Journal of Genetics and Genomics 45 (2018) 527–538

Contents lists available at ScienceDirect

Journal of Genetics and Genomics

Journal homepage: www.journals.elsevier.com/journal-of-genetics-and-genomics/

Original research

Genomic landscapes of Chinese sporadic autism spectrum disorders revealed by whole-genome sequencing

Jinyu Wu [a, 1], Ping Yu [a, 1], Xin Jin [b, 1], Xiu Xu [c, 1], Jinchen Li [d], Zhongshan Li [a], Mingbang Wang [b], Tao Wang [a], Xueli Wu [b], Yi Jiang [a], Wanshi Cai [e], Junpu Mei [b], Qingjie Min [a], Qiong Xu [c], Bingrui Zhou [c], Hui Guo [d], Ping Wang [f], Wenhao Zhou [c], Zhengmao Hu [d], Yingrui Li [b], Tao Cai [a], Yi Wang [c], Kun Xia [d, *], Yong-Hui Jiang [f, *], Zhong Sheng Sun [a, e, *]

[a] Institute of Genomic Medicine, Wenzhou Medical University, Wenzhou 325000, China
[b] BGI-Shenzhen, Shenzhen 518083, China
[c] Department of Child Healthcare, Children's Hospital of Fudan University, Shanghai 200032, China
[d] State Key Laboratory of Medical Genetics, Central South University, Changsha 410078, China
[e] Beijing Institutes of Life Science, Chinese Academy of Sciences, Beijing 100101, China
[f] Department of Pediatrics and Neurobiology, Duke University School of Medicine, Durham, NC 27710, USA

ARTICLE INFO

Article history:
Received 15 May 2018
Received in revised form
25 August 2018
Accepted 9 September 2018
Available online 21 October 2018

Keywords:
Autism spectrum disorders
De novo mutations
Microcephaly-associated genes
Whole-genome sequencing

ABSTRACT

Autism spectrum disorder (ASD) is a neurodevelopmental disorder with considerable clinical and genetic heterogeneity. In this study, we identified all classes of genomic variants from whole-genome sequencing (WGS) dataset of 32 Chinese trios with ASD, including *de novo* mutations, inherited variants, copy number variants (CNVs) and genomic structural variants. A higher mutation rate (Poisson test, $P < 2.2 \times 10^{-16}$) in exonic (1.37×10^{-8}) and 3′-UTR regions (1.42×10^{-8}) was revealed in comparison with that of whole genome (1.05×10^{-8}). Using an integrated model, we identified 87 potentially risk genes ($P < 0.01$) from 4832 genes harboring various rare deleterious variants, including *CHD8* and *NRXN2*, implying that the disorders may be in favor to multiple-hit. In particular, frequent rare inherited mutations of several microcephaly-associated genes (*ASPM*, *WDR62*, and *ZNF335*) were found in ASD. In chromosomal structure analyses, we found four *de novo* CNVs and one *de novo* chromosomal rearrangement event, including a *de novo* duplication of *UBE3A*-containing region at 15q11.2-q13.1, which causes Angelman syndrome and microcephaly, and a disrupted *TNR* due to *de novo* chromosomal translocation t(1; 5)(q25.1; q33.2). Taken together, our results suggest that abnormalities of centrosomal function and chromatin remodeling of the microcephaly-associated genes may be implicated in pathogenesis of ASD. Adoption of WGS as a new yet efficient technique to illustrate the full genetic spectrum in complex disorders, such as ASD, could provide novel insights into pathogenesis, diagnosis and treatment.

1. Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder with considerable clinical and genetic heterogeneity, characterized by significant impairments in reciprocal social interactions, restricted interests and repetitive behaviors. They have been reported to affect 1 of 88 children in the United States (Mullins et al., 2016). ASD has a significantly high heritability (52%–90%) and hundreds of genes or more are involved in ASD etiology (Mullins et al., 2016; Yin and Schaaf, 2017). Recent whole-exome sequencing (WES) (De Rubeis et al.,. 2014; Iossifov et al., 2014; Toma et al., 2014; Tammimies et al., 2015) and whole-genome sequencing (WGS) (Michaelson et al., 2012; Jiang et al., 2013; Yuen et al., 2015; Yuen et al., 2016; RK et al., 2017; Turner et al., 2017) studies by us and others have successfully identified a

* Corresponding authors.
E-mail addresses: xiakun@sklmg.edu.cn (K. Xia), yong-hui.jiang@duke.edu (Y.-H. Jiang), sunzs@mail.biols.ac.cn (Z.S. Sun).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jgg.2018.09.002

significant set of *de novo* or rare inherited mutations in exonic regions for ASD, such as *SCN2A* and *GRIN2B*, which are implicated in sodium-selective channel and synaptic function, respectively.

Despite the progress in the discovery of risk loci and genes, less than 20% of individuals with ASD have an identifiable genetic etiology and it is anticipated that additional genomic variants can be discovered using improved technologies, like WGS (Vorstman et al., 2017). WGS is an appropriate but also challenging application for a deeper characterization of the underlying genomic architecture of this complex psychiatric disorder (RK et al., 2017). For instance, WGS study of 235 ASD subjects has revealed an important role of *de novo* structural variants (SVs) in ASD (Brandler et al., 2016). Another WGS study found that 189 of 2620 (7.2%) subjects diagnosed with ASD carry one or more pathogenic chromosomal abnormalities, megabase copy number variants (CNVs) (RK et al., 2017). However, most current WGS studies have been conducted primarily for Caucasian populations. Few studies have delineated a comprehensive genomic landscape of genetic variants in Chinese ASD patients.

In this study, we performed WGS of 32 Chinese ASD trios and searched all classes of genomic variants in WGS, including *de novo* mutations (DNMs), rare inherited mutation, CNVs and genomic SVs. Using an integrated model to analyze both *de novo* and transmitted rare deleterious variants, we found that many rare genetic variant-containing genes were strongly implicated in psychiatric disorders and neuronal activities. In addition, we detected frequent rare inherited mutations of several microcephaly-associated genes in ASD. Extensive genetic and biomedical analyses of associated signaling pathways involved in the pathogenesis of ASD are also presented.

2. Results

2.1. 32 ASD trios and WGS

We recruited a cohort of 32 Chinese ASD patients (30 males and 2 females) between the ages of 3 and 13 years (6.16 ± 2.41, mean $\pm$ SD) and their parents (fathers 33.41 ± 6.02; mothers 28.72 ± 4.44). ASD diagnosis was carried out using ADOS/DSM-IV-TR by certified professional and experienced clinicians. Furthermore, 21 patients (65.6%) were evaluated to have intellectual disability (ID) by WISC-R examination. 96 whole genomes from the 32 trios were sequenced at BGI-Shenzhen using HiSeq2000 with an average coverage of 36 folds per sample (Tables S1 and S2). Using BWA/GATK (Li and Durbin, 2009; McKenna et al., 2010), we identified 309,774,761 single nucleotide variants (SNVs) (Table S3) and 72,018,176 small insertions/deletions (InDels) (Table S4).

We aimed to identify all classes of genomic variants, including DNMs, inherited variants, copy CNVs and SVs, simultaneously (Fig. S1). Using the Random Forest classifier (Michaelson et al., 2012; Jiang et al., 2013), a total of 1917 *de novo* SNVs (59.91 per sample, Table S5) and 174 *de novo* InDels (5.44 per sample, Table S6) were identified from all probands. To evaluate the power of DNM calling, we sequenced all 64 identified *de novo* SNVs and InDels in coding regions and UTRs. Only five of them (7.81%) were false positive (Table S7). We also sequenced additional 41 potential *de novo* SNVs obtained from in-house-developed bioinformatic method (Supplementary data) and confirmed 2 of them false negative (4.88%) (Table S7). The average *de novo* SNV rate was estimated to be $\sim 1.05 \times 10^{-8}$, similar to previous observation (Michaelson et al., 2012).

Investigation of all DNM spectrum revealed a higher transition rate (43%) of strong G-C bonds to weak A-T hydrogen bonds (e.g., C > T/G > A) than other type of transitions (Fig. 1A). Interestingly, the C > T/G > A alteration was dramatically increased in the form of CpG (Fig. 1B), suggesting possible events of spontaneous deamination of methylated cytosine to thymine in ASD. Also, higher ApT contexts were found in the form of T > C/A > G substitution (Fig. 1C). However, whether this transition pattern is specific for ASD or also presents in other psychiatric diseases needs further observations.

Analysis of *de novo* SNV distribution revealed 30% higher mutation rate (Poisson test, $P < 2.2 \times 10^{-16}$) in exonic region (1.37×10^{-8}) compared with that of whole genome (1.05×10^{-8}) (Fig. 1D). 68% of patients (22/32) were found with $\geq$1 DNMs in coding regions (Table S5) with a ratio of nonsynonymous *vs.* synonymous variants of 5:1. A higher rate of *de novo* SNVs was also observed in 3′-UTR region (1.42×10^{-8}), suggesting a potential selective advantage for autism. Furthermore, *de novo* SNVs were found to be mainly derived from the fathers with a male to female ratio of 292:88 in whole genome. Multiple regression analysis showed that the number of *de novo* SNVs was significantly correlated with the paternal age ($P = 0.005$, Fig. 1E) but not the maternal age ($P = 0.228$), consistent with previous observation in Caucasian ASD patients (Kong et al., 2012; Michaelson et al., 2012; Jiang et al., 2013). The rate of *de novo* InDel ($\sim 9.45 \times 10^{-10}$, Table S6) was 11.02 folds lower than that of DNMs. Approximate 80% of these InDels occurred within or adjacent to homopolymeric or polynucleotide repeats and 72% of the first base of insert sequence were identical to the base being inserted, which was likely induced by a replication slippage (Abecasis et al., 2012).

2.2. Integrated model for identification of ASD risk genes

To prioritize ASD risk genes, we applied a recently developed TADA program to incorporate both *de novo* and inherited genetic variants by a gene-based likelihood model involving parameters for allele frequencies and gene-specific penetrance (He et al., 2013). As a result, 87 genes were revealed to be significantly enriched in comparison with random expectation from 4832 genes with rare variants in the coding region, suggesting that these genes may be potential ASD risk genes ($P_{TADA} < 0.01$, Fig. 2A and Table S8).

Among the *de novo* events (Tables S9 and S10), we observed that 2 genes, *CHD8* (c.5661-5662delAG) and *NRXN2* (c.809-810insG), possessed frame-shift/truncated mutations, which were previously reported to be causative for ASD and define a subtype of autism (Bernier et al., 2014). Furthermore, a *de novo* frame-shift mutation of *RFX3* (c.1835-1836delTA) gene was found in proband A5, which is a strong candidate gene for ASD considering its crucial role in early brain development by regulating *GLI3* and *FGF8* activities (Benadiba et al., 2012). Damaging missense DNMs were also identified in 6 other genes, including *BRD3*, *CHD9*, *CHRM3*, *CNTNAP4, PASK*, and *PXDN*. Functional evidence for these genes supported a strong association with ASD (Table S10). For example, *CNTNAP4* is implicated in synaptic development and its family gene *CNTNAP2* is strongly implicated in ASD (Ebert and Greenberg, 2013). *BRD3* (c.731G > A) is calcium-sensitive and highly expressed in neural cells during amphibian development (Bibonne et al., 2013). *CHRM3* (c.1423A > T) influenced acetylcholine pathway in ASD as well as learning and memory (Petersen et al., 2013). In addition, 5 genes (*WDR62*, *DIP2B*, *DGKA*, *PC*, and *TF*), linked to other neuropsychiatric disorders previously, were considered to be potential ASD candidates. For example, mutation of *DIP2B* (c.1337T > C) occurred at a highly conserved site in the AMP-dependent synthetase/ligase domain, likely affecting neurocognitive function and ASD risk (Winnepenninckx et al., 2007). *WDR62* (c.1949G > A) is a possible candidate for ASD, in which recessive mutations can cause severe cerebral cortical malformations and microcephaly (Bilguvar et al., 2010; Yu et al., 2010). With respect to the role of other six *de novo* mutated genes (*CD276*, *HGFAC*, *KRT7*, *PLCH1*, *PRDM15*, and

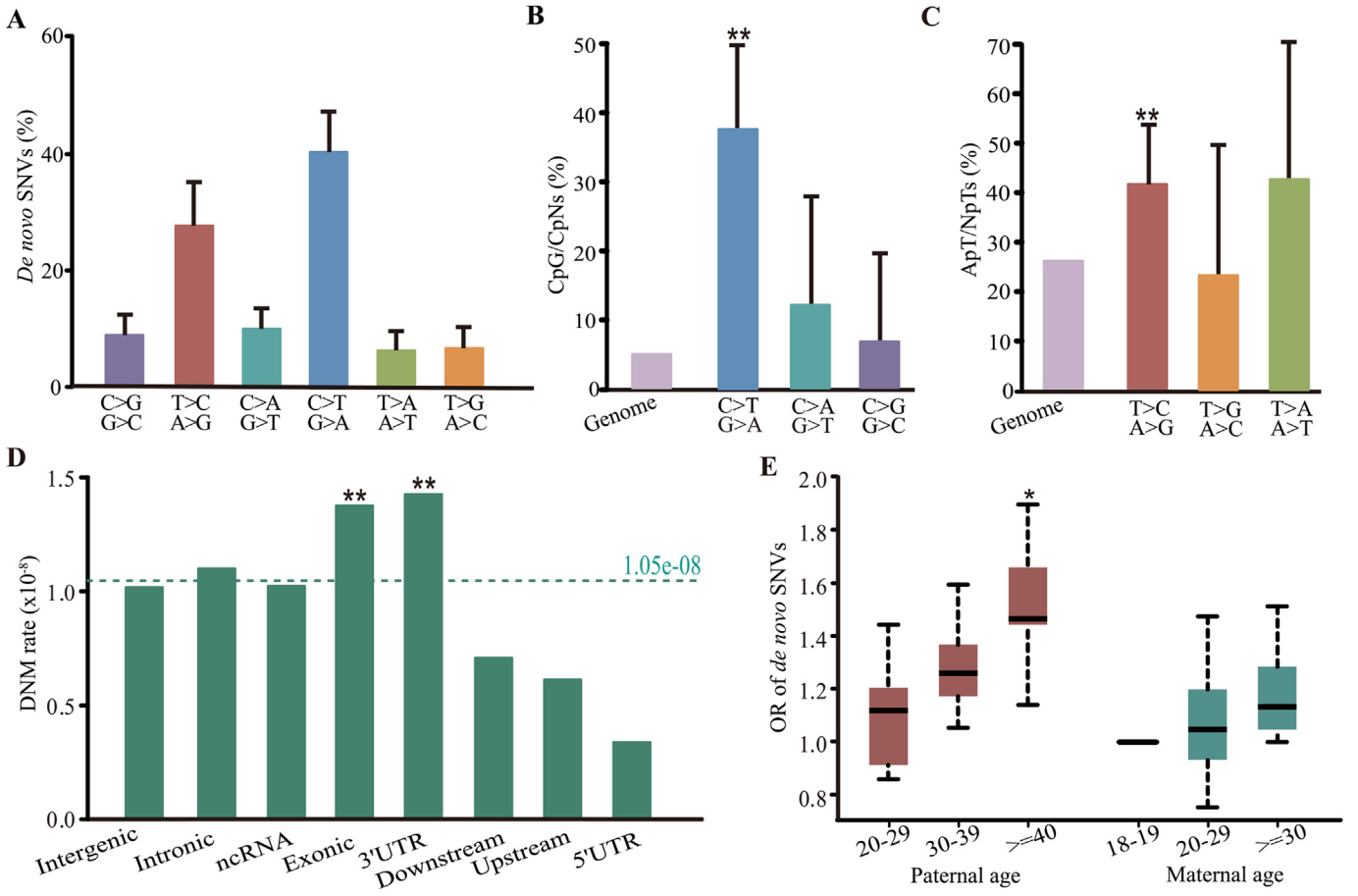

Fig. 1. Characterization of genome-wide DNMs in ASD. **A**: Mutation spectrum of *de novo* SNVs. Analysis of substitutions of DNM base composition revealed higher (43%) transition of strong G-C bonds than weak A-T hydrogen bonds (e.g., C > T/G > A). **B**: Average percentages of three kinds of cytosine substitutions at CpG. The C > T/G > A is dramatically increased in the 5′-CpG regions compared with the percentages in reference genome (hg19) (37.5%, two-tailed execute binomial test, $P < 2.2e\text{-}16$). **C**: Average percentage of three kinds of thymine substitutions at NpT regions. The distribution of T > C/A > G is dramatically increased in the 3′-base of ApT dinucleotides compared with the percentages in reference genome (hg19) (42.5%, two-tailed execute binomial test, $P < 2.2e\text{-}16$). **D**: Distribution of the number of *de novo* SNVs in different genomic regions as indicated. **E:** Association of the number of *de novo* SNVs and parent ages. The numbers of *de novo* SNVs correlate with the paternal age ($P = 0.005$, 95% CI), but not the maternal age ($P = 0.228$, 95% CI). Upstream, 1 kb prior to the 5′-UTR; downstream, 1 kb following 3′-UTR regions.

VAV2) in neuronal activities, further investigation is required to clarify their association with ASD.

Genes with inherited rare loss-of-function or damaging variants may also contribute to ASD pathogenesis (Lim et al., 2013; Stein et al., 2013; Yu et al., 2013; Toma et al., 2014). Based on the TADA program, several inherited hemizygous (*PDZD4*, *RENBP*, and *SLC6A14*), homozygous (*LFNG*), and heterozygous (*CSMD3*, *NRXN2*, and *PASK*) mutations were predicted to be ASD risk and additional 18 mutated genes (*GPR112* and *PLXNA3*, etc.) were associated with ID or neuronal activities (Tables S9 and S11–S13).

2.3. Frequent mutations of microcephaly-associated genes in ASD

Unexpectedly, we detected frequent rare novel mutations (1 DNM, 11 transmitted) in three microcephaly-associated genes (*ASPM*, *WDR62*, and *ZNF335*) in 10 probands of our cohort (31.3%, Fig. 2B–F). As summarized in Table 1, five different mutations in *ASPM*, a gene encoding a centrosomal protein regulating the Wnt pathway during neurogenesis, were identified in six probands (6/32, 18.75%, $P_{TADA} = 0.000346$; Tables S8 and S13). All of them were accompanied with ID (IQ score 40–54). None of these *ASPM* mutations were found in 1300 Chinese individuals (in-house data at BGI-Shenzhen) or 6500 Caucasian individuals (ESP6500 database). The frequency of damaging SNVs in *ASPM* in our ASD cohort was significantly higher than that of 1300 Chinese (Fisher's exact test, $P = 1.4 \times 10^{-8}$). Furthermore, the significant difference for the frequency of damaging SNVs in *ASPM* was also observed between ASD patients from ~1576 published Caucasians autistic exomes (He et al., 2013) and 6500 Caucasian individuals (Fisher's exact test, $P = 2.2 \times 10^{-16}$). At clinic, homozygous mutations of *ASPM* have been found to account for approximately 40% of patients with primary autosomal recessive microcephaly (MCPH5), with a majority in Middle Eastern and South Asian descent (Tan et al., 2014). In animal studies, the total number of cells in cortical plate was significantly decreased in both $Aspm^{+/-}$ ($P < 0.05$) and $Aspm^{-/-}$ mice ($P < 0.01$), compared with that in $Aspm^{+/+}$ mice (Fujimori et al., 2014), indicating a dose-dependent effect of *ASPM* on proliferation and differentiation of neural progenitor cells. These findings support a strong association between these heterozygous deleterious mutations in *ASPM* and the cohort with ASD.

Further analysis of additional microcephaly-associated genes identified 30 transmitted novel mutations of 12 microcephaly-associated genes in 14 probands (Table 1). Four of them (*VPS13B*, *PCNT*, *CDK5RAP2*, and *MBD5*) carried recurrent novel mutations. A single novel transmitted mutation was found in the rest of eight genes (*ATR*, *CENPJ*, *DYNC1H1*, *KIF5C*, *CEP63*, *KIF11*, *TUBGCP6*, and *NIN*), among which mutations of *DYNC1H1*, *KIF5C* and *KIF11* cause microcephaly-associated autosomal dominant disorders. Taken

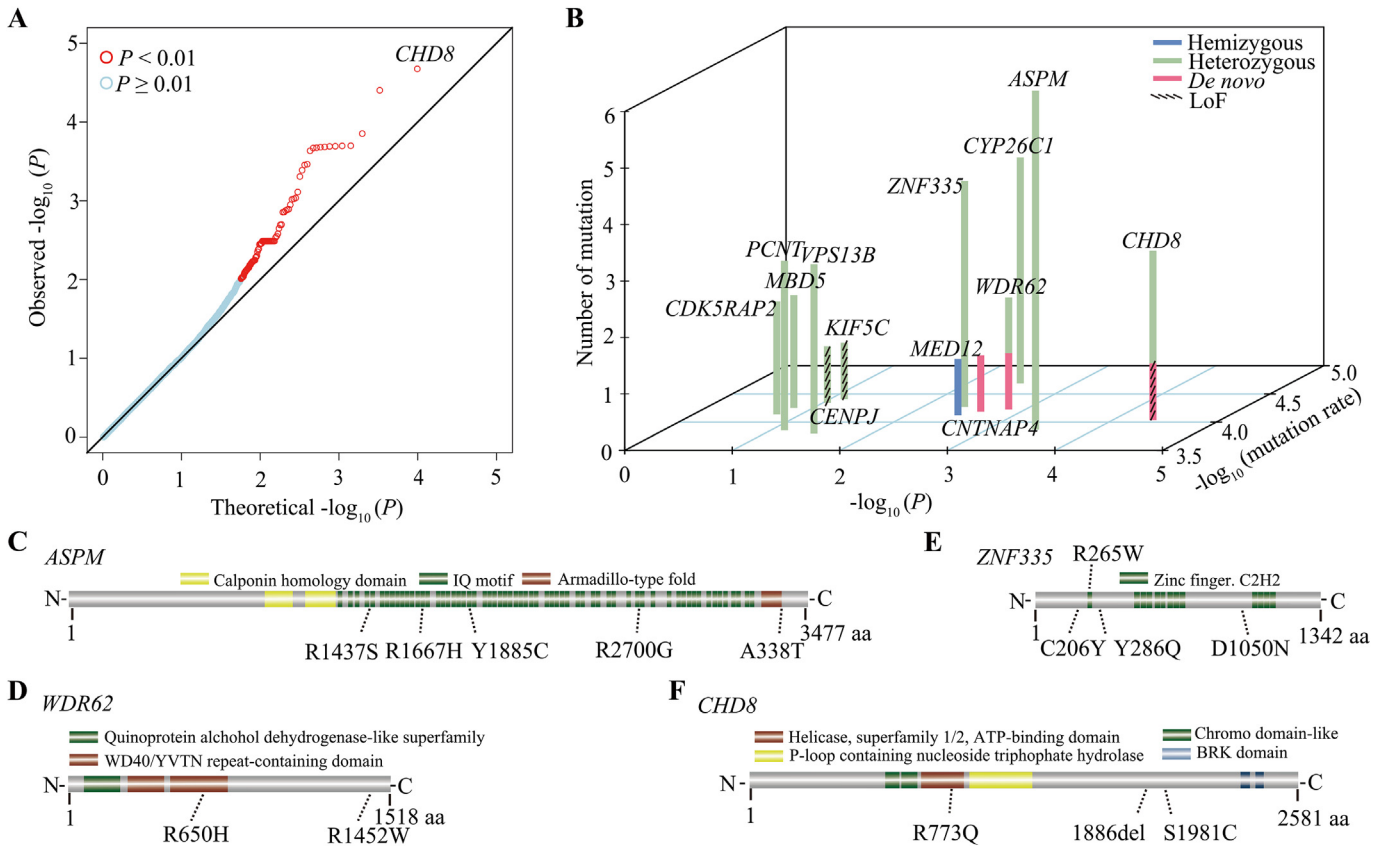

Fig. 2. Identification of ASD risk genes. **A**: Statistical significance of ASD risk genes. Q-Q plot of the *P*-values of the 4832 genes with rare mutations identified by WGS based on a combined analysis of loss-of-function (LoF) and deleterious missense mutations. One of the genes with an extremely low *P*-value (3.97×10^{-5}), located in the tail of the Q-Q plot, is *CHD8*, which is attributed to a *de novo* LoF and two inherited missense mutations. Besides the 12 LoF or DNMs on the top of the list, most of the tail in the Q-Q plot was driven by the remaining 75 genes harboring predicted damaging rare mutations. **B**: Three-dimensional view illustrating the probability of observing a specified number of mutations in genes with specified estimated mutations rates. Ranges of *P*-values and mutation rate are indicated for each of the genes on X- and Y-axis. **C**–**F**: Schematic diagrams of four ASD risk genes with recurrent mutations. Domain structures of each protein structure are predicted by the SMART program.

together, our results suggest that frequent heterozygous mutations of microcephaly-associated genes are likely to be associated with ASD.

Similar to the results in Caucasian patients (O'Roak et al., 2012; Bernier et al., 2014; Krumm et al., 2014), recurrent mutations of the macrocephaly-associated gene *CHD8* ($P_{TADA} = 0.0000381$, Fig. 2F and Table 1) were also detected in the Chinese cohort. In addition, we identified a novel hemizygous mutation in *MED12* ($P_{TADA} = 0.003263$), which plays an important role in chromatin binding and pathogenesis of macrocephaly (Ding et al., 2008). It is also worth to mention that a DNM (c.38A > G) in *CHD9* ($P_{TADA} = 0.003265$) was identified in proband A5, and is a possible candidate gene for ASD because of its interaction with a strong causative gene *CHD8*.

2.4. Identification of de novo CNVs and chromosomal rearrangements

We further explored potential CNVs and chromosomal rearrangements from these WGS data. Four *de novo* CNVs (ranging from 7.3-kb to 220-kb, Table S14) were confirmed by qPCR and Sanger sequencing (Table S14 and Fig. S2), including the *UBE3A*-containing 6.49 Mb *de novo* duplication at 15q11.2–13.1 ($P_{TADA} = 0.000204$, Fig. 3A and B). Different lengths of duplications of this region have been linked to ASD or other behavioral disorders (Fig. 3C) (Stewart et al., 2011; Abdelmoity et al., 2012; Christofolini et al., 2012; Urraca et al., 2013). Several genes in this region, such as *CYFIP1* and *MAGEL2*, are believed to be attributable to ASD. Growing evidences from clinical findings (Table S15) as well as animal models have indicated that *UBE3A*, which encodes an E3 ubiquitin ligase and regulates excitatory–inhibitory balance during brain development (Ebert and Greenberg, 2013), is a strong causative gene for Angelman syndrome including microcephaly and ID. In addition, we detected a *de novo* interstitial deletion of exon 8–10 (lost 109 amino acids in the Vinculin domain for microfilament attachment) in *CTNNA3* at 10q21.3 (Fig. S2). The encoded alpha-T-Catenin seemed to function as a negative regulator in the Wnt/β-catenin pathway (Busby et al., 2004), and was linked to autism and schizophrenia (Girirajan et al., 2013; Borglum et al., 2014).

Utilizing the pair-end read information and the CREST program, we discovered a *de novo* chromosomal translocation between chromosome 1 and 5. Our further analysis by Sanger sequencing confirmed the two breakpoints located in the first intron of *TNR* on 1q25.1 and the intergenic region between *KIF4B* and *SGCD* in 5q33.2, respectively (Fig. 4). Karyotype of this event [46, XY, t(1; 5)(q25.1; q33.2) *de novo*] was confirmed by G-banding karyotype analysis. The breakpoint in the first intron of *TNR* on 1q25.1 was expected to disrupt the transcription of *TNR* for the encoded tenascin-R protein. Clinically, this patient was diagnosed as ASD with significantly reduced learning ability (IQ score <40 in either verbal or performance test by WISC-R, compared to the IQ ≥ 120 from his parents). The encoded tenascin-R protein, a member of the tenascin family of extracellular matrix glycoproteins, is important in synaptic plasticity and affects learning, memory and cognition in

Table 1
Mutations of microcephaly- and macrocephaly-associated genes detected in patients with ASD.

Gene	Mutation	Proband	IQ	Cellular location	Function	Human (MIM#)
Microcephaly-associated genes						
ASPM	p.R1437S	A13	<40	Peri-centrosomal region (spindle pole or midbody) of mitotic, neural progenitor cells	Orientation of mitotic spindles during embryonic neurogenesis	MCPH5 (608716)
($P_{TADA} = 0.000353$)	p.R1667H	A18	54			
	p.Y1885C	A19	<40			
	p.R2700G	A32	41			
	p.R2700G	A3	<45			
	p.A3386T	A11	<40			
WDR62	p.R650H[#]	A29	<20	Spindle pole	Proliferation, neurogenesis	MCPH2 (604317)
($P_{TADA} = 0.001404$)	p.R1447W	A29	<20			
ZNF335	p.R265W	A10	<40	Component of histone methyltransferase complexes	Neural progenitor self-renewal, neurogenesis	MCPH10 (615095)
($P_{TADA} = 0.00408$)	p.C206Y	A12	<40			
	p.R286Q	A23	40			
	p.D1051N	A10	<40			
UBE3A	Duplication[*,#]	A22	61	E3 ubiquitin-protein ligase	Dendritic neurogenesis	AS (601623)
($P_{TADA} = 0.000204$)						
*VPS13*B	p.K2363R	A19	<40	Transmembrane protein in Golgi matrix	Vesicle-mediated transport	Cohen syndrome$^{-/-\text{ or }+/-}$ (216550)
($P_{TADA} = 0.034692$)	p.S3641T	A31	51			
	p.P3885S	A29	<20			
PCNT	p.E1185D	A22	61	Component of the pericentriolar material (PCM)	Centrosome splitting	MOPD$^{-/-\text{ or }+/-}$ (210720)
	p.R1939W	A15	94			
	p.K2393T	A17	<45			
CDK5RAP2	p.K126R	A1	<45	A centrosomal protein	Centriole replication	MCPH3 (604804)
	p.L1726F	A21	<20			
MBD5	p.I850M	A20	<45	Methyl-CpG-binding	Chromatin remodeling	ID, microcephaly, ASD$^{+/-}$ (156200)
	p.I497V	A31	61			
ATR	p.A1488V	A4	51	Phosphorylate checkpoint Chk1	DNA damage	SCKL1$^{-/-}$ (210600)
CENPJ	p.S407fs*	A22	61	Centromere protein, spindle	Centrosome integrity	MCPH6 (608393)
CEP63	p.R191Q	A17	<45	Component of centrosomes	Centriole duplication	SCKL6$^{-/-}$ (614728)
DYNC1H1	p.A187V	A32	41	ATPase	Axonal transport	MRD13$^{+/-}$ (614563)
KIF5C	c.1905+2-> G*	A10	<40	Microtubule motor and ATP binding	Dendritic trafficking	Cortical dysplasia$^{+/-}$ (615282)
KIF11	p.L517M	A13	<40	A motor protein in kinesin family	Centrosome separation	Microcephaly$^{+/-}$ (152950)
NIN	p.Q1164E	A3	<45	Centrosomal protein	Centrosome maturation	SCKL7 (614851)
TUBGCP6	p.R295S	A21	<20	γ-tubulin ring complex	Microtubule at centrosome	MCPHCR$^{-/-}$ (251270)
Macrocephaly-associated genes						
CHD8	p.L1887fs*[,#]	A14	45	DNA helicase binds beta-catenin, etc.	Chromatin remodeller	AUTS18$^{+/-}$ (615032)
($P_{TADA} = 0.000039$)	p.R773Q	A7	52			
	p.S1981C	A17	<45			
MED12	p.A50S	A10	<40	Chromatin binding mediator	Initiation of transcription	FGS1 (305450), X-linked
($P_{TADA} = 0.003263$)						

IQ, intelligence quotient; *, loss-of-function mutation; [#], *de novo* mutation; $^{-/-}$, homozygous; $^{+/-}$, autosomal dominant disorder or heterozygous mutant mice. 21 probands (65.6%) were detected with mutations in microcephaly- or macrocephaly-associated genes.

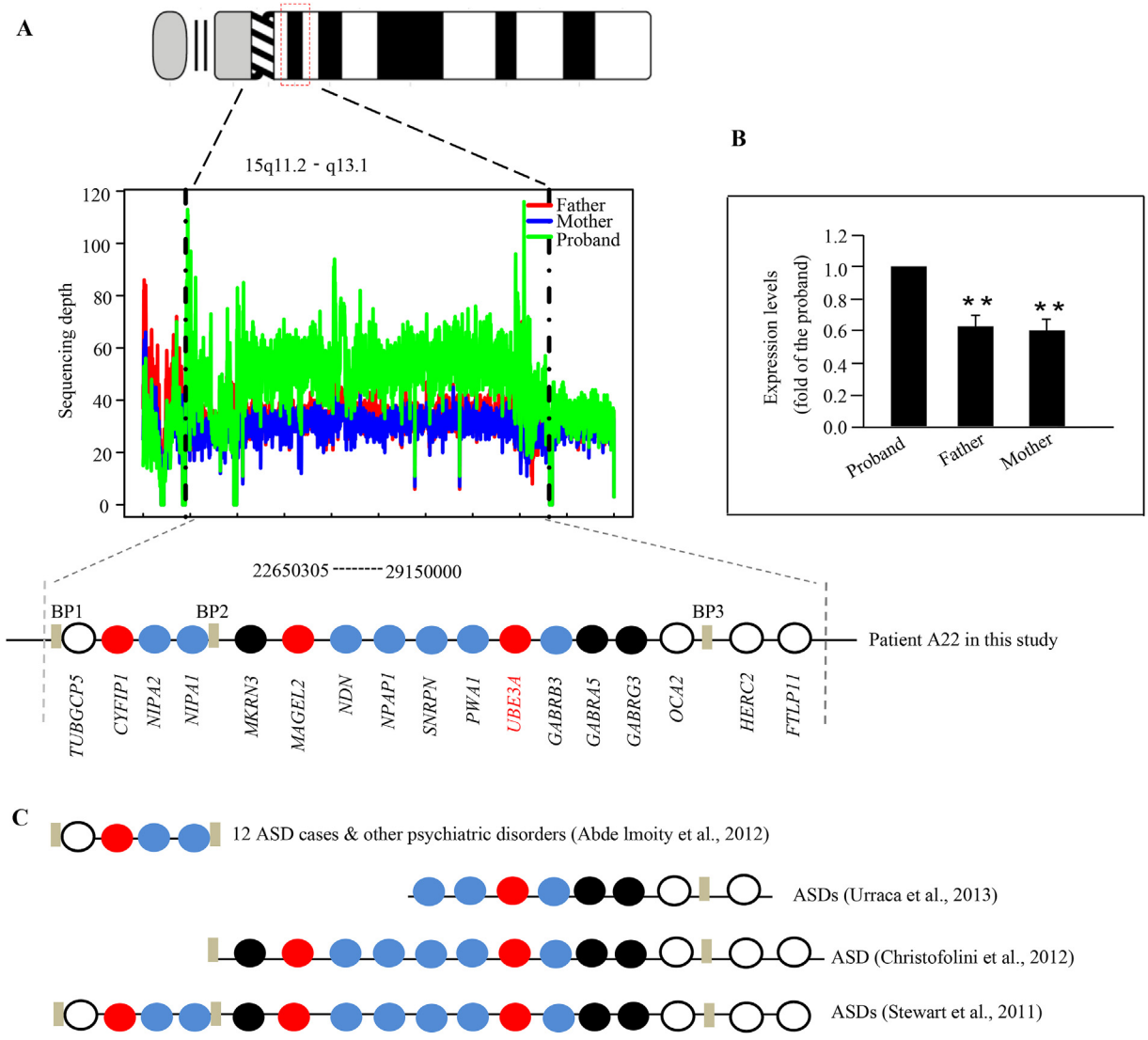

Fig. 3. A *de novo* CNV detected by WGS. **A**: An approximate 6.5 Mb duplication in the region of 15q11.2–13.1 detected in proband A22. The depths of sequencing reads are shown in different colors for the trio. 17 investigated genes are plotted along the region (encompassing 30 genes). Red ovals represent ASD-associated, blue ovals represent other psychiatric disorders, black ovals represent gene function in neuronal cells, and white ovals represent unrelated biological functions or disorders. The vertical dashed lines indicate boundaries. BP, breakpoint. **B**: qPCR results showing that expression level of duplicated genes like *UBE3A* in the proband is increased nearly 1.5-fold compared with his parents (**$P < 0.01$). **C**: Evidence from four representative reports linking different lengths of duplications to ASD or other behavioral disorders.

mice (Weber et al., 1999; Saghatelyan et al., 2004). More supporting evidences with respect of *TNR*'s role in ASD were found from recent microarray studies that identified a 4.97-Mb interstitial deletion (1q23.3–24.2) in a boy with mental retardation and autism (Della Monica et al., 2007) as well as a 394-kb homozygous deletion of *TNR* and the non-coding portions of *KIAA0040* in a girl with intellectual disability (Dufresne et al., 2012). This is the first identification of a *de novo* chromosomal translocation event in ASD using WGS.

3. Discussion

ASD is predominantly complex disorder with strong genetic heterogeneity (Rosti et al., 2014). Despite strong evidence supporting Mendelian inheritance in a small set of ASD, as a single-hit such as mutations in *FMRP* which usually accounts for <1% of cases (Steinberg and Webber, 2013), the exact inheritance remains elusive in majority of cases with ASD. In the absence of Mendelian inheritance patterns, ASD can be considered to be multiple-hit model, i.e., combination of multiple genetic risk factors with weak effect in each, such as in the cases with *SHANK2* mutation as well as CNV at 15q11-q13 (Leblond et al., 2012). In this study, we found that 25 probands in this study harbored multiple damaging mutations involving different loci in coding regions as well as potential genetic underpinnings in non-coding regions, such as UTRs or promoters (Table S16). For example, proband A11 was found to carry an inherited damaging mutation in *ASPM* and also a *de novo* loss-of-function mutation in *NRXN2*. Moreover, proband A14 carried a *de novo* loss-of-function mutation in *CHD8* and also several mutations in other loci, including a rare inherited damaging mutation in *STAB1*, a homozygous mutation in *GDPD4*, and a DNM in *KRT7*. Therefore, it is possible that genetic interaction in the multiple-hit model may lead to an aggravating genetic epistasis. Although these concurrent mutations have been prioritized for autism or other psychiatric disorders, understanding of the interactions will be critical for dissecting the complex inheritance pattern of ASD. Thus, precise genetic interactions should be pursued in the future.

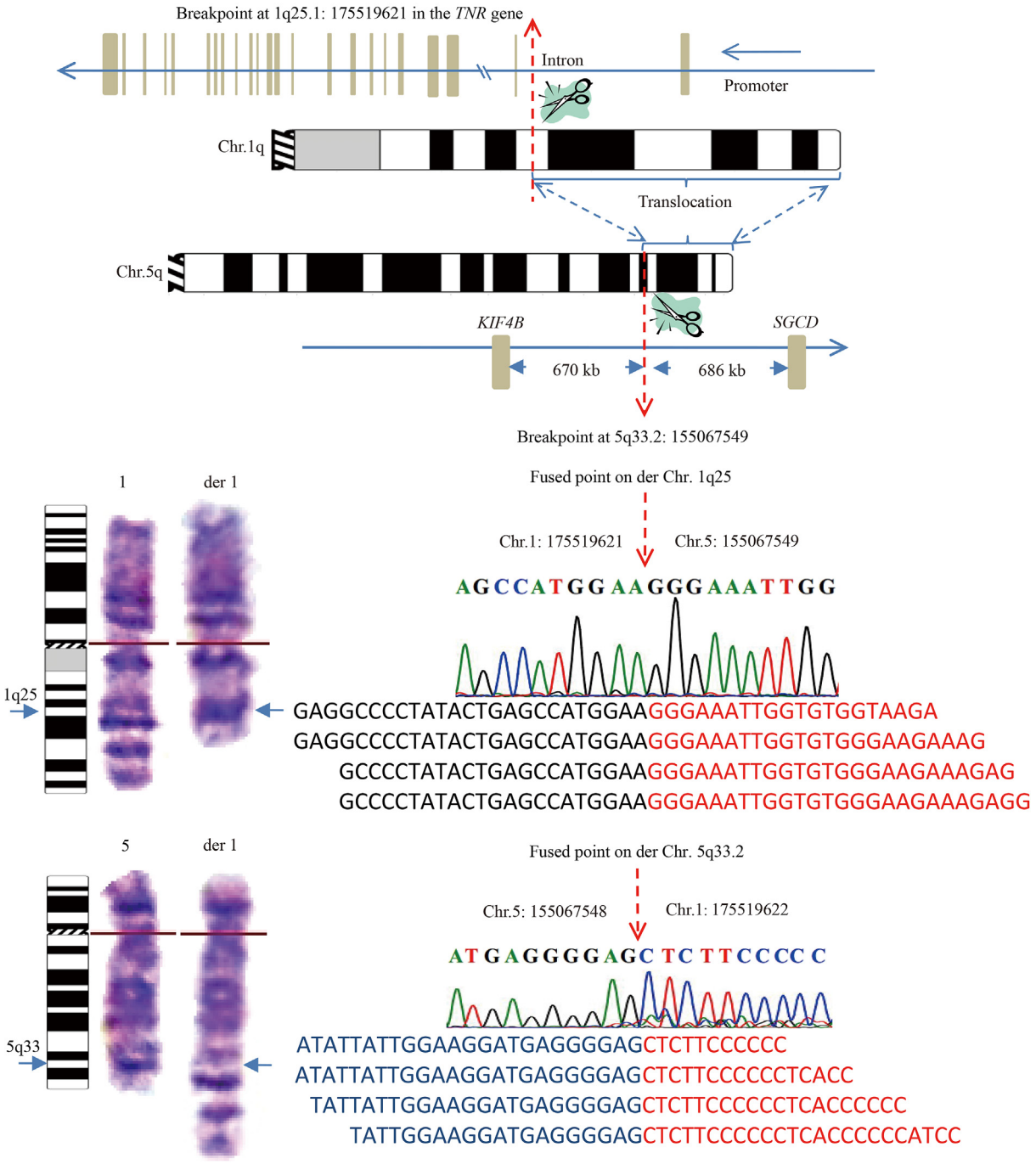

Fig. 4. Identification of a *de novo* chromosomal translocation in proband A12. The breakpoints are found in the intron 1 of *TNR* on 1q25.1 and the intergenic region between *KIF4B* and *SGCD* on 5q33.2, respectively. The vertical dashed lines in red indicate breakpoints. The gray vertical bars crossing with horizontal lines in red represent exons. The arrow points to the transcriptional direction. Sanger sequencing confirms the breakpoints on Chr.1:175519621 and Chr.5:155067548, respectively. Representative Illumina reads around breakpoints are also shown. Chromosomal karyotype confirmation of the *de novo* balance translocation of 46, XY, t(1; 5)(q25.1; q33.2).

Compared to WES, WGS can detect *de novo* CNVs or SVs more efficiently. Moreover, WGS can provide more uniform coverage over coding regions, a higher sensitivity and more protein-affecting variant detection than WES (Wilfert et al., 2017; Jin et al., 2018). In the present study, except for *de novo* CNVs and *de novo* SVs which were well-validated, we also successfully detected rare inherited mutations of several microcephaly-associated genes and potential genetic underpinnings in non-coding regions. With the gradual reduction in costs, WGS will be used widely as the main approach for ASD diagnosis or disease-causing gene discovery.

Despite substantial genetic heterogeneity in ASD, using ingenuity pathway analysis (IPA), investigation of identified microcephaly- and macrocephaly-associated genes revealed a network that would affect neuronal migration and neurogenesis as observed in MCPH disorders and related syndromes. The corresponding biological convergence of three modules is shown in Fig. 5: 1) ASPM, WDR62, VPS13B, CNTANP4, MBD5, NIN and CYP26C1 are implicated in Akt as well as Wnt pathways and are degraded via the

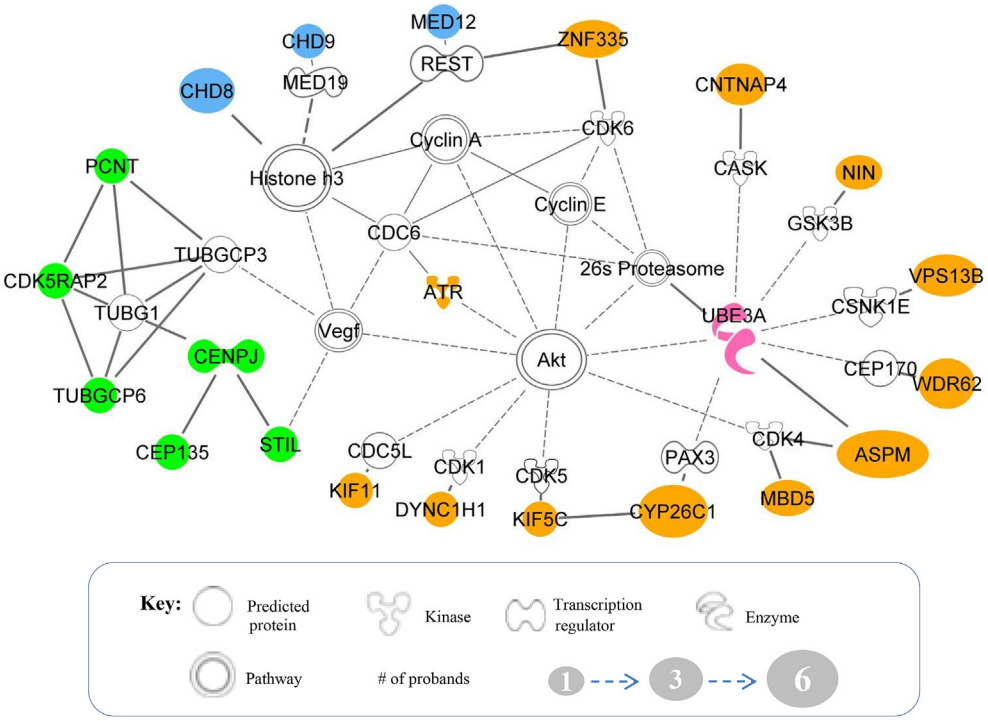

Fig. 5. Network analysis of genes implicated in ASD. IPA analysis was used to explore the relationship of genes harboring various mutations identified in this study (nodes in color), and intermediate proteins (nodes in white color). More than 20 predicted proteins corresponding to mutated genes are interconnected directly (solid line) or indirectly (dot line) according to IPA and IntAct database. Nodes in orange, green or red represent corresponding genes with mutations resulting in microcephaly, while nodes in blue denote mutated genes causing macrocephaly. These colored seed genes in the protein-protein interaction network were found to have a significantly higher probability in the convergence of ASD pathways than genes not in these networks.

UBE3A-proteasome system. In addition, KIF11, KIF5C and DYNC1H1 are also involved in the Akt pathway; 2) PCNT, CDK5RRAP2, TUBGCP6, CEP72, CENPJ, and STIL are related to microtubule nucleation at the centrosome; 3) CHD8, CHD9, and MED12 are involved in chromatin remodeling and histone H3 as well as cell cycle regulation. It is worth noting that ZNF335 was reported to interact with REST-histone H3 as well as CDK6-cell cycle proteins and cause microcephaly (MCPH10), but not macrocephaly (Yang et al., 2012). Thus, our result implied that a highly connected network associated with microcephaly through dysfunctions of centrosomal function and chromatin remodeling likely play a role in the pathogenesis of ASD.

It has been reported that ID is seen in 50%–80% of individuals with ASD, and conversely ASD is also seen in 40% of ID patients (Matson and Shoemaker, 2009). The overlapped phenotypes between ASD and ID suggest that there would be a common pathway(s) for two disorders. As microcephaly genes have been linked with ID, by combining with our finding of their roles in autism, we speculate that the genes involved in microcephaly may contribute to the pathogenesis of certain subtype of ASD, such as ASD with ID, in the manner of a single gene mutation like *UBE3A* or with multiple hits like the loss-of-function mutation of *NRXN2* and missense mutations of *ASPM*. It is expected that more related genes with mutations and polymorphisms will be identified from the implementation of WGS and functional analysis of network convergence.

Taken together, we performed comprehensive analysis of WGS data from a cohort of Chinese ASD trios for better understanding of their genetic architecture. Our results demonstrated that various DNMs occurred more frequently in coding regions with statistical significance. The co-occurrence of DNMs and inherited harmful mutations or combinations of any types of pathogenic structural alterations may play a role in the genetic susceptibility to ASD. Remarkably, for the first time, our findings revealed that transmitted rare mutations of microcephaly-associated genes are implicated in the pathogenesis of ASD, which may be correlated to certain subtype of ASD, such as ID. Given by the fact that most patients in this study were found to carry multiple rare DNMs and/or transmitted variants, despite the limited sample size, it is suggestive that ASD is in favor to multiple-hit. These findings in turn suggest that adoption of WGS as a new yet efficient technique to illustrate the full genetic spectrum in complex disorders, such as ASD, could provide novel insights into pathogenesis, diagnosis, and treatment.

4. Materials and methods

4.1. Patients and samples

This study included 32 unrelated trios (30 males and 2 females) of Han Chinese ancestry with ASD and the average age of probands was 6-year-old ranging from 3 to 13-year-old. The human subjects were recruited at the Children's Hospital of Fudan University, Shanghai, China and the Second Xiangya Hospital of Central South University, Changsha, China by following the protocol approved by the Ethics Committee of each hospital stated above. Written informed consents were signed by the parents and on behalf of their children before peripheral blood and clinical data were collected from ASD patients and their parents. Diagnosis was undertaken by a team of autism specialists based upon comprehensive analyses of DSM-IV-TR (Diagnostic and Statistical Manual of Mental Disorders-4th edition) criteria and Autism Diagnostic Observation Schedule (ADOS). In addition, clinical evaluation was performed to exclude other known medical disorders, such as tuberous sclerosis, congenital rubella syndrome and

neurofibromatosis.

4.2. Intelligence assessment

Psychometric intelligence was assessed using Wechsler Intelligence Scale for Children 4th edition-Chinese version (WISC-R) as described previously (Feng et al., 2011). Only the full scale of intelligence quotient (FSIQ) is shown if verbal intelligence quotient (VIQ) equals to the performance intelligence quotient (PIQ). All intelligence evaluations for the subjects were assessed by a single investigator who was blinded to the subjects' clinical information.

4.3. WGS and variant detection

Genomic DNAs (gDNAs) from peripheral blood of 32 trios were extracted using Genomic DNA Extraction Kit (Invitrogen, Grand Island, NY, USA). Genomic DNA libraries were prepared and sequenced according to Illumina paired-end sequencing protocols. WGS data (90 pair-end reads, >30 ×) were generated from gDNA libraries (500-bp, Illumina library preparation protocol) by Illumina Hiseq2000 sequencing system. Alignment of the reads (hg19, GRCh37) was performed using Burrows-Wheeler Aligner (BWA) (Li and Durbin, 2009) with the following parameters: "aln -o 1 -e 40 −i 15 -L -l 31 -k 2 -t 12". Duplicated reads were removed using Sequence Alignment/Map tools (SAMtools) (Li et al., 2009) and only uniquely mapping reads were used for variation detection. After local realignment and quality recalibration, SNVs and small InDels were identified using GATK (McKenna et al., 2010) for all trios, in which the mapping results of three members were used simultaneously as inputs. Annotation of the variants, such as locations (exonic, intronic and intergenic region, etc.) and effects on protein coding (synonymous, missense, nonsense, frameshift, etc.), was performed by ANNOVAR (Wang et al., 2010) with RefSeq (hg19, from UCSC) RefGene annotation.

Since the variants initially detected by BWA and GATK harbored considerable false variants, mirTrios and ForestDNM methods were used to remove false positives and refine the *de novo* SNV calls in all trios (Michaelson et al., 2012; Jiang et al., 2013; Li et al., 2015). Furthermore, an in-house-developed method was used to evaluate potential false negatives missed by ForestDNM method in the initial *de novo* SNV list and *de novo* InDels. The average rate of *de novo* mutation (M) across different genomic features was calculated as follows: $M = N/(P \times G \times L)$, where N = number of *de novo* SNVs (InDels) in a given genomic feature, L = region size of genomic feature (hg19), P = number of probands (32), G = number of generations (2).

SegSeq (Chiang et al., 2009) and CNVnator (Abyzov et al., 2011) were used to detect *de novo* CNVs. Firstly, SegSeq detected *de novo* CNVs by comparing proband with each of parents under the same filtering criterion separately (copy ratio ≥ 1.5 or ≤ 0.5, sequence depth ranging from 0.1-fold to 10-fold of average depth). Then CNVs in the two sets with same type variant and high coverage (overlap length/min $\geq$0.5) would be combined as candidate *de novo* CNVs. CNVnator detected CNVs in all family members independently. Proband-specific CNVs that did not meet the following criteria would be discarded: 1) q0 (the fraction of reads with zero mapping quality in the called CNV regions) < 50%; 2) corrected *P*-value < 0.05; 3) overlap rate of repeat database (STR, DGV, genomic SuperDups) < 50%. If the normalized RD (read depth) signal is < 1.4 for deletions or >2.6 for duplicates in proband and >1.6 for deletions or <2.4 for duplicates in both parents, this CNV would be defined as *de novo* CNV.

The CREST (Wang et al., 2011) program was used to detect the SVs in a proband. In CREST, the soft-clipping signatures and assembly-mapping-searching-assembly-alignment were used to identify the first breakpoint and the second breakpoint, respectively. We retained a candidate SV with scores $\geq$30 and number of supporting read pairs $\geq$3. Since CREST only allows two input files at the same time, we generated two sets of SVs which were not present in father and mother, respectively. Then, we intersected the two sets of SVs and obtained the SVs neither present in father nor in mother simultaneously, which generated the *de novo* SVs in the proband. Lastly, the *de novo* SVs identified were refined manually by an html view of the multiple alignment for the breakpoint built by visualization module in the CREST, which must meet the following criteria: 1) the *de novo* SVs are supported by $\geq$ 3 soft-clipping reads and absent in their parents; 2) the according loci of the breakpoint in parents have $\geq$2 reads.

To identify homozygous, compound heterozygous or hemizygous rare inherited mutation likely contributing to ASD, we only selected mutations in coding region with further exclusion of any mutations in dbSNP138, or with MAF >0.1% in 1000 Genomes Project (released in October 21, 2013), NHLBI-6500 Exomes, and CG69 database. Like DNMs in coding region, only loss-of-function mutation and pathogenic non-synonymous mutations were further analyzed. Pathogenic non-synonymous mutations were predicted by SIFT, PolyPhen2, MutationTaster and GERP++ and must satisfy at least two of the four criteria. For homozygous mutations, they must be inherited from both father and mother and their parent must be heterozygous. For compound heterozygous mutations, a gene in a proband has at least two rare inherited mutations. One mutation is inherited from the father and the other from the mother, which means two chromosomes contain at least one deleterious mutation. For hemizygous mutations on X chromosome in male, the mutation must be inherited from his mother who must be heterozygous.

4.4. Parent-of-origin assignment for the variants

Parent-of-origin for *de novo* mutation was determined by informative inherited variant sites, based on mapped reads encompassing *de novo* mutation site. The informative inherited variant sites were determined by two possible scenarios: 1) one parent is homozygous reference alleles, while the other parent is homozygous alternate alleles, and child is heterozygous; 2) one parent is homozygous reference or the other is heterozygous, and child is heterozygous.

The *de novo* sites where the informative inherited sites within region were twice the library size at both upstream and downstream of *de novo* position were included firstly. We then extracted reads within the interval defined by read length both upstream and downstream of *de novo* position. The reads which meet the following criteria were used for inferring the parent-of-origin: 1) the reads must contain the *de novo* site; 2) either the read or its mate-pair must contain at least one informative inherited site; 3) the base quality of the *de novo* site and inherited site on the candidate read or its mate-pair should be above the cut-off limit. The parent who contributes the base at inherited site is not the parent-of-origin for the *de novo* variant if the base at inherited site is reference genotype. This step will be repeated for all extracted reads or its mate-pair. Higher number of such reads will be translated to higher confidence level for the parent-of-origin information. This process was replicated for the remaining *de novo* sites.

4.5. Correlation analysis of parents' age and DNMs

To explore whether any correlation exists between DNMs and paternal age or maternal age, correlation analysis was conducted using R (http://www.r-project.org/). Firstly, we calculated the Pearson's correlation between DNM number and father's age.

Secondly, since the mother's age is substantially correlated with father's age ($r = 0.63$), a multivariate regression analysis was performed considering the father's age and mother's age together. Furthermore, the parents were divided into three groups based on their age, and odds ratio (OR) was calculated for each group of father and mother, respectively.

4.6. Exploration of genome-wide mutation patterns of de novo SNVs

For each proband, we calculated the number of *de novo* SNVs and InDels, as well as the mutation spectrum. We classified mutations into 6 classes: C > T/G > A, C > A/G > T, C > G/G > C, T > C/A > G, T > G/A > C, and T > A/A > T (equivalent changes were combined, for example, T > C and A > G). To explore mutation patterns of C > T/G > A and T > C/A > G classes, we compared the base 5′ end of C > T, C > A, and C > G mutations with the base after cytosine in reference genome (hg19) which is used as background. It is noted that G > A has counted by complementary strand same to C > T, the same as other class of mutation. In the same way, we also compared the base 3′ end of T > C, T > G, and T > A mutations with the base after thymine in reference genome. To execute binomial test using R (http://www.r-project.org/), we fist calculated the expected probability of each pattern (such as, CA/CT/CG/CC) by the whole reference genome hg19. Then, we counted the number of observed patterns in different classes of mutations.

4.7. Prioritization of ASD candidate genes

To identify risk-conferring variants to ASD, we firstly removed any variants with MAF >0.1%, according to dbSNP138, 1000-Genomes (released in October 21, 2013) (Abecasis et al., 2012), ESP6500 (Fu et al., 2013), CG69database (Drmanac et al., 2010) and 1300 in-house Chinese exome database at BGI (obtained from SureSelect Human All Exon V4 with larger than 50× sequencing depth). For mutations in coding region, loss-of-function mutations (such as frameshift, stopgain and splicing) were considered to contribute to ASD risk directly (Stein et al., 2013). For non-synonymous SNVs, the selected pathogenic mutations and deleterious non-synonymous SNVs must satisfy at least two of the following criteria: 1) SIFT scores <0.05 predicted as “deleterious” (Kumar et al., 2009); 2) PolyPhen scores> 0.15 predicted as “probably damaging” or “possibly damaging” (Adzhubei et al., 2010); 3) MutationTaster score >0.5 predicted as “disease causing” (Schwarz et al., 2010); 4) GERP++ score >2 predicted as “rejected substitutions” (Davydov et al., 2010). Then the TADA program (He et al., 2013) was used to calculated *P* value of each gene harboring rare deleterious mutation. The ASD-risk genes were submitted to Ingenuity Pathways Analysis (IPA, http://www.ingenuity.com/) to identify neurological disease or function-associated genes and pathways with statistically significant scores.

4.8. Genetic variation confirmation

SNVs and InDels were validated by PCR-Sanger sequencing in regions flanking given mutations. PCR-Sanger sequencing across the breakpoints was also performed to confirm identified *de novo* SVs. In addition, we applied TA Clone, coupled with Sanger sequencing, to verify the *de novo* SNVs of the *NRXN2* mutation (c.809_810insG, p.L1887fs) in proband A11.

Real-time PCR was used to verify the *de novo* CNVs. Briefly, the concentration of each DNA sample was diluted to 10 ng/μL with ddH_2O. Real-time PCR analysis was performed on an ABI 7300 Real time PCR system using the SYBR Green probe (TaKaRa, Dalian, China). We used total 20 μL reaction mixture containing 10 μL SYBR Green reagent and 1 μL DNA sample. Technical triplication was performed for each DNA sample. *β-actin* was used as the control with primer sequences: 5′-ccaaccgcgagaagatga-3′ and 5′-ccagaggcgtacagggatag-3′. Three pairs of primer, flanking at the start, middle, and end position of the CNV sequence were respectively chosen by using GenScript Real-time PCR Primer Design Tool (https://www.genscript.com/ssl-bin/app/primer). PCR conditions were as follows: 94 °C 15 s, 60 °C 30 s, 72 °C 30 s, 40 cycles. The relative quantitation was calculated according to the $2^{-\Delta\Delta Ct}$ formula.

Karyotype analysis was used to validate the *de novo* SV. We made air-dry preparations by dropping small droplets of blood cell suspension on a glass slide. After incubating the slide in 2 × SSC at 60–65 °C for 1 h, the slide was transferred to 0.9% NaCl solution at room temperature and rinsed in fresh NaCl solution. After staining for 4–6 min in Trypsin-Giemsa solution, the slide was transferred to fresh 1× PBS buffer and followed by washing with 1 × PBS buffer twice. After shaking off excess liquid and blowing dry with air, G-banding karyotype analysis was carried out with VideoTesT-Karyo 3.1 software (Digital Imaging Systems, Buckinghamshire, UK).

Acknowledgments

We thank the patients and their families for their participation in this study. This work was supported by the grants from the Major State Basic Research Development Program of China (2012CB517902 and 2012CB517904), National Key Technology Research and Development Program of China (2012BAI03B00), Special Research Program of National Health and Family Planning Commission of China (201302002), International S&T Cooperation Program of China (2011DFA30670) and National Natural Science Foundation of China (31571357/31771404). This work was also supported in part by research funding from AstraZeneca Innovation Center China and Wenzhou Medical University.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jgg.2018.09.002.

References

Abdelmoity, A.T., LePichon, J.B., Nyp, S.S., Soden, S.E., Daniel, C.A., Yu, S., 2012. 15q11.2 proximal imbalances associated with a diverse array of neuropsychiatric disorders and mild dysmorphic features. J. Dev. Behav. Pediatr. 33, 570–576.

Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., McVean, G.A., 2012. An integrated map of genetic variation from 1,092 human genomes. Nature 491, 56–65.

Abyzov, A., Urban, A.E., Snyder, M., Gerstein, M., 2011. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. Genome Res. 21, 974–984.

Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., Sunyaev, S.R., 2010. A method and server for predicting damaging missense mutations. Nat. Methods 7, 248–249.

Benadiba, C., Magnani, D., Niquille, M., Morle, L., Valloton, D., Nawabi, H., Ait-Lounis, A., Otsmane, B., Reith, W., Theil, T., Hornung, J.P., Lebrand, C., Durand, B., 2012. The ciliogenic transcription factor RFX3 regulates early midline distribution of guidepost neurons required for corpus callosum development. PLoS Genet. 8 e1002606.

Bernier, R., Golzio, C., Xiong, B., Stessman, H.A., Coe, B.P., Penn, O., Witherspoon, K., Gerdts, J., Baker, C., Vulto-van Silfhout, A.T., Schuurs-Hoeijmakers, J.H., Fichera, M., Bosco, P., Buono, S., Alberti, A., Failla, P., Peeters, H., Steyaert, J., Vissers, Lelm, Francescatto, L., Mefford, H.C., Rosenfeld, J.A., Bakken, T., O'Roak, B.J., Pawlus, M., Moon, R., Shendure, J., Amaral, D.G., Lein, E., Rankin, J., Romano, C., de Vries, B.B.A., Katsanis, N., Eichler, E.E., 2014. Disruptive *CHD8* mutations define a subtype of autism early in development. Cell 158, 263–276.

Bibonne, A., Neant, I., Batut, J., Leclerc, C., Moreau, M., Gilbert, T., 2013. Three calcium-sensitive genes, *fus, brd3* and *wdr5*, are highly expressed in neural and renal territories during amphibian development. Biochim. Biophys. Acta 1833, 1665–1671.

Bilguvar, K., Ozturk, A.K., Louvi, A., Kwan, K.Y., Choi, M., Tatli, B., Yalnizoglu, D.,

Tuysuz, B., Caglayan, A.O., Gokben, S., Kaymakcalan, H., Barak, T., Bakircioglu, M., Yasuno, K., Ho, W., Sanders, S., Zhu, Y., Yilmaz, S., Dincer, A., Johnson, M.H., Bronen, R.A., Kocer, N., Per, H., Mane, S., Pamir, M.N., Yalcinkaya, C., Kumandas, S., Topcu, M., Ozmen, M., Sestan, N., Lifton, R.P., State, M.W., Gunel, M., 2010. Whole-exome sequencing identifies recessive *WDR62* mutations in severe brain malformations. Nature 467, 207–210.

Borglum, A.D., Demontis, D., Grove, J., Pallesen, J., Hollegaard, M.V., Pedersen, C.B., Hedemand, A., Mattheisen, M., Uitterlinden, A., Nyegaard, M., Orntoft, T., Wiuf, C., Didriksen, M., Nordentoft, M., Nothen, M.M., Rietschel, M., Ophoff, R.A., Cichon, S., Yolken, R.H., Hougaard, D.M., Mortensen, P.B., Mors, O., 2014. Genome-wide study of association and interaction with maternal cytomegalovirus infection suggests new schizophrenia loci. Mol. Psychiatr. 19, 325–333.

Brandler, W.M., Antaki, D., Gujral, M., Noor, A., Rosanio, G., Chapman, T.R., Barrera, D.J., Lin, G.N., Malhotra, D., Watts, A.C., Wong, L.C., Estabillo, J.A., Gadomski, T.E., Hong, O., Fajardo, K.V., Bhandari, A., Owen, R., Baughn, M., Yuan, J., Solomon, T., Moyzis, A.G., Maile, M.S., Sanders, S.J., Reiner, G.E., Vaux, K.K., Strom, C.M., Zhang, K., Muotri, A.R., Akshoomoff, N., Leal, S.M., Pierce, K., Courchesne, E., Iakoucheva, L.M., Corsello, C., Sebat, J., 2016. Frequency and complexity of *de novo* structural mutation in autism. Am. J. Hum. Genet. 98, 667–679.

Busby, V., Goossens, S., Nowotny, P., Hamilton, G., Smemo, S., Harold, D., Turic, D., Jehu, L., Myers, A., Womick, M., Woo, D., Compton, D., Doil, L.M., Tacey, K.M., Lau, K.F., Al-Saraj, S., Killick, R., Pickering-Brown, S., Moore, P., Hollingworth, P., Archer, N., Foy, C., Walter, S., Lendon, C., Iwatsubo, T., Morris, J.C., Norton, J., Mann, D., Janssens, B., Hardy, J., O'Donovan, M., Jones, L., Williams, J., Holmans, P., Owen, M.J., Grupe, A., Powell, J., van Hengel, J., Goate, A., Van Roy, F., Lovestone, S., 2004. Alpha-T-catenin is expressed in human brain and interacts with the Wnt signaling pathway but is not responsible for linkage to chromosome 10 in Alzheimer's disease. NeuroMolecular Med. 5, 133–146.

Chiang, D.Y., Getz, G., Jaffe, D.B., O'Kelly, M.J., Zhao, X., Carter, S.L., Russ, C., Nusbaum, C., Meyerson, M., Lander, E.S., 2009. High-resolution mapping of copy-number alterations with massively parallel sequencing. Nat. Methods 6, 99–103.

Christofolini, D.M., Meloni, V.A., Ramos, M.A., Oliveira, M.M., de Mello, C.B., Pellegrino, R., Takeno, S.S., Melaragno, M.I., 2012. Autistic disorder phenotype associated to a complex 15q intrachromosomal rearrangement. Am. J. Med. Genet. B Neuropsychiatr. Genet. 159B, 823–828.

Davydov, E.V., Goode, D.L., Sirota, M., Cooper, G.M., Sidow, A., Batzoglou, S., 2010. Identifying a high fraction of the human genome to be under selective constraint using GERP plus. PLoS Comput. Biol. 6 e1001025.

De Rubeis, S., He, X., Goldberg, A.P., Poultney, C.S., Samocha, K., Cicek, A.E., Kou, Y., Liu, L., Fromer, M., Walker, S., Singh, T., Klei, L., Kosmicki, J., Shih-Chen, F., Aleksic, B., Biscaldi, M., Bolton, P.F., Brownfeld, J.M., Cai, J., Campbell, N.G., Carracedo, A., Chahrour, M.H., Chiocchetti, A.G., Coon, H., Crawford, E.L., Curran, S.R., Dawson, G., Duketis, E., Fernandez, B.A., Gallagher, L., Geller, E., Guter, S.J., Hill, R.S., Ionita-Laza, J., Jimenz Gonzalez, P., Kilpinen, H., Klauck, S.M., Kolevzon, A., Lee, I., Lei, I., Lei, J., Lehtimaki, T., Lin, C.F., Ma'ayan, A., Marshall, C.R., McInnes, A.L., Neale, B., Owen, M.J., Ozaki, N., Parellada, M., Parr, J.R., Purcell, S., Puura, K., Rajagopalan, D., Rehnstrom, K., Reichenberg, A., Sabo, A., Sachse, M., Sanders, S.J., Schafer, C., Schulte-Ruther, M., Skuse, D., Stevens, C., Szatmari, P., Tammimies, K., Valladares, O., Voran, A., Li-San, W., Weiss, L.A., Willsey, A.J., Yu, T.W., Yuen, R.K., Study, D.D.D., Autism Homozygosity Mapping Collaborative for, Uk K, Consortium, Cook, E.H., Freitag, C.M., Gill, M., Hultman, C.M., Lehner, T., Palotie, A., Schellenberg, G.D., Sklar, P., State, M.W., Sutcliffe, J.S., Walsh, C.A., Scherer, S.W., Zwick, M.E., Barett, J.C., Cutler, D.J., Roeder, K., Devlin, B., Daly, M.J., Buxbaum, J.D., 2014. Synaptic, transcriptional and chromatin genes disrupted in autism. Nature 515, 209–215.

Della Monica, M., Lonardo, F., Faravelli, F., Pierluigi, M., Luquetti, D.V., De Gregori, M., Zuffardi, O., Scarano, G., 2007. A case of autism with an interstitial 1q deletion (1q23.3-24.2) and a *de novo* translocation of chromosomes 1q and 5q. Am. J. Med. Genet. 143A, 2733–2737.

Ding, N., Zhou, H., Esteve, P.O., Chin, H.G., Kim, S., Xu, X., Joseph, S.M., Friez, M.J., Schwartz, C.E., Pradhan, S., Boyer, T.G., 2008. Mediator links epigenetic silencing of neuronal gene expression with x-linked mental retardation. Mol. Cell 31, 347–359.

Drmanac, R., Sparks, A.B., Callow, M.J., Halpern, A.L., Burns, N.L., Kermani, B.G., Carnevali, P., Nazarenko, I., Nilsen, G.B., Yeung, G., Dahl, F., Fernandez, A., Staker, B., Pant, K.P., Baccash, J., Borcherding, A.P., Brownley, A., Cedeno, R., Chen, L.S., Chernikoff, D., Cheung, A., Chirita, R., Curson, B., Ebert, J.C., Hacker, C.R., Hartlage, R., Hauser, B., Huang, S., Jiang, Y., Karpinchyk, V., Koenig, M., Kong, C., Landers, T., Le, C., Liu, J., McBride, C.E., Morenzoni, M., Morey, R.E., Mutch, K., Perazich, H., Perry, K., Peters, B.A., Peterson, J., Pethiyagoda, C.L., Pothuraju, K., Richter, C., Rosenbaum, A.M., Roy, S., Shafto, J., Sharanhovich, U., Shannon, K.W., Sheppy, C.G., Sun, M., Thakuria, J.V., Tran, A., Vu, D., Zaranek, A.W., Wu, X.D., Drmanac, S., Oliphant, A.R., Banyai, W.C., Martin, B., Ballinger, D.G., Church, G.M., Reid, C.A., 2010. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. Science 327, 78–81.

Dufresne, D., Hamdan, F.F., Rosenfeld, J.A., Torchia, B., Rosenblatt, B., Michaud, J.L., Srour, M., 2012. Homozygous deletion of *Tenascin-R* in a patient with intellectual disability. J. Med. Genet. 49, 451–454.

Ebert, D.H., Greenberg, M.E., 2013. Activity-dependent neuronal signalling and autism spectrum disorder. Nature 493, 327–337.

Feng, J.J., Xu, X., Wang, W.P., Guo, S.J., Yang, H., 2011. Pattern visual evoked potential performance in preterm preschoolers with average intelligence quotients. Early Hum. Dev. 87, 61–66.

Fu, W.Q., O'Connor, T.D., Jun, G., Kang, H.M., Abecasis, G., Leal, S.M., Gabriel, S., Altshuler, D., Shendure, J., Nickerson, D.A., Bamshad, M.J., Akey, J.M., NHLBI Exome Sequencing Project, 2013. Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. Nature 493, 216–220.

Fujimori, A., Itoh, K., Goto, S., Hirakawa, H., Wang, B., Kokubo, T., Kito, S., Tsukamoto, S., Fushiki, S., 2014. Disruption of Aspm causes microcephaly with abnormal neuronal differentiation. Brain Dev. 36, 661–669.

Girirajan, S., Dennis, M.Y., Baker, C., Malig, M., Coe, B.P., Campbell, C.D., Mark, K., Vu, T.H., Alkan, C., Cheng, Z., Biesecker, L.G., Bernier, R., Eichler, E.E., 2013. Refinement and discovery of new hotspots of copy-number variation associated with autism spectrum disorder. Am. J. Hum. Genet. 92, 221–237.

He, X., Sanders, S.J., Liu, L., De Rubeis, S., Lim, E.T., Sutcliffe, J.S., Schellenberg, G.D., Gibbs, R.A., Daly, M.J., Buxbaum, J.D., State, M.W., Devlin, B., Roeder, K., 2013. Integrated model of *de novo* and inherited genetic variants yields greater power to identify risk genes. PLoS Genet. 9 e1003671.

Iossifov, I., O'Roak, B.J., Sanders, S.J., Ronemus, M., Krumm, N., Levy, D., Stessman, H.A., Witherspoon, K.T., Vives, L., Patterson, K.E., Smith, J.D., Paeper, B., Nickerson, D.A., Dea, J., Dong, S., Gonzalez, L.E., Mandell, J.D., Mane, S.M., Murtha, M.T., Sullivan, C.A., Walker, M.F., Waqar, Z., Wei, L., Willsey, A.J., Yamrom, B., Lee, Y.H., Grabowska, E., Dalkic, E., Wang, Z., Marks, S., Andrews, P., Leotta, A., Kendall, J., Hakker, I., Rosenbaum, J., Ma, B., Rodgers, L., Troge, J., Narzisi, G., Yoon, S., Schatz, M.C., Ye, K., McCombie, W.R., Shendure, J., Eichler, E.E., State, M.W., Wigler, M., 2014. The contribution of *de novo* coding mutations to autism spectrum disorder. Nature 515, 216–221.

Jiang, Y.H., Yuen, R.K., Jin, X., Wang, M., Chen, N., Wu, X., Ju, J., Mei, J., Shi, Y., He, M., Wang, G., Liang, J., Wang, Z., Cao, D., Carter, M.T., Chrysler, C., Drmic, I.E., Howe, J.L., Lau, L., Marshall, C.R., Merico, D., Nalpathamkalam, T., Thiruvahindrapuram, B., Thompson, A., Uddin, M., Walker, S., Luo, J., Anagnostou, E., Zwaigenbaum, L., Ring, R.H., Wang, J., Lajonchere, C., Shih, A., Szatmari, P., Yang, H., Dawson, G., Li, Y., Scherer, S.W., 2013. Detection of clinically relevant genetic variants in autism spectrum disorder by whole-genome sequencing. Am. J. Hum. Genet. 93, 249–263.

Jin, Z.B., Li, Z., Liu, Z., Jiang, Y., Cai, X.B., Wu, J., 2018. Identification of *de novo* germline mutations and causal genes for sporadic diseases using trio-based whole-exome/genome sequencing. Biol. Rev. Camb. Philos. Soc. 93, 1014–1031.

Kong, A., Frigge, M.L., Masson, G., Besenbacher, S., Sulem, P., Magnusson, G., Gudjonsson, S.A., Sigurdsson, A., Jonasdottir, A., Wong, W.S., Sigurdsson, G., Walters, G.B., Steinberg, S., Helgason, H., Thorleifsson, G., Gudbjartsson, D.F., Helgason, A., Magnusson, O.T., Thorsteinsdottir, U., Stefansson, K., 2012. Rate of *de novo* mutations and the importance of father's age to disease risk. Nature 488, 471–475.

Krumm, N., O'Roak, B.J., Shendure, J., Eichler, E.E., 2014. A *de novo* convergence of autism genetics and molecular neuroscience. Trends Neurosci. 37, 95–105.

Kumar, P., Henikoff, S., Ng, P.C., 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat. Protoc. 4, 1073–1082.

Leblond, C.S., Heinrich, J., Delorme, R., Proepper, C., Betancur, C., Huguet, G., Konyukh, M., Chaste, P., Ey, E., Rastam, M., Anckarsater, H., Nygren, G., Gillberg, I.C., Melke, J., Toro, R., Regnault, B., Fauchereau, F., Mercati, O., Lemiere, N., Skuse, D., Poot, M., Holt, R., Monaco, A.P., Jarvela, I., Kantojarvi, K., Vanhala, R., Curran, S., Collier, D.A., Bolton, P., Chiocchetti, A., Klauck, S.M., Poustka, F., Freitag, C.M., Waltes, R., Kopp, M., Duketis, E., Bacchelli, E., Minopoli, F., Ruta, L., Battaglia, A., Mazzone, L., Maestrini, E., Sequeira, A.F., Oliveira, B., Vicente, A., Oliveira, G., Pinto, D., Scherer, S.W., Zelenika, D., Delepine, M., Lathrop, M., Bonneau, D., Guinchat, V., Devillard, F., Assouline, B., Mouren, M.C., Leboyer, M., Gillberg, C., Boeckers, T.M., Bourgeron, T., 2012. Genetic and functional analyses of *shank2* mutations suggest a multiple hit model of autism spectrum disorders. PLoS Genet. 8 e1002521.

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079.

Li, J., Jiang, Y., Wang, T., Chen, H., Xie, Q., Shao, Q., Ran, X., Xia, K., Sun, Z.S., Wu, J., 2015. mirTrios: an integrated pipeline for detection of *de novo* and rare inherited mutations from trios-based next-generation sequencing. J. Med. Genet. 52, 275–281.

Lim, E.T., Raychaudhuri, S., Sanders, S.J., Stevens, C., Sabo, A., MacArthur, D.G., Neale, B.M., Kirby, A., Ruderfer, D.M., Fromer, M., Lek, M., Liu, L., Flannick, J., Ripke, S., Nagaswamy, U., Muzny, D., Reid, J.G., Hawes, A., Newsham, I., Wu, Y., Lewis, L., Dinh, H., Gross, S., Wang, L.S., Lin, C.F., Valladares, O., Gabriel, S.B., dePristo, M., Altshuler, D.M., Purcell, S.M., State, M.W., Boerwinkle, E., Buxbaum, J.D., Cook, E.H., Gibbs, R.A., Schellenberg, G.D., Sutcliffe, J.S., Devlin, B., Roeder, K., Daly, M.J., 2013. Rare complete knockouts in humans: population distribution and significant role in autism spectrum disorders. Neuron 77, 235–242.

Matson, J.L., Shoemaker, M., 2009. Intellectual disability and its relationship to autism spectrum disorders. Res. Dev. Disabil. 30, 1107–1114.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303.

Michaelson, J.J., Shi, Y., Gujral, M., Zheng, H., Malhotra, D., Jin, X., Jian, M., Liu, G., Greer, D., Bhandari, A., Wu, W., Corominas, R., Peoples, A., Koren, A., Gore, A.,

Kang, S., Lin, G.N., Estabillo, J., Gadomski, T., Singh, B., Zhang, K., Akshoomoff, N., Corsello, C., McCarroll, S., Iakoucheva, L.M., Li, Y., Wang, J., Sebat, J., 2012. Whole-genome sequencing in autism identifies hot spots for *de novo* germline mutation. Cell 151, 1431–1442.

Mullins, C., Fishell, G., Tsien, R.W., 2016. Unifying views of autism spectrum disorders: a consideration of autoregulatory feedback loops. Neuron 89, 1131–1156.

O'Roak, B.J., Vives, L., Fu, W., Egertson, J.D., Stanaway, I.B., Phelps, I.G., Carvill, G., Kumar, A., Lee, C., Ankenman, K., Munson, J., Hiatt, J.B., Turner, E.H., Levy, R., O'Day, D.R., Krumm, N., Coe, B.P., Martin, B.K., Borenstein, E., Nickerson, D.A., Mefford, H.C., Doherty, D., Akey, J.M., Bernier, R., Eichler, E.E., Shendure, J., 2012. Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. Science 338, 1619–1622.

Petersen, A.K., Ahmad, A., Shafiq, M., Brown-Kipphut, B., Fong, C.T., Anwar Iqbal, M., 2013. Deletion 1q43 encompassing only *CHRM3* in a patient with autistic disorder. Eur. J. Med. Genet. 56, 118–122.

RK, Yuen, C., Merico, D., Bookman, M., Howe J, L., Thiruvahindrapuram, B., Patel, R.V., Whitney, J., Deflaux, N., Bingham, J., Wang, Z., Pellecchia, G., Buchanan, J.A., Walker, S., Marshall, C.R., Uddin, M., Zarrei, M., Deneault, E., D'Abate, L., Chan, A.J., Koyanagi, S., Paton, T., Pereira, S.L., Hoang, N., Engchuan, W., Higginbotham, E.J., Ho, K., Lamoureux, S., Li, W., MacDonald, J.R., Nalpathamkalam, T., Sung, W.W., Tsoi, F.J., Wei, J., Xu, L., Tasse, A.M., Kirby, E., Van Etten, W., Twigger, S., Roberts, W., Drmic, I., Jilderda, S., Modi, B.M., Kellam, B., Szego, M., Cytrynbaum, C., Weksberg, R., Zwaigenbaum, L., Woodbury-Smith, M., Brian, J., Senman, L., Iaboni, A., Doyle-Thomas, K., Thompson, A., Chrysler, C., Leef, J., Savion-Lemieux, T., Smith, I.M., Liu, X., Nicolson, R., Seifer, V., Fedele, A., Cook, E.H., Dager, S., Estes, A., Gallagher, L., Malow, B.A., Parr, J.R., Spence, S.J., Vorstman, J., Frey, B.J., Robinson, J.T., Strug, L.J., Fernandez, B.A., Elsabbagh, M., Carter, M.T., Hallmayer, J., Knoppers, B.M., Anagnostou, E., Szatmari, P., Ring, R.H., Glazer, D., Pletcher, M.T., Scherer, S.W., 2017. Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. Nat. Neurosci. 20, 602–611.

Rosti, R.O., Sadek, A.A., Vaux, K.K., Gleeson, J.G., 2014. The genetic landscape of autism spectrum disorders. Dev. Med. Child Neurol. 56, 12–18.

Saghatelyan, A., de Chevigny, A., Schachner, M., Lledo, P.M., 2004. Tenascin-R mediates activity-dependent recruitment of neuroblasts in the adult mouse forebrain. Nat. Neurosci. 7, 347–356.

Schwarz, J.M., Rodelsperger, C., Schuelke, M., Seelow, D., 2010. MutationTaster evaluates disease-causing potential of sequence alterations. Nat. Methods 7, 575–576.

Stein, J.L., Parikshak, N.N., Geschwind, D.H., 2013. Rare inherited variation in autism: beginning to see the forest and a few trees. Neuron 77, 209–211.

Steinberg, J., Webber, C., 2013. The roles of FMRP-regulated genes in autism spectrum disorder: single- and multiple-hit genetic etiologies. Am. J. Hum. Genet. 93, 825–839.

Stewart, L.R., Hall, A.L., Kang, S.H., Shaw, C.A., Beaudet, A.L., 2011. High frequency of known copy number abnormalities and maternal duplication 15q11-q13 in patients with combined schizophrenia and epilepsy. BMC Med. Genet. 12, 154.

Tammimies, K., Marshall, C.R., Walker, S., Kaur, G., Thiruvahindrapuram, B., Lionel, A.C., Yuen, R.K., Uddin, M., Roberts, W., Weksberg, R., Woodbury-Smith, M., Zwaigenbaum, L., Anagnostou, E., Wang, Z., Wei, J., Howe, J.L., Gazzellone, M.J., Lau, L., Sung, W.W., Whitten, K., Vardy, C., Crosbie, V., Tsang, B., D'Abate, L., Tong, W.W., Luscombe, S., Doyle, T., Carter, M.T., Szatmari, P., Stuckless, S., Merico, D., Stavropoulos, D.J., Scherer, S.W., Fernandez, B.A., 2015. Molecular diagnostic yield of chromosomal microarray analysis and whole-exome sequencing in children with autism spectrum disorder. J. Am. Med. Assoc. 314, 895–903.

Tan, C., Del Gaudio, D., Dempsey, M., Arndt, K., Botes, S., Reeder, A., Das, S., 2014. Analysis of ASPM in an ethnically diverse cohort of 400 patient samples: perspectives of the molecular diagnostic laboratory. Clin. Genet. 85, 353–358.

Toma, C., Torrico, B., Hervas, A., Valdes-Mas, R., Tristan-Noguero, A., Padillo, V., Maristany, M., Salgado, M., Arenas, C., Puente, X.S., Bayes, M., Cormand, B., 2014. Exome sequencing in multiplex autism families suggests a major role for heterozygous truncating mutations. Mol. Psychiatr. 19, 784–790.

Turner, T.N., Coe, B.P., Dickel, D.E., Hoekzema, K., Nelson, B.J., Zody, M.C., Kronenberg, Z.N., Hormozdiari, F., Raja, A., Pennacchio, L.A., Darnell, R.B., Eichler, E.E., 2017. Genomic patterns of *de novo* mutation in simplex autism. Cell 171, 710–722 e712.

Urraca, N., Cleary, J., Brewer, V., Pivnick, E.K., McVicar, K., Thibert, R.L., Schanen, N.C., Esmer, C., Lamport, D., Reiter, L.T., 2013. The interstitial duplication 15q11.2-q13 syndrome includes autism, mild facial anomalies and a characteristic EEG signature. Autism Res. 6, 268–279.

Vorstman, J.A.S., Parr, J.R., Moreno-De-Luca, D., Anney, R.J.L., Nurnberger Jr., J.I., Hallmayer, J.F., 2017. Autism genetics: opportunities and challenges for clinical translation. Nat. Rev. Genet. 18, 362–376.

Wang, J., Mullighan, C.G., Easton, J., Roberts, S., Heatley, S.L., Ma, J., Rusch, M.C., Chen, K., Harris, C.C., Ding, L., Holmfeldt, L., Payne-Turner, D., Fan, X., Wei, L., Zhao, D., Obenauer, J.C., Naeve, C., Mardis, E.R., Wilson, R.K., Downing, J.R., Zhang, J., 2011. CREST maps somatic structural variation in cancer genomes with base-pair resolution. Nat. Methods 8, 652–654.

Wang, K., Li, M., Hakonarson, H., 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 38, e164.

Weber, P., Bartsch, U., Rasband, M.N., Czaniera, R., Lang, Y., Bluethmann, H., Margolis, R.U., Levinson, S.R., Shrager, P., Montag, D., Schachner, M., 1999. Mice deficient for tenascin-R display alterations of the extracellular matrix and decreased axonal conduction velocities in the CNS. J. Neurosci. 19, 4245–4262.

Wilfert, A.B., Sulovari, A., Turner, T.N., Coe, B.P., Eichler, E.E., 2017. Recurrent *de novo* mutations in neurodevelopmental disorders: properties and clinical implications. Genome Med. 9, 101.

Winnepenninckx, B., Debacker, K., Ramsay, J., Smeets, D., Smits, A., FitzPatrick, D.R., Kooy, R.F., 2007. CGG-repeat expansion in the *DIP2B* gene is associated with the fragile site FRA12A on chromosome 12q13.1. Am. J. Hum. Genet. 80, 221–231.

Yang, Y.J., Baltus, A.E., Mathew, R.S., Murphy, E.A., Evrony, G.D., Gonzalez, D.M., Wang, E.P., Marshall-Walker, C.A., Barry, B.J., Murn, J., Tatarakis, A., Mahajan, M.A., Samuels, H.H., Shi, Y., Golden, J.A., Mahajnah, M., Shenhav, R., Walsh, C.A., 2012. Microcephaly gene links trithorax and REST/NRSF to control neural stem cell proliferation and differentiation. Cell 151, 1097–1112.

Yin, J., Schaaf, C.P., 2017. Autism genetics - an overview. Prenat. Diagn. 37, 14–30.

Yu, T.W., Chahrour, M.H., Coulter, M.E., Jiralerspong, S., Okamura-Ikeda, K., Ataman, B., Schmitz-Abe, K., Harmin, D.A., Adli, M., Malik, A.N., D'Gama, A.M., Lim, E.T., Sanders, S.J., Mochida, G.H., Partlow, J.N., Sunu, C.M., Felie, J.M., Rodriguez, J., Nasir, R.H., Ware, J., Joseph, R.M., Hill, R.S., Kwan, B.Y., Al-Saffar, M., Mukaddes, N.M., Hashmi, A., Balkhy, S., Gascon, G.G., Hisama, F.M., LeClair, E., Poduri, A., Oner, O., Al-Saad, S., Al-Awadi, S.A., Bastaki, L., Ben-Omran, T., Teebi, A.S., Al-Gazali, L., Eapen, V., Stevens, C.R., Rappaport, L., Gabriel, S.B., Markianos, K., State, M.W., Greenberg, M.E., Taniguchi, H., Braverman, N.E., Morrow, E.M., Walsh, C.A., 2013. Using whole-exome sequencing to identify inherited causes of autism. Neuron 77, 259–273.

Yu, T.W., Mochida, G.H., Tischfield, D.J., Sgaier, S.K., Flores-Sarnat, L., Sergi, C.M., Topcu, M., McDonald, M.T., Barry, B.J., Felie, J.M., Sunu, C., Dobyns, W.B., Folkerth, R.D., Barkovich, A.J., Walsh, C.A., 2010. Mutations in *WDR62*, encoding a centrosome-associated protein, cause microcephaly with simplified gyri and abnormal cortical architecture. Nat. Genet. 42, 1015–1020.

Yuen, R.K., Merico, D., Cao, H., Pellecchia, G., Alipanahi, B., Thiruvahindrapuram, B., Tong, X., Sun, Y., Cao, D., Zhang, T., Wu, X., Jin, X., Zhou, Z., Liu, X., Nalpathamkalam, T., Walker, S., Howe, J.L., Wang, Z., MacDonald, J.R., Chan, A., D'Abate, L., Deneault, E., Siu, M.T., Tammimies, K., Uddin, M., Zarrei, M., Wang, M., Li, Y., Wang, J., Wang, J., Yang, H., Bookman, M., Bingham, J., Gross, S.S., Loy, D., Pletcher, M., Marshall, C.R., Anagnostou, E., Zwaigenbaum, L., Weksberg, R., Fernandez, B.A., Roberts, W., Szatmari, P., Glazer, D., Frey, B.J., Ring, R.H., Xu, X., Scherer, S.W., 2016. Genome-wide characteristics of *de novo* mutations in autism. NPJ Genom. Med. 1, 160271–1602710.

Yuen, R.K., Thiruvahindrapuram, B., Merico, D., Walker, S., Tammimies, K., Hoang, N., Chrysler, C., Nalpathamkalam, T., Pellecchia, G., Liu, Y., Gazzellone, M.J., D'Abate, L., Deneault, E., Howe, J.L., Liu, R.S., Thompson, A., Zarrei, M., Uddin, M., Marshall, C.R., Ring, R.H., Zwaigenbaum, L., Ray, P.N., Weksberg, R., Carter, M.T., Fernandez, B.A., Roberts, W., Szatmari, P., Scherer, S.W., 2015. Whole-genome sequencing of quartet families with autism spectrum disorder. Nat. Med. 21, 185–191.