A., Krawczak, M., Sampson, J.R. and Cheadle, J.P. (1997) Molecular genetic and phenotypic analysis reveals differences between TSC1 and TSC2 associated familial and sporadic tuberous sclerosis. *Hum. Mol. Genet.*, **6**, 2155–2161.
- Wong, V. (2006) Study of the relationship between tuberous sclerosis complex and autistic disorder. *J. Child Neurol.*, **21**, 199–204.
- Amir, R.E., Van den Veyver, I.B., Wan, M., Tran, C.Q., Francke, U. and Zoghbi, H.Y. (1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat. Genet.*, **23**, 185–188.
- Wilson, H.L., Wong, A.C., Shaw, S.R., Tse, W.Y., Stapleton, G.A., Phelan, M.C., Hu, S., Marshall, J. and McDermid, H.E. (2003) Molecular characterisation of the 22q13 deletion syndrome supports the role of haploinsufficiency of SHANK3/PROSAP2 in the major neurological symptoms. *J. Med. Genet.*, **40**, 575–584.
- Klein, R.J., Xu, X., Mukherjee, S., Willis, J. and Hayes, J. (2010) Successes of genome-wide association studies. *Cell*, **142**, 350–351; author reply 353–355.
- Ma, D., Salyakina, D., Jaworski, J.M., Konidari, I., Whitehead, P.L., Andersen, A.N., Hoffman, J.D., Slifer, S.H., Hedges, D.J., Cukier, H.N. *et al.* (2009) A genome-wide association study of autism reveals a common novel risk locus at 5p14.1. *Ann. Hum. Genet.*, **73**, 263–273.
- Wang, K., Zhang, H., Ma, D., Bucan, M., Glessner, J.T., Abrahams, B.S., Salyakina, D., Imielinski, M., Bradfield, J.P., Sleiman, P.M. *et al.* (2009) Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature*, **459**, 528–533.
- Weiss, L.A. and Arking, D.E. Gene Discovery Project of Johns Hopkins & the Autism Consortium, Daly, M.J. and Chakravarti, A. (2009) A genome-wide linkage and association scan reveals novel loci for autism. *Nature*, **461**, 802–808.
- Anney, R., Klei, L., Pinto, D., Regan, R., Conroy, J., Magalhaes, T.R., Correia, C., Abrahams, B.S., Sykes, N., Pagnamenta, A.T. *et al.* (2010) A genome-wide scan for common alleles affecting risk for autism. *Hum. Mol. Genet.*, **19**, 4072–4082.
- Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorf, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A. *et al.* (2009) Finding the missing heritability of complex diseases. *Nature*, **461**, 747–753.
- Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., Yamrom, B., Yoon, S., Krasnitz, A., Kendall, J. *et al.* (2007) Strong association of de novo copy number mutations with autism. *Science*, **316**, 445–449.
- Walsh, T., McClellan, J.M., McCarthy, S.E., Addington, A.M., Pierce, S.B., Cooper, G.M., Nord, A.S., Kusenda, M., Malhotra, D., Bhandari, A. *et al.* (2008) Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science*, **320**, 539–543.
- International Schizophrenia Consortium. (2008) Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature*, **455**, 237–241.



18. Xu, B., Roos, J.L., Levy, S., van Rensburg, E.J., Gogos, J.A. and Karayiorgou, M. (2008) Strong association of de novo copy number mutations with sporadic schizophrenia. *Nat. Genet.*, **40**, 880–885.
19. Elia, J., Gai, X., Xie, H.M., Perin, J.C., Geiger, E., Glessner, J.T., D'arcy, M., Deberardinis, R., Frackelton, E., Kim, C. *et al.* (2010) Rare structural variants found in attention-deficit hyperactivity disorder are preferentially associated with neurodevelopmental genes. *Mol. Psychiatry*, **15**, 637–646.
20. Pinto, D., Pagnamenta, A.T., Klei, L., Anney, R., Merico, D., Regan, R., Conroy, J., Magalhaes, T.R., Correia, C., Abrahams, B.S. *et al.* (2010) Functional impact of global rare copy number variation in autism spectrum disorders. *Nature*, **466**, 368–372.
21. Awadalla, P., Gauthier, J., Myers, R.A., Casals, F., Hamdan, F.F., Griffing, A.R., Cote, M., Henrion, E., Spiegelman, D., Tarabeux, J. *et al.* (2010) Direct measure of the de novo mutation rate in autism and schizophrenia cohorts. *Am. J. Hum. Genet.*, **87**, 316–324.
22. Rees, E., Moskva, V., Owen, M.J., O'Donovan, M.C. and Kirov, G. (2011) De novo rates and selection of schizophrenia-associated copy number variants. *Biol. Psychiatry*, **70**, 1109–1114.
23. Miles, J.H. (2011) Autism spectrum disorders—a genetics review. *Genet. Med.*, **13**, 278–294.
24. Schaaf, C.P., Goin-Kochel, R.P., Nowell, K.P., Hunter, J.V., Aleck, K.A., Cox, S., Patel, A., Bacino, C.A. and Shinawi, M. (2011) Expanding the clinical spectrum of the 16p11.2 chromosomal rearrangements: three patients with syringomyelia. *Eur. J. Hum. Genet.*, **19**, 152–156.
25. Niklasson, L. and Gillberg, C. (2010) The neuropsychology of 22q11 deletion syndrome. A neuropsychiatric study of 100 individuals. *Res. Dev. Disabil.*, **31**, 185–194.
26. Gu, W. and Lupski, J.R. (2008) CNV and nervous system diseases—what's new? *Cytogenet. Genome Res.*, **123**, 54–64.
27. Marshall, C.R., Noor, A., Vincent, J.B., Lionel, A.C., Feuk, L., Skaug, J., Shago, M., Moessner, R., Pinto, D., Ren, Y. *et al.* (2008) Structural variation of chromosomes in autism spectrum disorder. *Am. J. Hum. Genet.*, **82**, 477–488.
28. Weiss, L.A., Shen, Y., Korn, J.M., Arking, D.E., Miller, D.T., Fossdal, R., Saemundsen, E., Stefansson, H., Ferreira, M.A., Green, T. *et al.* (2008) Association between microdeletion and microduplication at 16p11.2 and autism. *N. Engl. J. Med.*, **358**, 667–675.
29. Christian, S.L., Brune, C.W., Sudi, J., Kumar, R.A., Liu, S., Karamohamed, S., Badner, J.A., Matsui, S., Conroy, J., McQuaid, D. *et al.* (2008) Novel submicroscopic chromosomal abnormalities detected in autism spectrum disorder. *Biol. Psychiatry*, **63**, 1111–1117.
30. Kumar, R.A., Karamohamed, S., Sudi, J., Conrad, D.F., Brune, C., Badner, J.A., Gilliam, T.C., Nowak, N.J., Cook, E.H. Jr, Dobyns, W.B. *et al.* (2008) Recurrent 16p11.2 microdeletions in autism. *Hum. Mol. Genet.*, **4**, 628–638.
31. Glessner, J.T., Wang, K., Cai, G., Korvatska, O., Kim, C.E., Wood, S., Zhang, H., Estes, A., Brune, C.W., Bradfield, J.P. *et al.* (2009) Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature*, **459**, 569–573.
32. Crespi, B., Stead, P. and Elliot, M. (2010) Evolution in health and medicine sackler colloquium: comparative genomics of autism and schizophrenia. *Proc. Natl Acad. Sci. USA*, **107**(Suppl. 1), 1736–1741.
33. Hussman, J.P., Chung, R.H., Griswold, A.J., Jaworski, J.M., Salyakina, D., Ma, D., Konidari, I., Whitehead, P.L., Vance, J.M., Martin, E.R. *et al.* (2011) A noise-reduction GWAS analysis implicates altered regulation of neurite outgrowth and guidance in autism. *Mol. Autism*, **2**, 1.
34. Raychaudhuri, S., Korn, J.M. and McCarroll, S.A., International Schizophrenia Consortium, Altshuler, D., Sklar, P., Purcell, S. and Daly, M.J. (2010) Accurately assessing the risk of schizophrenia conferred by rare copy-number variation affecting genes with brain function. *PLoS Genet.*, **6**, e1001097.
35. Levy, D., Ronemus, M., Yamrom, B., Lee, Y.H., Leotta, A., Kendall, J., Marks, S., Lakshmi, B., Pai, D., Ye, K. *et al.* (2011) Rare de novo and transmitted copy-number variation in autistic spectrum disorders. *Neuron*, **70**, 886–897.
36. Gilman, S.R., Iossifov, I., Levy, D., Ronemus, M., Wigler, M. and Vitkup, D. (2011) Rare de novo variants associated with autism implicate a large functional network of genes involved in formation and function of synapses. *Neuron*, **70**, 898–907.
37. Autism Genome Project Consortium, Szatmari, P., Paterson, A.D., Zwaigenbaum, L., Roberts, W., Brian, J., Liu, X.Q., Vincent, J.B., Skaug, J.L., Thompson, A.P., Senman, L. *et al.* (2007) Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat. Genet.*, **39**, 319–328.
38. Buxbaum, J.D., Cai, G., Nygren, G., Chaste, P., Delorme, R., Goldsmith, J., Rastam, M., Silverman, J.M., Hollander, E., Gillberg, C. *et al.* (2007) Mutation analysis of the NSD1 gene in patients with autism spectrum disorders and macrocephaly. *BMC Med. Genet.*, **8**, 68.
39. Lindstrand, A., Malmgren, H., Verri, A., Benetti, E., Eriksson, M., Nordgren, A., Anderlid, B.M., Golovleva, I., Schoumans, J. and Blennow, E. (2010) Molecular and clinical characterization of patients with overlapping 10p deletions. *Am. J. Med. Genet. A*, **152A**, 1233–1243.
40. Li, F., Shen, Y., Kohler, U., Sharkey, F.H., Menon, D., Coulleaux, L., Malan, V., Rio, M., McMullan, D.J., Cox, H. *et al.* (2010) Interstitial microduplication of Xp22.31: causative of intellectual disability or benign copy number variant? *Eur. J. Med. Genet.*, **53**, 93–99.
41. Mefford, H.C., Sharp, A.J., Itsara, A., Jiang, Z., Buysse, K., Huang, S., Maloney, V.K., Crolla, J.A., Baralle, D. *et al.* (2008) Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. *N. Engl. J. Med.*, **359**, 1685–1699.
42. Van der Aa, N., Rooms, L., Vandeweyer, G., van den Ende, J., Reyniers, E., Fichera, M., Romano, C., Delle Chiaie, B., Mortier, G., Menten, B. *et al.* (2009) Fourteen new cases contribute to the characterization of the 7q11.23 microduplication syndrome. *Eur. J. Med. Genet.*, **52**, 94–100.
43. Miller, D.T., Shen, Y., Weiss, L.A., Korn, J., Anselm, I., Bridgemohan, C., Cox, G.F., Dickinson, H., Gentile, J., Harris, D.J. *et al.* (2009) Microdeletion/duplication at 15q13.2q13.3 among individuals with features of autism and other neuropsychiatric disorders. *J. Med. Genet.*, **46**, 242–248.
44. Baker, P., Piven, J., Schwartz, S. and Patil, S. (1994) Brief report: duplication of chromosome 15q11–13 in two individuals with autistic disorder. *J. Autism Dev. Disord.*, **24**, 529–535.
45. Gadow, K.D., Roohi, J., DeVincent, C.J., Kirsch, S. and Hatchwell, E. (2009) Association of COMT (Val158Met) and BDNF (Val66Met) gene polymorphisms with anxiety, ADHD and tics in children with autism spectrum disorder. *J. Autism Dev. Disord.*, **39**, 1542–1551.
46. Bonaglia, M.C., Giorda, R., Mani, E., Aceti, G., Anderlid, B.M., Baroncini, A., Pramparo, T. and Zuffardi, O. (2006) Identification of a recurrent breakpoint within the SHANK3 gene in the 22q13.3 deletion syndrome. *J. Med. Genet.*, **43**, 822–828.
47. Edelman, L., Pandita, R.K., Spiteri, E., Funke, B., Goldberg, R., Palanisamy, N., Chaganti, R.S., Magenis, E., Shprintzen, R.J. and Morrow, B.E. (1999) A common molecular basis for rearrangement disorders on chromosome 22q11. *Hum. Mol. Genet.*, **8**, 1157–1167.
48. Sanders, S.J., Ercan-Sencicek, A.G., Hus, V., Luo, R., Murtha, M.T., Moreno-De-Luca, D., Chu, S.H., Moreau, M.P., Gupta, A.R., Thomson, S.A. *et al.* (2011) Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron*, **70**, 863–885.
49. Pinto, D., Pagnamenta, A.T., Klei, L., Anney, R., Merico, D., Regan, R., Conroy, J., Magalhaes, T.R., Correia, C., Abrahams, B.S. *et al.* (2010) Functional impact of global rare copy number variation in autism spectrum disorders. *Nature*, **466**, 368–372.
50. Schinzel, A. (1981) Duplication-deletion with partial trisomy 1q and partial monosomy 3p resulting from a maternal reciprocal translocation rcp (1;3) (q32;p25). *J. Med. Genet.*, **18**, 64–68.
51. Shinawi, M., Liu, P., Kang, S.H., Shen, J., Belmont, J.W., Scott, D.A., Probst, F.J., Craigen, W.J., Graham, B., Pursley, A. *et al.* (2009) Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioral problems, dysmorphism, epilepsy, and abnormal head size. *J. Med. Genet.*, **47**, 332–341.
52. Laje, G., Morse, R., Richter, W., Ball, J., Pao, M. and Smith, A.C. (2010) Autism spectrum features in Smith-Magenis syndrome. *Am. J. Med. Genet. C Semin. Med. Genet.*, **154C**, 456–462.
53. Berg, J.S., Brunetti-Pierri, N., Peters, S.U., Kang, S.H., Fong, C.T., Salamone, J., Freedman, D., Hannig, V.L., Prock, L.A., Miller, D.T. *et al.* (2007) Speech delay and autism spectrum behaviors are frequently associated with duplication of the 7q11.23 Williams-Beuren syndrome region. *Genet. Med.*, **9**, 427–441.
54. Phelan, M.C. (2008) Deletion 22q13.3 syndrome. *Orphanet. J. Rare Dis.*, **3**, 14.
55. Delahaye, A., Toutain, A., Aboura, A., Dupont, C., Tabet, A.C., Benzacken, B., Elion, J., Verloes, A., Pipiras, E. and Drunat, S. (2009) Chromosome 22q13.3 deletion syndrome with a de novo interstitial

- 22q13.3 cryptic deletion disrupting SHANK3. *Eur. J. Med. Genet.*, **52**, 328–332.
56. Mukaddes, N.M. and Herguner, S. (2007) Autistic disorder and 22q11.2 duplication. *World J. Biol. Psychiatry*, **8**, 127–130.
57. Ramelli, G.P., Silacci, C., Ferrarini, A., Cattaneo, C., Visconti, P. and Pescia, G. (2008) Microduplication 22q11.2 in a child with autism spectrum disorder: clinical and genetic study. *Dev. Med. Child Neurol.*, **50**, 953–955.
58. Szafranski, P., Schaaf, C.P., Person, R.E., Gibson, I.B., Xia, Z., Mahadevan, S., Wiszniewska, J., Bacino, C.A., Lalani, S., Potocki, L. *et al.* (2010) Structures and molecular mechanisms for common 15q13.3 microduplications involving CHRNA7: Benign or pathological? *Hum. Mutat.*, **31**, 840–850.
59. Shinawi, M., Schaaf, C.P., Bhatt, S.S., Xia, Z., Patel, A., Cheung, S.W., Lanpher, B., Nagl, S., Herding, H.S., Nevinny-Stickel, C. *et al.* (2009) A small recurrent deletion within 15q13.3 is associated with a range of neurodevelopmental phenotypes. *Nat. Genet.*, **41**, 1269–1271.
60. Fernandez, B.A., Roberts, W., Chung, B., Weksberg, R., Meyn, S., Szatmari, P., Joseph-George, A.M., Mackay, S., Whitten, K., Noble, B. *et al.* (2010) Phenotypic spectrum associated with de novo and inherited deletions and duplications at 16p11.2 in individuals ascertained for diagnosis of autism spectrum disorder. *J. Med. Genet.*, **47**, 195–203.
61. Goodship, J., Cross, I., Scambler, P. and Burn, J. (1995) Monozygotic twins with chromosome 22q11 deletion and discordant phenotype. *J. Med. Genet.*, **32**, 746–748.
62. Kirov, G., Grozeva, D., Norton, N., Ivanov, D., Mantripragada, K.K. and Holmans, P., International Schizophrenia Consortium, Wellcome Trust Case Control Consortium, Craddock, N., Owen, M.J. *et al.* (2009) Support for the involvement of large copy number variants in the pathogenesis of schizophrenia. *Hum. Mol. Genet.*, **18**, 1497–1503.
63. Oroszi, G., Enoch, M.A., Chun, J., Virkkunen, M. and Goldman, D. (2005) Thr105Ile, a functional polymorphism of histamine N-methyltransferase, is associated with alcoholism in two independent populations. *Alcohol. Clin. Exp. Res.*, **29**, 303–309.
64. Cottrell, C.E., Bir, N., Varga, E., Alvarez, C.E., Bouyain, S., Zernzach, R., Thrush, D.L., Evans, J., Trimarchi, M., Butter, E.M. *et al.* (2011) Contactin 4 as an autism susceptibility locus. *Autism Res.*, **4**, 189–199.
65. Nicholas, B., Rudrasingham, V., Nash, S., Kirov, G., Owen, M.J. and Wimpory, D.C. (2007) Association of Per1 and Npas2 with autistic disorder: support for the clock genes/social timing hypothesis. *Mol. Psychiatry*, **12**, 581–592.
66. Griswold, A.J., Ma, D., Sacharow, S.J., Robinson, J.L., Jaworski, J.M., Wright, H.H., Abramson, R.K., Lybaek, H., Oyen, N., Cuccaro, M.L. *et al.* (2011) A de novo 1.5 mb microdeletion on chromosome 14q23.2–23.3 in a patient with autism and spherocytosis. *Autism Res.*, **4**, 221–227.
67. Martin, C.L., Duvall, J.A., Ilkin, Y., Simon, J.S., Arreaza, M.G., Wilkes, K., Alvarez-Retuerto, A., Whichello, A., Powell, C.M., Rao, K. *et al.* (2007) Cytogenetic and molecular characterization of A2BP1/FOX1 as a candidate gene for autism. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*, **144B**, 869–876.
68. Babatz, T.D., Kumar, R.A., Sudi, J., Dobyms, W.B. and Christian, S.L. (2009) Copy number and sequence variants implicate APBA2 as an autism candidate gene. *Autism Res.*, **2**, 359–364.
69. Myers, R.A., Casals, F., Gauthier, J., Hamdan, F.F., Keebler, J., Boyko, A.R., Bustamante, C.D., Piton, A.M., Spiegelman, D., Henrion, E. *et al.* (2011) A population genetic approach to mapping neurological disorder genes using deep resequencing. *PLoS Genet.*, **7**, e1001318.
70. O’Roak, B.J., Deriziotis, P., Lee, C., Vives, L., Schwartz, J.J., Girirajan, S., Karakoc, E., Mackenzie, A.P., Ng, S.B., Baker, C. *et al.* (2011) Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat. Genet.*, **43**, 585–589.
71. Vaags, A.K., Lionel, A.C., Sato, D., Goodenberger, M., Stein, Q.P., Curran, S., Ogilvie, C., Ahn, J.W., Drmic, I., Senman, L. *et al.* (2011) Rare deletions at the neurexin 3 locus in autism spectrum disorder. *Am. J. Hum. Genet.*, **90**, 133–141.
72. Durand, C.M., Betancur, C., Boeckers, T.M., Bockmann, J., Chaste, P., Fauchereau, F., Nygren, G., Rastam, M., Gillberg, I.C., Anckarsater, H. *et al.* (2007) Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. *Nat. Genet.*, **39**, 25–27.
73. Sykes, N.H., Toma, C., Wilson, N., Volpi, E.V., Sousa, I., Pagnamenta, A.T., Tancredi, R., Battaglia, A., Maestrini, E., Bailey, A.J. *et al.* (2009) Copy number variation and association analysis of SHANK3 as a candidate gene for autism in the IMGSAC collection. *Eur. J. Hum. Genet.*, **17**, 1347–1353.
74. Voineagu, I., Wang, X., Johnston, P., Lowe, J.K., Tian, Y., Horvath, S., Mill, J., Cantor, R.M., Blencowe, B.J. and Geschwind, D.H. (2011) Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature*, **474**, 380–384.
75. Ma, D.Q., Whitehead, P.L., Menold, M.M., Martin, E.R., Ashley-Koch, A.E., Mei, H., Ritchie, M.D., DeLong, G.R., Abramson, R.K., Wright, H.H. *et al.* (2005) Identification of significant association and gene-gene interaction of GABA receptor subunit genes in autism. *Am. J. Hum. Genet.*, **77**, 377–388.
76. Collins, A.L., Ma, D., Whitehead, P.L., Martin, E.R., Wright, H.H., Abramson, R.K., Hussman, J.P., Haines, J.L., Cuccaro, M.L., Gilbert, J.R. *et al.* (2006) Investigation of autism and GABA receptor subunit genes in multiple ethnic groups. *Neurogenetics*, **7**, 167–174.
77. Delahanty, R.J., Kang, J.Q., Brune, C.W., Kistner, E.O., Courchesne, E., Cox, N.J., Cook, E.H. Jr, Macdonald, R.L. and Sutcliffe, J.S. (2009) Maternal transmission of a rare GABRB3 signal peptide variant is associated with autism. *Mol. Psychiatry*, **16**, 86–96.
78. Hussman, J.P. (2001) Suppressed GABAergic inhibition as a common factor in suspected etiologies of autism. *J. Autism Dev. Disord.*, **31**, 247–248.
79. Gray, P.W., Glaister, D., Seeburg, P.H., Guidotti, A. and Costa, E. (1986) Cloning and expression of cDNA for human diazepam binding inhibitor, a natural ligand of an allosteric regulatory site of the gamma-aminobutyric acid type A receptor. *Proc. Natl Acad. Sci. USA*, **83**, 7547–7551.
80. Mansuy, V., Boireau, W., Fraichard, A., Schlick, J.L., Jouvenot, M. and Delage-Mourroux, R. (2004) GEC1, a protein related to GABARAP, interacts with tubulin and GABA(A) receptor. *Biochem. Biophys. Res. Commun.*, **325**, 639–648.
81. Minelli, A., DeBiasi, S., Brecha, N.C., Zuccarello, L.V. and Conti, F. (1996) GAT-3, a high-affinity GABA plasma membrane transporter, is localized to astrocytic processes, and it is not confined to the vicinity of GABAergic synapses in the cerebral cortex. *J. Neurosci.*, **16**, 6255–6264.
82. Ahuja, R., Pinyol, R., Reichenbach, N., Custer, L., Klingensmith, J., Kessels, M.M. and Qualmann, B. (2007) Cordon-bleu is an actin nucleation factor and controls neuronal morphology. *Cell*, **131**, 337–350.
83. Eiraku, M., Tohgo, A., Ono, K., Kaneko, M., Fujishima, K., Hirano, T. and Kengaku, M. (2005) DNER acts as a neuron-specific notch ligand during Bergmann glial development. *Nat. Neurosci.*, **8**, 873–880.
84. Saito, S.Y. and Takeshima, H. (2006) DNER as key molecule for cerebellar maturation. *Cerebellum*, **5**, 227–231.
85. Lin, L.F., Doherty, D.H., Lile, J.D., Bektess, S. and Collins, F. (1993) GDNF: a glial cell line-derived neurotrophic factor for midbrain dopaminergic neurons. *Science*, **260**, 1130–1132.
86. Carroll, L.S. and Owen, M.J. (2009) Genetic overlap between autism, schizophrenia and bipolar disorder. *Genome Med.*, **1**, 102.
87. Shibata, H., Shibata, A., Ninomiya, H., Tashiro, N. and Fukumaki, Y. (2002) Association study of polymorphisms in the GluR6 kainate receptor gene (GRIK2) with schizophrenia. *Psychiatry Res.*, **113**, 59–67.
88. Bah, J., Quach, H., Ebstein, R.P., Segman, R.H., Melke, J., Jamain, S., Rietschel, M., Modai, I., Kanas, K., Karni, O. *et al.* (2004) Maternal transmission disequilibrium of the glutamate receptor GRIK2 in schizophrenia. *Neuroreport*, **15**, 1987–1991.
89. Treutlein, J., Muhleisen, T.W., Frank, J., Mattheisen, M., Herms, S., Ludwig, K.U., Treutlein, T., Schmael, C., Strohmaier, J., Bosshenz, K.V. *et al.* (2009) Dissection of phenotype reveals possible association between schizophrenia and glutamate receptor delta 1 (GRID1) gene promoter. *Schizophr. Res.*, **111**, 123–130.
90. Porton, B., Ferreira, A., DeLisi, L.E. and Kao, H.T. (2004) A rare polymorphism affects a mitogen-activated protein kinase site in synapsin III: possible relationship to schizophrenia. *Biol. Psychiatry*, **55**, 118–125.
91. Riley, B., Thiselton, D., Maher, B.S., Bigdeli, T., Wormley, B., McMichael, G.O., Fanous, A.H., Vladimirov, V., O’Neill, F.A., Walsh, D. *et al.* (2010) Replication of association between schizophrenia and ZNF804A in the Irish case-control study of schizophrenia sample. *Mol. Psychiatry*, **15**, 29–37.
92. Clark, A.J., Rosenberg, E.H., Almeida, L.S., Wood, T.C., Jakobs, C., Stevenson, R.E., Schwartz, C.E. and Salomons, G.S. (2006) X-linked creatine transporter (SLC6A8) mutations in about 1% of males with mental retardation of unknown etiology. *Hum. Genet.*, **119**, 604–610.

93. Kremer, E.J., Pritchard, M., Lynch, M., Yu, S., Holman, K., Baker, E., Warren, S.T., Schlessinger, D., Sutherland, G.R. and Richards, R.I. (1991) Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)<sub>n</sub>. *Science*, **252**, 1711–1714.
94. Autism Genetics Resource Exchange. (2008) *AGRE*. Autism Speaks, Los Angeles, CA.
95. S.S., Balla, D and Cicchetti, D. (1984) *Vineland Adaptive Behavior Scales*. AGS Publishing, Circle Pines, MN.
96. Sparrow, S.S., Cicchetti, D.V. and Balla, D. (2005) *Vineland Adaptive Behavior Scales*. 2nd edn. AGS, Circle Pines, MN.
97. Rutter, M., Bailey, A., Berument, S.K., LeCouteur, A., Lord, C. and Pickles, A. (2003) *Social Communication Questionnaire (SCQ)*. Western Psychological Services, Los Angeles, CA.
98. Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S.F., Hakonarson, H. and Bucan, M. (2007) PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res.*, **17**, 1665–1674.
99. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A. and Reich, D. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.*, **38**, 904–909.
100. Colella, S., Yau, C., Taylor, J.M., Mirza, G., Butler, H., Clouston, P., Bassett, A.S., Seller, A., Holmes, C.C. and Ragoussis, J. (2007) QuantiSNP: an objective Bayes Hidden-Markov model to detect and accurately map copy number variation using SNP genotyping data. *Nucleic Acids Res.*, **35**, 2013–2025.
101. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–575.
102. Applied Biosystems by Life Technologies. (2011) *Applied Biosystems* Life Technologies, Carlsbad, CA.
103. Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B. (Meth.)*, **57**, 289–300.