Journal of Genetics and Genomics xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Genetics and Genomics

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

Original research

Targeted sequencing and integrative analysis of 3195 Chinese patients with neurodevelopmental disorders prioritized 26 novel candidate genes

Tao Wang [a, b, c, 1], Yi Zhang [d, 1], Liqui Liu [b], Yan Wang [b], Huiqian Chen [e], Tianda Fan [f], Jinchen Li [d, g, *], Kun Xia [a, h, i, *], Zhongsheng Sun [b, f, j, k, *]

[a] Center for Medical Genetics & Hunan Key Laboratory of Medical Genetics, School of Life Sciences, Central South University, Changsha 410083, Hunan, China
[b] Beijing Institutes of Life Science, Chinese Academy of Sciences, Beijing 100101, China
[c] DIAGenes Precision Medicine, Beijing 102600, China
[d] National Clinical Research Centre for Geriatric Disorders, Department of Geriatrics, Xiangya Hospital, Central South University, Changsha 410083, Hunan, China
[e] Shanghai Adeptus Biotechnology, Shanghai 200126, China
[f] Institute of Genomic Medicine, Wenzhou Medical University, Wenzhou 325035, China
[g] Department of Neurology, Xiangya Hospital, Central South University, Changsha 410083, Hunan, China
[h] CAS Center for Excellence in Brain Science and Intelligences Technology (CEBSIT), Shanghai 200031, China
[i] School of Basic Medical Science, Central South University, Changsha 410083, Hunan, China
[j] CAS Center for Excellence in Biotic Interactions, University of Chinese Academy of Sciences, Beijing 100049, China
[k] State Key Laboratory of Integrated Management of Pest Insects and Rodents, Chinese Academy of Sciences, Beijing 100101, China

ARTICLE INFO

Article history:
Received 8 December 2020
Received in revised form
5 March 2021
Accepted 7 March 2021
Available online xxx

Keywords:
De novo mutations
X-linked variants
Rare inherited variants
Homozygous mutations
Neurodevelopmental disorders
Autism spectrum disorders
Candidate genes

ABSTRACT

Neurodevelopmental disorders (NDDs) are a set of complex disorders characterized by diverse and co-occurring clinical symptoms. The genetic contribution in patients with NDDs remains largely unknown. Here, we sequence 519 NDD-related genes in 3195 Chinese probands with neurodevelopmental phenotypes and identify 2522 putative functional mutations consisting of 137 *de novo* mutations (DNMs) in 86 genes and 2385 rare inherited mutations (RIMs) with 22 X-linked hemizygotes in 13 genes, 2 homozygous mutations in 2 genes and 23 compound heterozygous mutations in 10 genes. Furthermore, the DNMs of 16,807 probands with NDDs are retrieved from public datasets and combine in an integrated analysis with the mutation data of our Chinese NDD probands by taking 3582 in-house controls of Chinese origin as background. We prioritize 26 novel candidate genes. Notably, six of these genes — *ITSN1*, *UBR3, CADM1, RYR3*, *FLNA*, and *PLXNA3* — preferably contribute to autism spectrum disorders (ASDs), as demonstrated by high co-expression and/or interaction with ASD genes confirmed via rescue experiments in a mouse model. Importantly, these genes are differentially expressed in the ASD cortex in a significant manner and involved in ASD-associated networks. Together, our study expands the genetic spectrum of Chinese NDDs, further facilitating both basic and translational research.

Copyright © 2021, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, and Genetics Society of China. Published by Elsevier Limited and Science Press. All rights reserved.

Introduction

Neurodevelopmental disorders (NDDs), commonly manifested as intellectual disabilities, behavior abnormalities, and social and cognitive impairments, affect approximately 1–2% of the global population (Sabus et al., 2019). Symptoms of heterogeneity are frequently observed in patients with NDDs, including different core symptoms in different types of NDDs and diversified symptoms in different patients with the same type of NDD, even within the same patient (Matson and Shoemaker, 2009; Kanner, 2016; Zheng et al., 2018; de Crescenzo et al., 2019; Taylor et al., 2019; Sierra-Arregui et al., 2020). In general, autism spectrum disorders (ASDs) are characterized by defective social communication and restricted/repetitive behaviors (Hegarty et al., 2019; Sullivan and Geschwind, 2019). Schizophrenia (SCZ) patients primarily experience

* Corresponding authors.
E-mail addresses: lijinchen@csu.edu.cn (J. Li), xiakun@sklmg.edu.cn (K. Xia), sunzs@biols.ac.cn (Z. Sun).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jgg.2021.03.002
1673-8527/Copyright © 2021, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, and Genetics Society of China. Published by Elsevier Limited and Science Press. All rights reserved.

Please cite this article as: T. Wang, Y. Zhang, L. Liu *et al.*, Targeted sequencing and integrative analysis of 3195 Chinese patients with neurodevelopmental disorders prioritized 26 novel candidate genes, Journal of Genetics and Genomics, https://doi.org/10.1016/j.jgg.2021.03.002

longstanding delusions and hallucinations (Sullivan and Geschwind, 2019). Intellectual disability (ID) is characterized by typical signs of severe deficits in cognition, while epileptic encephalopathy (EE) is characterized by recurrent unprovoked seizures (Kearney et al., 2019). A previous meta-analysis has estimated that nearly 70% of patients with ASD experience at least one additional psychiatric disorder (Lugo-Marín et al., 2019). For instance, in the 22,176 patients with ASD, 11.8% of them were diagnosed with SCZ. Another meta-analysis of 19 studies on patients with epilepsy found a coupled ASD prevalence of 6.3% (Strasser et al., 2018). Furthermore, Breuillard et al. (2016) recruited eight female patients with EE carrying PCDH-19 mutations, of which six were diagnosed with ASD and ID simultaneously (Breuillard et al., 2016). Therefore, NDDs complexity severely affects the diagnostic accuracy, etiological exploration, and ultimately treatment.

We have previously indicated that shared and unique clinical symptoms in NDDs are attributed to convergent and divergent genetic features, respectively (Li et al., 2016). Previous genetic studies have revealed that some convergent patterns among NDDs are caused by copy number variations (CNVs) (Coe et al., 2019), common variants (Xia et al., 2019), and rare variants (Coe et al., 2019). Therefore, integration analysis of genetic data on different types of NDDs may facilitate the identification of novel risk genes and further promote the stratification of NDDs, thereby improving diagnostic accuracy.

NDDs associated disabilities have affected many families' quality of life (Sullivan and Geschwind, 2019), which, despite extensive research, their needs remain unmet. For instance, worldwide epidemiological investigations have indicated that the incidence of ASD has considerably increased in the past few years (Christensen et al., 2016; Ji et al., 2020). In China, approximately 0.7% of children between the age of 6 and 12 years suffer from ASD (Zhou et al., 2020). However, only a small portion of children diagnosed with ASD can be explained by genetic findings; the etiology of the remaining 75% of cases remains undetermined (Fernandez and Scherer, 2017). Hence, discovering novel risk genes for NDDs remains challenging.

De novo mutation (DNM), a major pathogenic risk factor, has been well established in identifying candidate genes in NDDs, such as ASD (de Rubeis et al., 2014; Iossifov et al., 2014; Sanders et al., 2015; Yuen et al., 2017; Takata et al., 2018; Wu et al., 2018; Satterstrom et al., 2020), ID (Matson and Shoemaker, 2009; Hamdan et al., 2014; Lelieveld et al., 2016), EE (Epi4K Consortium et al., 2013; Kanner, 2016; Sierra-Arregui et al., 2020) and SCZ (Xu et al., 2011, 2012; Zheng et al., 2018; de Crescenzo et al., 2019; Sullivan and Geschwind, 2019). Via integrative analysis of DNM data generated using whole-exome sequencing (WES) for 10,927 patients with NDDs, and targeted sequencing of 125 genes for 17,429 NDD patients with a Caucasian origin, we and other studies, have identified an abundance of shared and unique pathogenic genes (Li et al., 2016; Coe et al., 2019; Wang et al., 2020). In addition to DNM, rare inherited variants (RIMs) and X-linked hemizygous variants are underlying pathogenic candidate factors for the etiology of NDDs (Lim et al., 2013; Toma et al., 2014; Krumm et al., 2015; Al-Mubarak et al., 2017; Jin et al., 2017; Sullivan and Geschwind, 2019). Noteworthily, the commonly used WES (Neale et al., 2012; O'Roak et al., 2012; Sanders et al., 2012; Iossifov et al., 2012, Iossifov et al., 2014; de Rubeis et al., 2014) and whole-genome sequencing (WGS) (Jiang et al., 2013; Yuen et al., 2015, Yuen et al., 2016, Yuen et al., 2017; Wu et al., 2018) in trio-based families rapidly promote our ability to detect DNMs, as well as rare inherited and X-linked hemizygous variants in NDDs, simultaneously (Xu et al., 2012; Hamdan et al., 2014; Lelieveld et al., 2016; Jin et al., 2017; Satterstrom et al., 2020). With well-curated panel genes, targeted sequencing can more efficiently and cost-effectively show the disease-related mutation spectrum and discover risk genes compared with WES and WGS in a large cohort study. Indeed, this approach has been extensively used in large cohort studies on patients with NDD patients. To date, five targeted sequencing studies of Chinese ASD cohorts have been reported with 85−358 targeted genes of 536−1543 probands (Wang et al., 2016; Li et al., 2017a; Guo et al., 2018; Zhou et al., 2019a; Zhou et al., 2019b). As for other types of Chinese NDDs, only ID and EE have been independently studied, reporting 454 targeted genes in 112 families (Yan et al., 2019) and 412 targeted genes in 63 trios (Wang et al., 2017), respectively. Moreover, in our previous studies (Li et al., 2017a; 2017b; Li et al., 2019; Zhang et al., 2020), gene expression and functional network analyses have complemented the investigation of ASD-related genes. We showed that integration of results from targeted genomic sequencing, transcriptome sequencing, and network analysis could be an effective method to identify candidate genes.

To comprehensively illustrate the genetic architecture of NDDs and identify candidate genes, we have conducted targeted deep sequencing of 519 NDD-related genes in 3195 Chinese probands with neurodevelopmental phenotypes. Then, we combined the mutation data of our Chinese NDD patients with the published DNMs reported in 16,807 NDD patients to jointly prioritize candidate genes. Finally, we identified a high-confidence candidate gene set by further integrating their differential expression status in the cortex from published NDD studies and their functional roles in NDD associated networks. To our knowledge, this is the first comprehensive genetic architecture investigation of NDDs focused on a higher number of mutation types and disorders compared to previous studies (Coe et al., 2019; Wang et al., 2020). This study aimed to (1) identify additional novel candidate genes and expand the genetic spectrum of Chinese patients with NDDs, (2) provide multiple evidence to confirm the novel candidate genes involved in the etiology of NDDs, and (3) illustrate the distribution of novel candidate genes among different types of NDDs. Via comprehensive integration analysis, our study has identified 26 novel NDD candidate genes, six of which may preferably contribute to the development of ASDs. Furthermore, we showed several DNMs and recessive mutated genes that may specifically contribute to the risk of Chinese patients on NDDs.

Results

Datasets, quality control, and variant summary

A total of 20,002 NDDs cases (3195 from our Chinese NDDs cohort and the remaining 16,807 from published NDDs cohorts) and 6973 controls (3391 from published unaffected siblings and 3582 from our Inhouse Chinese controls) were included in this study (shown in Methods; Table S1). As for our Chinese NDDs cohort, we analyzed 519 targeted genes with the different levels of evidence association with NDDs (shown in Methods; Table S2) and performed targeted sequencing on 3195 Chinese probands with neurodevelopmental phenotypes (Fig. 1). On average, we obtained 1635.57 MB clean data for each sample with approximately 231× sequence depth in target regions (Table S3). Additionally, approximately 94% of target regions were covered with at least 10× sequence depth (Table S3). After quality control, read mapping, variant detection and annotation, protein-truncating variants (PTVs, including frameshift-, splicing-, stop-gain-, and stop-loss-derived variants) and deleterious missense (Dmis) variants with minor-allele frequencies of less than 0.001 were selected for validation using Sanger sequencing proband and the patient's parents (Fig. 1). In total, 2522 putative functional variants (501 PTVs and 2021 Dmis) in 418 genes were detected and validated through Sanger sequencing (Fig. S1A; Table S4). For published NDDs cohorts, we collected 6511 ASD probands, 1094 SCZ probands, 1331 ID probands, 933 EE probands, 2645 congenital heart defects (CHD) probands, and 4293

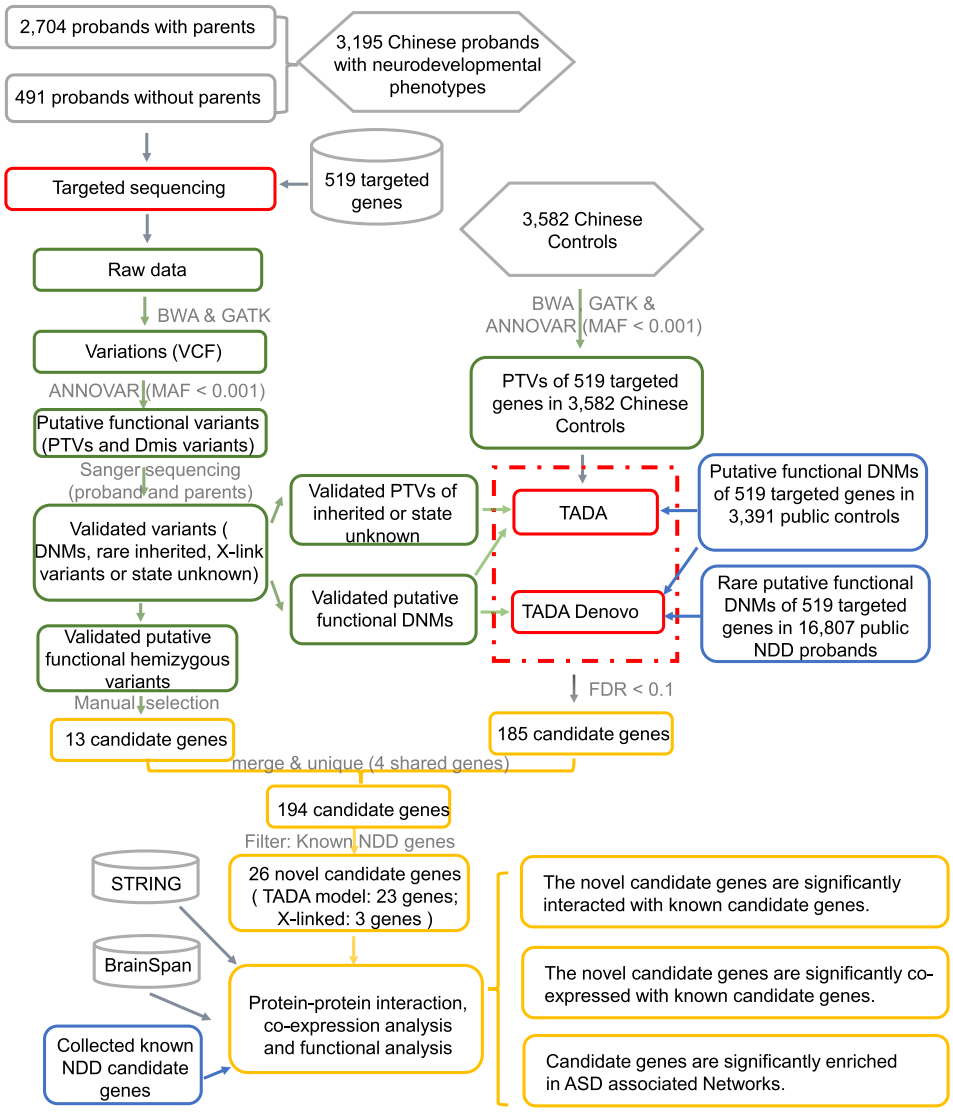

Fig. 1. Workflow of data analysis. Targeted sequencing of 519 NDD-related genes is employed in 3195 probands. Putative functional variants in the 519 genes from 3195 probands are validated by Sanger sequencing. These putative functional variants from Chinese NDD probands, PTVs of 519 genes from our 3582 Chinese controls and putative functional DNMs of 519 genes from public WES/WGS studies (16,807 NDD probands plus 3391 unaffected siblings) are used to prioritize candidate genes using the TADA model. Then, PPI, gene co-expression and functional analysis are employed to investigate the functional impact of candidate genes. Finally, the 26 novel candidate genes and their frequently connected known NDD candidates are used for network enrichment by integrating the results of PPI and co-expression analysis. PPI, protein-protein interaction; TADA, Transmission and *De novo* Association; PTV means protein-truncating variants, which include frameshift-, splicing-, stop-gain-, and stop-loss-derived variants; Dmis variants, deleterious missense variants with ReVe score >0.7; Putative functional variants contained PTVs and Dmis variants.

undiagnosed developmental disorder (UDD) probands, which carried with 567, 69, 274, 144, 156, and 871 putative functional DNMs in targeted gene regions, respectively (Table S1).

DNMs in our Chinese cohort

In 2704 unrelated trios-families, we identified 137 (5.4%) putative functional DNMs residing in 86 genes of 132 probands, of which 16 genes had multiple DNMs (Table 1). These recurrent *de novo* mutated genes have been reported in several different types of NDDs cohorts (Neale et al., 2012; Talkowski et al., 2012; de Rubeis et al., 2014; Iossifov et al., 2014; Wang et al., 2016; Lim et al., 2017; Yuen et al., 2017; Stessman et al., 2017; An et al., 2018; Guo et al., 2019; Zhou et al., 2019; Satterstrom et al., 2020), suggesting shared genetic components in different types of NDD disorders (Tables 1 and S5). Additionally, compared with published Chinese studies on NDDs, our study reported 48 out of the 86 genes with DNMs for the first time (Table S6). We also found that *SCN2A* is the most frequently mutated gene harboring 13 putative functional DNMs (9 PTVs and 4 Dmis variants) in ~0.48% (13/2704) of the probands, followed by *MECP2*, *DEAF1*, *SHANK3*, *TANC2*, *ARID1B*, *KDM5B*, *CHD8*, *TCF4*, *DYNC1H1*, *TSC2*, *CREBBP*, *SETBP1*, *MEF2C*, *SYNE1*, and *EHMT1*. Interestingly, three DNMs in *MECP2* were recurrently observed, including a triplicate stop-gain variant (c.538C>T, p.R180X), a duplicate stop-gain variant (c.916C>T, p.R306X), and a duplicate Dmis variant (c.509C>T, p.T170M) (Table S5), indicating that these DNMs are mutational hotspots, that might be associated with NDDs.

To further evaluate the functional impact of mutated genes, the probability of loss-of-function intolerance (pLI) (Lek et al., 2016), and the residual variation intolerance score (RVIS) (Petrovski et al., 2013) were used. Genes with higher pLI may pose a higher risk for patients if the putative functional mutations occur on these genes and vice-versa for genes with lower RVIS. As expected, our results showed that genes with recurrent DNMs exhibited significantly higher pLI and

Table 1
Sixteen genes with recurrent DNMs identified in our Chinese samples.

Gene symbol	DNMs ($n = 2704$)	Inherited or state unknown ($n = 3195$)	DNMs in NDD from Gene4Denovo ($n = 16,807$)	RVIS (percentile)	pLi_EXAC (percentile)	Gene function
SCN2A	9 PTV, 4 Dmis	2 PTV, 7 Dmis	18 PTV, 32 Dmis	−2.51 (1.1%)	1.00 (0.72%)	Behavior and regulation of ion transport
MECP2	7 PTV, 4 Dmis	6 Dmis	9 PTV, 10 Dmis	−0.32 (30.3%)	0.70 (24.62%)	Synapse organization and chromatin modification
DEAF1	6 Dmis	1 PTV, 2 Dmis	4 Dmis	−0.59 (18.8%)	1.93e-04 (69.85%)	Behavior and associative learning
SHANK3	5 PTV	5 PTV	9 PTV	–	1.00 (4.13%)	Synapse organization and behavior
TANC2	4 PTV	1 PTV, 12 Dmis	3 PTV, 4 Dmis	−3.88 (0.3%)	1.00 (0.90%)	Synapse organization
ARID1B	1 PTV, 3 Dmis	4 Dmis	49 PTV, 2 Dmis	−2.21 (1.5%)	1.00 (3.21%)	Synapse organization and chromatin remodeling
KDM5B	1 PTV, 2 Dmis	1 PTV, 3 Dmis	11 PTV, 3 Dmis	−2.30 (1.3%)	5.09e-05 (73.77%)	Chromatin modification and remodeling
CHD8	3 PTV	7 PTV, 11 Dmis	16 PTV, 4 Dmis	−2.36 (1.2%)	1.00 (0.35%)	Behavior and chromatin remodeling
TCF4	3 Dmis	–	15 PTV, 4 Dmis	−0.79 (13.1%)	1.00 (3.96%)	Synapse organization
DYNC1H1	1 PTV, 1 Dmis	1 PTV, 9 Dmis	2 PTV, 14 Dmis	−8.60 (0.01%)	1.00 (0.01%)	Neuron migration
TSC2	2 PTV	26 Dmis	6 Dmis	−2.95 (0.7%)	1.00 (0.82%)	Synapse organization and behavior
CREBBP	2 PTV	6 Dmis	4 PTV, 13 Dmis	−4.23 (0.2%)	1.00 (0.17%)	Covalent chromatin modification and associative learning
SETBP1	1 PTV, 1 Dmis	1 PTV, 8 Dmis	5 PTV, 3 Dmis	−0.96 (9.7%)	0.99 (5.85%)	Regulation of transcription, DNA-templated
MEF2C	1 PTV, 1 Dmis	–	5 PTV, 7 Dmis	−0.58 (19.0%)	4.25e-03 (58.81%)	Synapse organization, behavior, and regulation of ion transport
SYNE1	2 Dmis	8 PTV, 23 Dmis	3 Dmis	−1.10 (7.7%)	3.75e-27 (99.45%)	Nuclear matrix anchoring at nuclear membrane
EHMT1	2 Dmis	2 PTV, 5 Dmis	10 PTV, 3 Dmis	−2.09 (1.7%)	1.00 (1.08%)	Covalent chromatin modification

FDR, false discovery rate; -, without information; RVIS, residual variation intolerance score. ExAC, Exome Aggregation Consortium; pLI, probability of loss-of-function intolerance (Genes with higher pLI may pose a higher risk for patients if putative functional mutations occur on these genes. Gene with pLI $\geq$ 0.9 was considered as highly constrained).

lower RVIS ($P = 6.55 \times 10^{-5}$ and $P = 7.75 \times 10^{-8}$, respectively, two-tailed Wilcoxon rank-sum test, Fig. S2A) than the background genes (all RefSeq genes after removing the known NDD-candidate genes); specifically, 12 (75.0%) and 15 (93.75%) of the 16 recurrent genes ranked in the top 50% of pLI and RVIS, respectively (Table 1). Overall, the results suggest that these genes with recurrent DNMs are more sensitive to putative functional variants than the background genes, and therefore might contribute to the disease-associated risks.

RIMs in our Chinese cohort

Apart from the 137 DNMs, the remaining 2385 variants of a total of 2522 (435 PTVs and 1950 Dmis variants) were inherited, or their inheritance status was unknown (Table S3). Interestingly, we observed a slight bias for maternally inherited PTVs (156 maternal vs. 128 paternal, OR = 1.23, $P = 0.06$), which was not observed for maternally inherited Dmis variants (OR = 0.96, $P = 0.73$) (Fig. S1B). We also detected 22 X-linked hemizygous variants in 13 genes (Table S7), including ten known and three newly identified genes (*CACNA1F*, *FLNA*, and *PLXNA3*). Moreover, all the three newly identified X-linked genes harbored two putative functional hemizygous variants, including one splicing variant (c.1114+1G>A) and one Dmis variant (c.3056C>T, p.T1019) in *CACNA1F*, two Dmis variants (c.7766C>G, p.P2589R and c.4762G>A, p.E1588K) in *FLNA*, and two Dmis variants (c.1249G>A, p.V417M and c.3671T>G, p.L1224R) in *PLXNA3* (Table S6). Among the other ten known candidate genes, *NLGN3* was the most frequently mutated gene harboring four putative functional X-linked hemizygous variants (one PTV and three Dmis variants; ~0.17% [4/2360] of male probands), followed by *MECP2*, *NLGN4X*, *TMLHE*, *ARX*, *BCORL1*, *KDM5C*, *PTCHD1*, *SLC35A2*, and *SLC9A6* (Table S6).

Prioritization of candidate genes and investigation of their functional impact on both cellular and molecular levels

TADA, an integrated hierarchical Bayesian framework, was used to prioritize the 467 out of the 519 targeted genes with putative functional variants identified in either our Chinese NDD cohort or public NDDs cohorts (shown in Methods; Table S8). In our Chinese NDD cohort analysis, candidate genes were prioritized based on putative functional genetic variants in the 3195 NDD probands against those in the 3582 in-house controls as background (shown in Methods). Eleven candidate genes (*SCN2A*, *CHD8*, *KDM5B*, *DEAF1*, *SHANK3*, *CREBBP*, *TANC2*, *ARID1B*, *MECP2*, *TCF4*, and *MEF2C*) showed an FDR < 0.1. All of which were previously associated with NDDs, and two (*KDM5B* and *CREBBP*) were reported for the first time in Chinese NDD cohorts with putative functional DNMs (Table S9).

Since many pathogenic genes are shared among different types of NDDs (Li et al., 2016) while the integration of datasets from multiple independent studies can increase the power of candidate gene prioritization (Liu et al., 2013; Ben-David and Shifman, 2013; de Rubeis et al., 2014; Hoischen et al., 2014; Nguyen et al., 2017; Stessman et al., 2017), we next combined the putative functional DNMs identified in our study with those reported in other NDDs cohorts with data publicly available (Table S10). In the combined cohort analysis, we prioritized 185 candidate genes (Table S11) (our cohort, $n = 2704$; public cohorts, $n = 16,807$; public controls, $n = 3391$) with FDR values < 0.1. Moreover, we noted that 137 out of the 185 candidate genes were considered as strong candidates, meeting the more stringent condition of FDR < 0.01. Among the 185 candidate genes, 23 were defined as novel candidates by comparison with 806 collected known NDD candidate genes and further screened in the PubMed database with no obvious genetic evidence of association with NDDs (shown in Methods; Tables 2 and S9). In addition, 13 genes harboring hemizygous variants were manually listed as candidates for their reported involvement in ASD, mental retardation, or other X-linked neurodevelopmental disorders. Three of these 13 X-linked genes were considered as novel candidates for their absence in collected known NDD candidate genes. Interestingly, four genes, namely *MECP2*, *SLC35A2*, *PTCHD1*, and *KDM5C*, were both with putative functional hemizygous variants and significantly *de novo* mutated. By removing redundant genes, in total, 194 genes were prioritized as NDD candidates, and 26 of them were considered as novel candidates. Both total candidate genes and novel candidate genes showed significantly higher pLI and lower RVIS than those of background genes ($n = 194$, $P < 2.20 \times 10^{-16}$ for pLI and $P < 2.20 \times 10^{-16}$ for RVIS, Fig. S2B; $n = 26$, $P = 5.87 \times 10^{-4}$ for pLI and $P = 5.13 \times 10^{-7}$ for RVIS, Fig. S2C). We also found that putative functional DNMs, X-linked hemizygous variants, and inherited or unknown status variants of the 194 candidate genes were detected in

~4.51% (122/2704), ~0.93% (22/2360), and ~27.07% (865/3195) of Chinese probands, respectively (Table 3).

Cell-specific enrichment analysis (CSEA) (Xu et al., 2014) showed that the 194 candidate genes were most significantly associated with medium spiny neurons of the striatum (striatum D1+ and D2+ medium spiny neurons, $P_{BH} = 3.00 \times 10^{-2}$ and $P_{BH} = 5.00 \times 10^{-3}$, one-tailed Fisher's exact test, Fig. S3), which was in accordance with the findings of previous studies (Tripathi et al., 2015). We further applied the MetaScape tool (Zhou et al., 2019a; Zhou et al., 2019b) to examine the GO/KEGG enrichment of the 194 candidate genes. The top five significantly enriched GO terms were synapse organization, covalent chromatin modification, head development, behavior, and regulation of ion transport (Fig. S4; Table S13). We compared the 194 NDD candidate genes with two previous gene sets discovered by the common variant approach from a large cohort of the Han Chinese population with autism and a large combined cohort of five psychiatric disorders in European ancestry (Xia et al., 2019, 2020). In the Han Chinese autism cohort, Xia et al. (2020) discovered 14 SNPs, annotated as either intronic or intergenic, associated with 10 genes. Of those, only *EEF1A2* reached a significant level with FDR $= 5.80 \times 10^{-10}$ (Table S12). Another study by Xia et al. (2019) identified 69 SNPs, most of them annotated as either intronic or intergenic, associated with 186 genes. Five of those genes were included in our panel, and each reached a significant level (Table S12). All the six shared significant genes are known NDD candidates reported previously (Tables S9 and S11). The 196 genes identified by common variants in studies by Xia et al. (2019, 2020) are significantly enriched in the GO term of "modulation of chemical synaptic transmission" (top 1) (Fig. S5). The top 1 significantly enriched GO terms by both our 194 candidate genes and Xia's 196 genes (2019, 2020) are associated with synapse function. These results indicate that genes identified by either rare or common variants showed a convergence pattern in the pathway.

The novel candidate genes are functionally associated with known genes

The 26 novel candidate genes showed significantly higher pLI (two-tailed Wilcoxon rank-sum test, $P = 7.42 \times 10^{-5}$) and lower RVIS ($P = 3.89 \times 10^{-8}$) than those of background genes, suggesting intolerance to putative functional variants (Fig. S2C). Furthermore, a permutation test (iterations = 1,000,000) was used to determine the relationship between the 26 novel candidate genes and the 806 collected known NDDs candidate genes, based on co-expression data from the human brain and PPI data from the STRING database (shown in Methods). We found that the novel candidate genes were significantly co-expressed and interacted with known NDD candidate genes compared with the random expectation. In the human brain expression data, 11 out of the 26 novel candidate genes ($P = 3.13 \times 10^{-2}$, Fig. S6A) were co-expressed with 168 known NDD candidate genes ($P = 6.20 \times 10^{-4}$, Fig. S6B), with 257 connections ($P = 7.46 \times 10^{-3}$, Fig. S6C). Based on PPI data, all 26 novel candidate genes ($P = 6.64 \times 10^{-4}$, Fig. S6D) interacted with 193 known NDD candidate genes ($P = 6.03 \times 10^{-2}$, Fig. S6E), with 284 connections ($P = 1.83 \times 10^{-2}$, Fig. S6F). Collectively, by removing redundant genes, 319 unique known NDD candidate genes were co-expressed and/or interacted with the novel candidate genes, of which 22 genes were validated by a mouse rescue model documented in the SFARI Gene database (Tables S9 and S14), were further considered as validated ASD candidates. Finally, we found 18 out of the 26 novel candidate genes co-expressed and/or interacted with these validated ASD candidates (Table S14), 10 of which (*UBR4*,

Table 2
Integrated analysis based on DNMs prioritized 23 novel candidate genes.

Gene symbol	DNMs ($n = 2704$)	Inherited or state unknown ($n = 3195$)	DNMs in NDD from Gene4Denovo ($n = 16,807$)	FDR (combined our and public NDD data)	pLi_EXAC (percentile)	Gene function
FRYL	–	1 PTV, 19 Dmis	6 PTV	1.16E-04	1.00 (0.72%)	Neuron development and neuron differentiation
LPHN2	1 Dmis	5 Dmis	1 PTV, 3 Dmis	1.61E-03	1.00 (0.96%)	Excitatory neurotoxin presented in black widow spider venom
ITSN1	–	6 Dmis	3 PTV, 2 Dmis	1.89E-03	1.00 (0.87%)	Nervous system and head development, synapse function
GALNT18	1 Dmis	7 Dmis	2 PTV, 1 Dmis	2.18E-03	4.77E-02 (47.21%)	Catalyzes the initial reaction in O-linked oligosaccharide biosynthesis
BIRC6	–	1 PTV, 27 Dmis	3 PTV, 4 Dmis	3.07E-03	1.00 (0.10%)	Cell cycle process and microtubule cytoskeleton organization
OR10Z1	1 Dmis	1 PTV, 4 Dmis	1 PTV, 1 Dmis	3.87E-03	7.36E-03 (56.32%)	Odorant receptor plays a key role in olfactory transduction
UBR3	–	3 Dmis	3 PTV	5.41E-03	1.00 (5.02%)	Behavior
UBR4	1 PTV	26 Dmis	2 PTV, 5 Dmis	9.10E-03	1 (0.01%)	Nervous system development
MYO7B	1 Dmis	5 PTV, 13 Dmis	4 Dmis	1.24E-02	6.38E-16 (97.34%)	Vesicle-mediated transport in synapse and regulation of ion transport
TECTA	1 Dmis	3 PTV, 17 Dmis	2 PTV, 2 Dmis	1.56E-02	1.07E-12 (95.18%)	Sensory perception of sound
PINK1	1 PTV	6 Dmis	1 PTV	2.31E-02	4.08E-07 (83.98%)	Chromatin organization and synapse function
LRRC4	–	1 PTV, 3 Dmis	1 PTV, 2 Dmis	2.56E-02	0.88 (18.68%)	Synapse organization and chemical synaptic transmission
CADM1	1 PTV	–	1 Dmis	3.14E-02	0.99 (8.63%)	Vesicle-mediated transport in synapse
RFX7	–	5 Dmis	2 PTV, 1 Dmis	4.19E-02	1.00 (4.67%)	Regulation of the transcription by RNA polymerase II
CACNA1S	–	11 Dmis	2 PTV, 2 Dmis	4.32E-02	9.56E-08 (86.11%)	Cholinergic synapse
ALS2	1 Dmis	9 Dmis	1 PTV, 1 Dmis	4.33E-02	7.70E-04 (65.57%)	Nervous system development and synapse function
MSL2	–	2 Dmis	1 PTV	5.47E-02	0.89 (18.19%)	Covalent chromatin modification
RYR3	1 Dmis	3 PTV, 28 Dmis	1 PTV, 5 Dmis	5.62E-02	1.00 (0.96%)	Cholinergic synapse
CHMP2A	–	1 Dmis	1 PTV, 1 Dmis	5.95E-02	0.91 (17.47%)	Vesicle-mediated transport in synapse
TSPAN4	–	2 PTV, 3 Dmis	1 PTV, 1 Dmis	7.55E-02	2.52E-06 (80.76%)	Regulation of cell development, activation, growth, and motility.
SPAG9	–	2 PTV, 3 Dmis	1 PTV	7.73E-02	1.00 (1.67%)	Nervous system development and vesicle-mediated transport in synapse
ATP13A4	1 PTV	2 PTV, 3 Dmis	1 Dmis	7.89E-02	6.08E-10 (91.54%)	Calcium ion homeostasis
LRP2	–	4 PTV, 22 Dmis	3 PTV, 3 Dmis	8.25E-02	1.00 (0.23%)	Nervous system development and synapse function

Table 3
Summary of 194 identified NDD candidate genes.

Class	known NDD candidate genes ($n = 168$*)	Novel NDD candidate genes ($n = 26$)	
$FDR_{TADA\_combined} < 0.01$ (Strong, $n = 137$)	129	8 genes: *LPHN2*, *GALNT18*, *OR10Z1*, *UBR4*, *FRYL*, *ITSN1*, *BIRC6,* and *UBR3*	
$0.01 \leq FDR_{TADA\_combined} < 0.05$ (Suggested, $n = 24$)	16	8 genes: *MYO7B*, *TECTA*, *PINK1*, *CADM1*, *ALS2*, *LRRC4*, *RFX7,* and *CACNA1S*	
$0.05 \leq FDR_{TADA\_combined} < 0.1$ (Positive, $n = 24$)	17	7 genes: *RYR3*, *ATP13A4*, *MSL2*, *CHMP2A*, *TSPAN4*, *SPAG9,* and *LRP2*	
X-linked genes ($n = 13$)	10	3 genes: *CACNA1F*, *FLNA,* and *PLXNA3*	
Chinese proband with DNMs ($n = 2704$)	111 (4.11%)	11 (0.41%)	Subtotal: 122 (4.51%)
Chinese proband with X-linked hemizygous variants ($n = 2360$)	16 (0.68%)	6 (0.25%)	Subtotal: 22 (0.93%)
Chinese proband with inherited or state unknown variants ($n = 3195$)	617 (19.31%)	248 (7.76%)	Subtotal: 865 (27.07%)

TADA, Transmission and *De novo* Association; The TADA combines DNMs in PTV and Dmis classes with proper weightings from our Chinese cohorts and public NDD cohorts to prioritize candidate genes. *, four genes (*MECP2*, *SLC35A2*, *PTCHD1,* and *KDM5C*) were both X-linked and significantly *de novo* mutated.

CADM1, *ITSN1*, *RYR3*, *RFX7*, *BIRC6, PLXNA3, FLNA, UBR3*, and *LPHN2*) were highly constrained (pLI $\geq$ 0.9).

The 26 novel candidates and the 319 known NDD candidates determined to interact with $\geq$2 novel NDD candidates were used for the development of a functional network based on co-expression and PPI connections (Fig. 2). As a result, we obtained three convergent functional networks involved in chromatin organization ($n_{novel} = 4$ and $n_{known} = 46$), nervous system development ($n_{novel} = 9$ and $n_{known} = 73$), and synaptic function ($n_{novel} = 13$ and $n_{known} = 67$) (Table S15), which are known networks involved in the pathogenicity of ASD (de Rubeis et al., 2014; Chang et al., 2015). These ASD-associated networks comprised 17 of the novel candidate genes (*FLNA*, *ALS2*, *ITSN1*, *PINK1*, *PLXNA3*, *CACNA1S*, *SPAG9*, *CACNA1F*, *CADM1*, *CHMP2A*, *LRP2*, *LRRC4*, *MSL2*, *MYO7B*, *RYR3*, *UBR3*, and *UBR4*) and 114 of the known NDD candidate genes (Fig. 2). In addition, the three networks were connected by certain hub genes, such as *FLNA* (in all of them), and *ALS2*, *ITSN1*, *PINK1*, *SPAG9*, *LRP2*, *CACNA1S*, and *PLXNA3* (in two of them, Fig. 2), suggesting that the heterogeneity of clinical symptoms of NDDs might be caused by these genes that are involved in multiple functional networks.

Discussion

Recently, the falling costs of sequencing have allowed the prioritization of dozens of definitive NDD candidate genes with an excess of putative functional DNMs (de Rubeis et al., 2014; Iossifov et al., 2014), which improved our understanding of their neurobiological and genetic basis. However, the complete pathogenic spectrum behind most NDDs remains unclear, especially in Chinese NDD probands. Compared with published genetic studies of NDDs in patients of Chinese origin (Wang et al., 2016; Li et al., 2017a; Guo et al., 2018; Zhou et al., 2019), our study, for the first time, comprehensively investigated genetic spectrum in a much larger Chinese cohort from aspects of DNM and rare inherited dominant, recessive, and X-linked inheritance. Via the characterization of putative functional DNMs, 86 genes with a total of 137 DNMs, including 16 recurrent genes, were detected in 132 probands from our Chinese

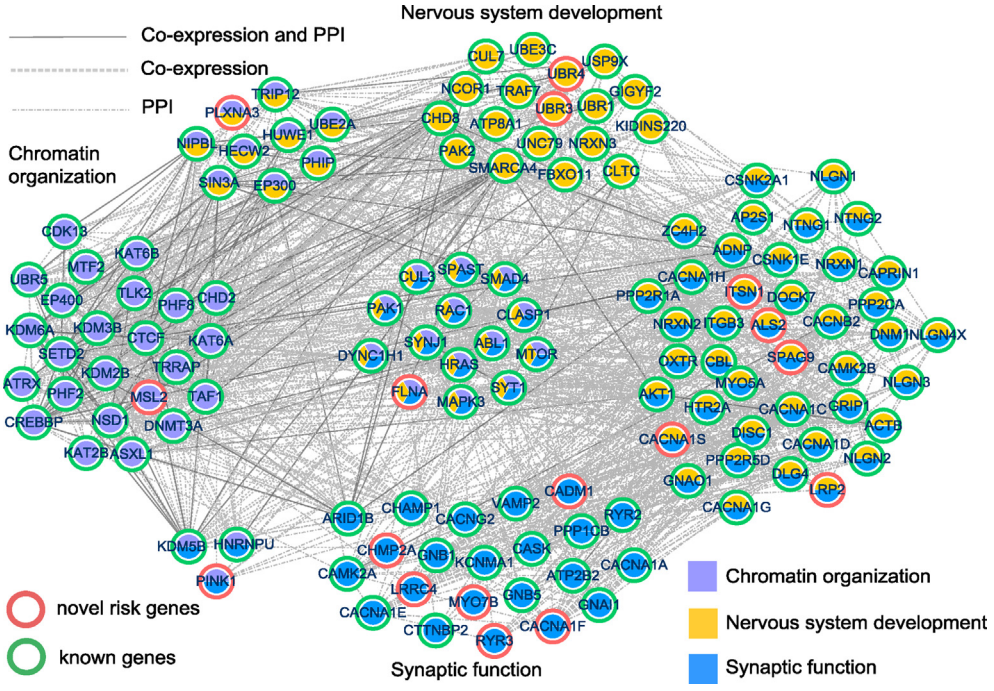

Fig. 2. Co-expression and PPI functional networks. Novel candidate genes from a biological network involved in chromatin organization, nervous system development, and synaptic function. Only genes implicated in these three functional networks are displayed. Biological process of GO is used to report the biological pathway. Novel candidate genes are co-expressed and/or interacted with each other, as well as with known NDD candidate genes.

cohort with neurodevelopmental phenotypes. Notably, 48 out of the 86 genes carrying DNMs in our cohort were first reported in Chinese patients with NDDs (Table S6). We also found that 117 out of 137 DNMs occurred in our Chinese rather than Western NDDs cohorts (Coe et al., 2019; Heyne et al., 2019; Du et al., 2020; Wang et al., 2020) and 91 of which were undetected in both, Chinese and Western, normal controls (Table S16). Interestingly, duplicate DNMs in three genes, *SHANK3* (frameshift deletion, c.4023_4024del, p.V1343Gfs*3), *SYNE1* (missense, c.10532T>C, p.F3511S), and *TANC2* (frameshift insertion, c.1730dupT, p.T578Nfs*2) were only observed from our Chinese NDD probands, respectively, indicating that these recurrent DNMs might be hotspots associated with Chinese NDDs. Meanwhile, even using 519 genes in our panel, we found that 21.31% ($n = 681$) of the 3195 Chinese probands with NDDs carry ≥ 2 putative functional mutations, supporting our previous notion in which NDDs were in favor of the oligogenic model (Du et al., 2020).

In our Chinese cohort, we also identified 13 probands with putative functional recessive mutations in 12 genes, including two genes *CTNND2* and *MACROD2,* with one homozygous mutation in each and ten genes with 23 compound heterozygous mutations (Table S17), which supports the previous conclusion that functional recessive mutations contribute to the etiology of NDDs (Doan et al., 2019). Of those, only *NEB* and *DNAH5* were showed putative functional compound heterozygous mutations in the Simons Simplex Collection (SSC) cohort (Fischbach and Lord, 2010). Two of the remaining ten genes (*TECAT* and *LRP2*) showed no obvious association with NDD pathogenesis. Interestingly, *TECAT* that carries putative functional compound heterozygous mutations in two unrelated probands, might be a novel gene specifically contributing to Chinese NDD etiology in recessive inheritance pattern.

Furthermore, by integrating the publicly available NDD DNM data with that of our Chinese probands, we prioritized 194 candidate genes, of which 168 are known NDD candidates ($P = 2.09 \times 10^{-31}$). Indeed, this convergent analysis highlighted three evidences supporting the pathogenicity of the remaining 26 novel candidate genes. First, the novel candidate genes showed a similar probability of genic intolerance with the known NDD candidate genes, which is significantly higher than the background gene set. Second, the novel candidate genes were significantly co-expressed and/or interacted with the known NDD candidate genes. Finally, the novel candidate genes and the interacting known candidate genes were significantly enriched in ASD-associated networks, such as the chromatin organization, nervous system development, and synaptic function networks. Taken together, the identified novel candidates are putative pathogenic genes that may contribute to the etiology of NDDs, although further functional validation studies are warranted.

Interestingly, 53.85% ($n = 14$) of the 26 novel candidate genes were highly constrained (pLI ≥ 0.9) (Jakob et al., 2017), and some carried putative functional DNMs for multiple types of NDDs (Fig. 3A and 3B). Additionally, 12 out of these 14 genes (*UBR4*, *CADM1*, *ITSN1*, *RYR3*, *SPAG9*, *RFX7*, *BIRC6, PLXNA3, FLNA, UBR3*, *LPHN2*, and *FRYL*) showed a differential gene expression in the brain cortex of patients with ASD and/or SCZ, were highly co-expressed and/or interacted together with validated ASD genes (Fig. 3C and 3D); six of these 12 genes (*ITSN1*, *UBR3, CADM1, RYR3*, *FLNA*, and *PLXNA3*) met the above two criteria. For example, intersection 1 (*ITSN1*), with the loss of function DNMs, was statistically significant (FDR $= 1.89 \times 10^{-3}$) compared with that in previous studies (FDR > 0.1). *ITSN1* is involved in two ASD-associated networks of chromatin organization and the development of the nervous system. In addition, a previous study has demonstrated that the loss of the *ITSN1* signaling in mice can inhibit the stimulation function of the Reelin pathway, thereby reducing neuronal migration and synaptic plasticity (Jakob et al., 2017). Meanwhile, given that *ITSN1* was highly co-expressed with the ASD-related gene *TBR1* (Notwell et al., 2016; Fazel Darbandi et al., 2018) and showed a significant differential expression in ASD cortex as compared with that in control, we speculate *ITSN1* to be a pathogenic gene in ASD. The actin cross-linking protein filamin A (*FLNA*) carried three Dmis hemizygous variants in three unrelated male probands, one of which was reported in our previous study (Li et al., 2017a). Network analysis showed that *FLNA* acts as a hub gene in chromatin organization, nervous system development, and synapse function networks. Interestingly, *FLNA* knockdown in the cortex of mice with an epilepsy phenotype reduced seizure activity (Zhang et al., 2020). Moreover, our previous study on both humans and mouse models found that the *FLNA*-interacting gene, *PAK2* is an ASD pathogenic gene (Wang et al., 2018). Therefore, *FLNA* might be a causal gene to the multiple NDD phenotypes, especially for ASD and EE.

As expected, a higher percentage of our Chinese patients with NDDs (~4.9%, 132/2074) carried putative functional DNMs in 519 targeted genes compared with that in Chinese patients with ASD from a previous study, with ~3.4% (34/1045) of patients in 189 targeted genes carrying DNMs (Wang et al., 2016), indicating that the integration analysis of similar disorders might be a good choice for risk gene discovery in NDDs.

This study has some limitations. First, for approximately 15.36% ($n = 491$) of the probands in this study, the blood samples from the progenitors (father and/or mother) were not available; hence the state of inheritance for ~20.38% of the variants (514/2522) remains undetermined. Second, the candidate NDD genes identified in this study were primarily based on predictions; therefore, further *in vitro* or *in vivo* investigations are warranted to confirm their contribution to the disease etiology.

In summary, using the large-scale sequencing of NDD-associated genes in Chinese probands with neurodevelopmental phenotypes, we have identified recurrent and new candidate genes associated with NDDs, some of which were first reported in Chinese individuals. Moreover, convergent analysis of the data obtained from our cohort, and that of publicly available on different NDD cohorts, identified 26 novel NDD candidate genes. Finally, we showed that six novel candidate genes, namely *ITSN1*, *UBR3, CADM1, RYR3, FLNA*, and *PLXNA3*, may preferably contribute to the development of ASDs due to the involvement of ASD-associated networks and significant differential expression in the ASD cortex when compared with those in normal controls (Fig. 3). Overall, our results showed novel high-confidence candidate genes that warrant further investigation.

Materials and methods

Dataset

Individual variant data used in this study were obtained from our Chinese NDD target-sequencing data and published NDD data, including autism spectrum disorder (ASD), schizophrenia (SCZ), intellectual developmental disorder (ID), epileptic encephalopathy (EE), congenital heart defects (CHD), undiagnosed developmental disorders (UDD) and controls were retrieved from Gene4Denovo database (version 1.0) (Zhao et al., 2020). For our Chinese NDD cohort, whole-blood DNA from 2704 unrelated trios (probands and their unaffected parents) and 491 probands (235 probands without both father' and mother' blood, 220 probands only without father' blood, and 36 probands only without mother' blood, totally 8859 samples) were recruited in mainland China. In reference to DSM-IV, all probands showed neurodevelopmental phenotypes. Genomic DNA (1 μg) from 3195 probands (male/female, 2360/835) was used to construct a DNA library for targeted sequencing as in a previous study (Li et al., 2017a). This study was approved by the DIAGenes Precision

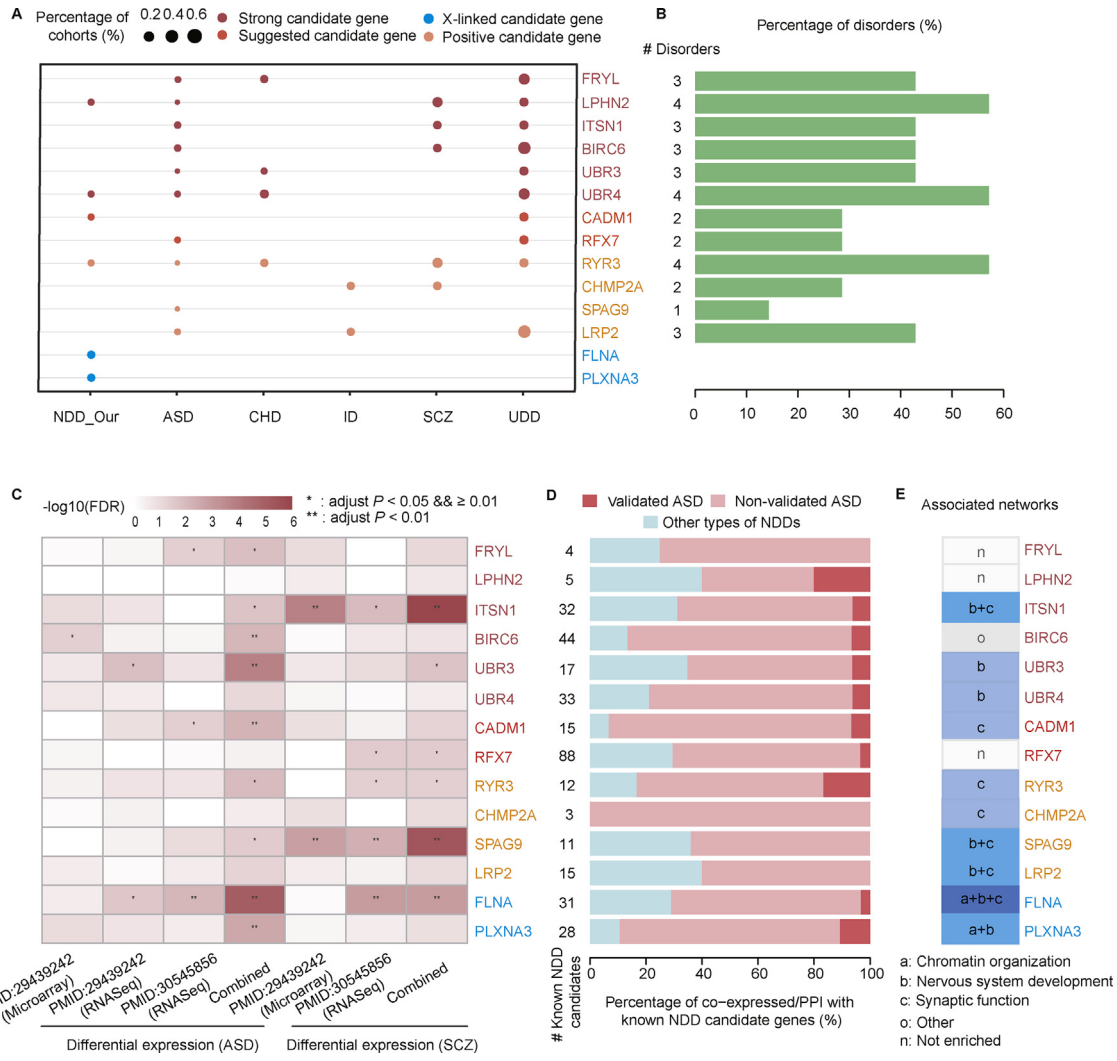

Fig. 3. Summary of 14 high confidence novel NDD candidates with pLI $\geq$ 0.9. **A:** The covered range of NDD disorder types by the novel NDD candidate genes with putative functional mutations. The size of the dots represents the percentage of all cohorts of the NDD disorder type in which the gene is identified with putative functional mutation. **B:** The number and percentage of disorder types in which the gene has putative functional mutations in all disorder types. In total, seven cohorts, while EE cohort is missing putative functional mutations in genes listed in (A) is represented in the bars. **C:** Differential expression results with *P* value and adjust *P* from public ASD and SCZ cortex transcriptome analysis were integrated. Then, meta-analysis was performed by combining the *P* values from all public transcriptome studies in ASD or SCZ for each of the reference genes using the Fisher's method, and further combined *P* values for all reference genes were corrected by the Benjamini-Hochberg method (FDR). Novel NDD candidate genes with an FDR < 0.05 were considered as significantly differential expression. **D:** The percentage of three classes of co-expressed/PPI interacted known NDD candidate genes in all co-expressed/PPI interacted known NDD candidate genes. **E:** The associated networks for each of the novel NDD candidates.

Medicine, Beijing, China, and Beijing Institutes of Life Science, Chinese Academy of Sciences, China. All subjects who participated in this study completed informed consent before the original sample collection. For the published NDD data, all detailed information has been described in the previous individual study. Notably, the majority of the NDD samples were family-based, excepting our in-house control cohort and 491 sporadic NDD cases from our Chinese NDD cohort.

Target-sequencing

We manually collected 519 NDD-related genes (Table S2) for targeted sequencing. Most of them were selected according to the following criteria: (1) they were reported to carry recurrent DNMs in large cohort studies of NDD patients; (2) they act as putative candidates for NDDs identified by dominant, recessive, or X-linked inheritance patterns based on the six databases: OMIM, SFARI, AutismKB, Developmental Disorders (DDG2P)s, Disease and gene associations (DisGeNET) and Human Phenotype Ontology (HPO); (3) they were involved in NDD-related functional pathways/GO terms. The remaining dozens of genes are our interesting genes, which were selected for their possible contribution to NDD etiology based on our experience.

Variant detection and annotation, and damaging variant selection

In quality control, Cutadapt and FastQC were employed to remove adapter sequences and unqualified sequences, respectively. The clean reads were aligned to the human reference genome (hg19) using BWA-MEM (Li and Durbin, 2010). Samtools (Li et al., 2009) was used to mark duplicate reads and generate position-sorted files. The

Genome Analysis Toolkit HaplotypeCaller was used to detect variants. Comprehensive annotation of all variants was performed by ANNOVAR (Wang et al., 2010), including functional implications (gene region, functional effect, mRNA GeneBank accession number, amino acid change, cytoband, etc.), functional predictions for missense variants and allele frequencies of gnomAD, ExAC and in-house control data ($n = 3582$). Deleterious missense (Dmis) variants were predicted by ReVe (Li et al., 2018) with a score of more than 0.7. Only rare protein-truncating variants (PTVs, including stop-gain-, stop-loss-, frameshift-, splicing-derived variants) and Dmis variants with minor-allele frequency less than 0.1% were included in subsequent analysis. Both PTVs and Dmis variants were defined as putative functional variants. All variants and their state of inheritance were validated by Sanger sequencing. The same criterion for variant selection was applied for the published NDD data and control data.

Candidate genes prioritization

In this study, we employed TADA (He et al., 2013) (version 1.2) to prioritize candidate genes, which integrated the following genetic data of 519 targeted genes: (1) putative functional DNMs from 2704 trios; (2) PTVs, which are rare inherited or whose inheritance unknown from 3195 probands; (3) PTVs from 3582 inhouse Chinese controls, and (4) putative functional DNMs from 3391 control from Gene4Denovo database (version 1.0) (Zhao et al., 2020), which involved in background DNM rate estimation for PTV and Dmis category for each gene. In addition, to increase the statistic power, putative functional DNMs of 16,807 probands with different types of neurodevelopmental disorders, including autism spectrum disorder (ASD), schizophrenia (SCZ), intellectual developmental disorder (ID), epileptic encephalopathy (EE), congenital heart defects (CHD) and undiagnosed developmental disorders (UDD) were integrated with prioritizing candidate genes using TADA. Genes with FDR < 0.1 were defined as candidates.

Identification of novel candidate genes

To compare candidate genes identified in our study, we collected known NDD candidates, which are statistically significant genes identified by 11 recent publications on NDDs without considering thresholds or alternative methods of significance (DNM and rare inherited variation) (Set 1) and two well-curated databases (Set 2): OMIM and SFARI, similar methods used in Evan E Eichler's NDD study (Coe et al., 2019). Set 1: candidate genes defined by the 11 publications, including six for ASD (de Rubeis et al., 2014; Sanders et al., 2015; Stessman et al., 2017; Yuen et al., 2017; Takata et al., 2018; Satterstrom et al., 2020), one for ID (Lelieveld et al., 2016), one for UDD (Deciphering Developmental Disorders Study, 2017), one for CHD (Jin et al., 2017), and two for integrated NDD (Nguyen et al., 2017; Coe et al., 2019), which involved large-scale exome, whole-genome, or targeted sequencing. Set 2: candidate genes collected from two well-curated databases: OMIM, which with an annotated phenotype of intellectual developmental disorder (ID), autism, mental retardation, epileptic encephalopathy/Epilepsy (EE), schizophrenia (SCZ), congenital heart defects (CHD) or neurodevelopmental disorder (NDD) and SFARI Gene database, which annotated as syndromic gene or score category of 1 or 2. The novel candidate genes are absent from the known NDD candidate gene sets mentioned above and have no obvious genetic evidence to be associated with neurodevelopmental disorders in PubMed. Then, the known NDD candidate genes were further classified into seven categories: Shared NDD candidate - for which is directly defined as NDD candidate (integration analysis and/or OMIM annotation) or defined as a candidate for at least two types of NDDs; ASD candidate only; EE candidate only; ID candidate only; SCZ candidate only; CHD candidate only and mental retardation candidate only.

Enrichment analysis

Biological pathways (denoted as Gene Ontology terms and KEGG pathways) over-represented among interested genes were identified using the MetaScape tool (Zhou et al., 2019) with default cutoffs. Cell-specific enrichment analysis (CSEA) for the 194 NDD candidate genes was performed using the CSEA tool (Xu et al., 2014) with mouse expression data.

Permutation test

The spatial and temporal expression data of the human brain were downloaded from BrainSpan (http://www.brainspan. org/). Human PPI networks were retrieved from the STRING database (http://string.embl.de/). $|R| > 0.8$ was used to select co-expressed gene pairs in the human brain as previous study (Gulsuner et al., 2013). For the Human PPI networks, gene pairs with a score of >400 were considered as connected and used for downstream analysis. The permutation test was performed between the novel candidate gene set and collected known candidate gene set to evaluate their functional connections. In brief, we compared the number of co-expressed and connected genes within the novel candidate gene set and collected known candidate gene set and their connections with a random gene set of 1,000,000 random iterations. The random gene set was retrieved from all reference genes, with the same size as the novel candidate gene set, and the connections between genes from the random gene set and collected known candidate gene set were treated as background.

Network analysis

We constructed a network according to gene pairs selected from PPI and co-expression with the same criteria used in the permutation test. Novel candidate genes were defined as 'seed genes' that were directly connected and used to form an interconnected functional network. Known NDD genes directly connected at least two novel candidate genes were added to the above networks. To further investigate the functional pathways of novel candidate genes, we performed gene enrichment analysis using the Metascape tool (https://metascape.org). Network figures were generated with Cytoscape v.3.7.2.

Differential expression integration analysis for published brain transcriptome on NDDs

The publicly available brain transcriptome data of ASD and SCZ were involved in this study (Gandal et al., 2018a, 2018b). The differential expression results for all reference genes in ASD and SCZ brain studies were downloaded from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5898828/bin/NIHMS955683-supplement-DataTableS1.xlsx and https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6443102/bin/NIHMS1007359-supplement-Table_S1.xlsx. The P values and adjust P values for all reference genes in ASD and SCZ differential expression results were retrieved and integrated into our study. Then, a combined P value was further calculated by combining the P values of differential expression for each gene forming different datasets in ASD and SCZ, respectively, using Fisher's method. Furthermore, all combined P values for reference genes were corrected by the Benjamini-Hochberg method, and

genes with adjusted P value ≤ 0.05 were considered as significantly differential expressions.

CRediT authorship contribution statement

Tao Wang: Investigation, Methodology, Analyzation, Writing-original draft, Writing-review & editing. **Yi Zhang:** Methodology & Analyzation. **Liqui Liu:** Data Collection & Analyzation. **Yan Wang:** Writing-review & editing, Funding acquisition. **Huiqian Chen:** Writing-review & editing. **Tianda Fan:** Data Collection. **Jinchen Li:** Investigation, Methodology, Provision of in-house control data, Funding acquisition, Writing-review & editing. **Kun Xia:** Conceptualization, Funding acquisition, Writing-review & editing. **Zhongsheng Sun:** Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing-review & editing.

Acknowledgments

This work was supported by the Guangdong Key Project in "Development of new tools for diagnosis and treatment of Autism" (2018B030335001 to Z. Sun) and "Early diagnosis and treatment of autism spectrum disorders" (202007030002 to Z. Sun), the National Natural Science Foundation of China (32070590 to Y. Wang), the National Natural Science Foundation of China (81730036 and 81525007 to K. Xia), Science and Technology Major Project of Hunan Provincial Science and Technology Department (2018SK1030 to K. Xia), the National Natural Science Foundation of China (81801133 to J. Li), the Young Elite Scientist Sponsorship Program by CAST (2018QNRC001 to J. Li), the Innovation-Driven Project of Central South University (20180033040004 to J. Li) and Natural Science Foundation of Hunan Province for outstanding Young Scholars (2020JJ3059 to J. Li).

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jgg.2021.03.002.

References

Al-Mubarak, B., Abouelhoda, M., Omar, A., AlDhalaan, H., Aldosari, M., Nester, M., Alshamrani, H.A., El-Kalioby, M., Goljan, E., Albar, R., et al., 2017. Whole exome sequencing reveals inherited and de novo variants in autism spectrum disorder: a trio study from Saudi families. Sci. Rep. 7, 5679.

An, J.Y., Lin, K., Zhu, L., Werling, D.M., Dong, S., Brand, H., Wang, H.Z., Zhao, X., Schwartz, G.B., Collins, R.L., et al., 2018. Genome-wide de novo risk score implicates promoter variation in autism spectrum disorder. Science 362, eaat6576.

Ben-David, E., Shifman, S., 2013. Combined analysis of exome sequencing points toward a major role for transcription regulation during brain development in autism. Mol. Psychiatr. 18, 1054–1056.

Breuillard, D., Leunen, D., Chemaly, N., Auclair, L., Pinard, J.M., Kaminska, A., Desguerre, I., Ouss, L., Nabbout, R., 2016. Autism spectrum disorder phenotype and intellectual disability in females with epilepsy and PCDH-19 mutations. Epilepsy Behav. 60, 75–80.

Chang, J., Gilman, S.R., Chiang, A.H., Sanders, S.J., Vitkup, D., 2015. Genotype to phenotype relationships in autism spectrum disorders. Nat. Neurosci. 18, 191–198.

Christensen, D.L., Bilder, D.A., Zahorodny, W., Pettygrove, S., Durkin, M.S., Fitzgerald, R.T., Rice, C., Kurzius-Spencer, M., Baio, J., Yeargin-Allsopp, M., 2016. Prevalence and characteristics of autism spectrum disorder among 4-year-old children in the autism and developmental disabilities monitoring network. J. Dev. Behav. Pediatr. 37, 1–8.

Coe, B.P., Stessman, H.A.F., Sulovari, A., Geisheker, M.R., Bakken, T.E., Lake, A.M., Dougherty, J.D., Lein, E.S., Hormozdiari, F., Bernier, R.A., et al., 2019. Neurodevelopmental disease genes implicated by de novo mutation and copy number variation morbidity. Nat. Genet. 51, 106–116.

de Crescenzo, F., Postorino, V., Siracusano, M., Riccioni, A., Armando, M., Curatolo, P., Mazzone, L., 2019. Autistic symptoms in schizophrenia spectrum disorders: a systematic review and meta-analysis. Front. Psychiatr. 10, 78.

de Rubeis, S., He, X., Goldberg, A.P., Poultney, C.S., Samocha, K., Cicek, A.E., Kou, Y., Liu, L., Fromer, M., Walker, S., et al., 2014. Synaptic, transcriptional and chromatin genes disrupted in autism. Nature 515, 209–215.

Deciphering Developmental Disorders Study, 2017. Prevalence and architecture of de novo mutations in developmental disorders. Nature 542, 433–438.

Doan, R.N., Lim, E.T., de Rubeis, S., Betancur, C., Cutler, D.J., Chiocchetti, A.G., Overman, L.M., Soucy, A., Goetze, S., Autism Sequencing, C., et al., 2019. Recessive gene disruptions in autism spectrum disorder. Nat. Genet. 51, 1092–1098.

Du, Y., Li, Z., Liu, Z., Zhang, N., Wang, R., Li, F., Zhang, T., Jiang, Y., Zhi, X., Wang, Z., et al., 2020. Nonrandom occurrence of multiple de novo coding variants in a proband indicates the existence of an oligogenic model in autism. Genet. Med. 22, 170–180.

Epilepsy Phenome/Genome Project, Allen, A.S., Berkovic, S.F., Cossette, P., Delanty, N., Dlugos, D., Eichler, E.E., Epstein, M.P., Glauser, T., et al., 2013. De novo mutations in epileptic encephalopathies. Nature 501, 217–221.

Fazel Darbandi, S., Robinson Schwartz, S.E., Qi, Q., Catta-Preta, R., Pai, E.L., Mandell, J.D., Everitt, A., Rubin, A., Krasnoff, R.A., Katzman, S., et al., 2018. Neonatal Tbr1 dosage controls cortical layer 6 connectivity. Neuron 100, 831–845 e7.

Fernandez, B.A., Scherer, S.W., 2017. Syndromic autism spectrum disorders: moving from a clinically defined to a molecularly defined approach. Dialogues Clin. Neurosci. 19, 353–371.

Fischbach, G.D., Lord, C., 2010. The Simons Simplex Collection: a resource for identification of autism genetic risk factors. Neuron 68, 192–195.

Gandal, M.J., Haney, J.R., Parikshak, N.N., Leppa, V., Ramaswami, G., Hartl, C., Schork, A.J., Appadurai, V., Buil, A., Werge, T.M., et al., 2018a. Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. Science 359, 693–697.

Gandal, M.J., Zhang, P., Hadjimichael, E., Walker, R.L., Chen, C., Liu, S., Won, H., van Bakel, H., Varghese, M., Wang, Y., et al., 2018b. Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. Science 362, eaat8127.

Gulsuner, S., Walsh, T., Watts, A.C., Lee, M.K., Thornton, A.M., Casadei, S., Rippey, C., Shahin, H., Consortium on the Genetics of, S, Group, P.S., et al., 2013. Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. Cell 154, 518–529.

Guo, H., Duyzend, M.H., Coe, B.P., Baker, C., Hoekzema, K., Gerdts, J., Turner, T.N., Zody, M.C., Beighley, J.S., Murali, S.C., et al., 2019. Genome sequencing identifies multiple deleterious variants in autism patients with more severe phenotypes. Genet. Med. 21, 1611–1620.

Guo, H., Wang, T., Wu, H., Long, M., Coe, B.P., Li, H., Xun, G., Ou, J., Chen, B., Duan, G., et al., 2018. Inherited and multiple de novo mutations in autism/developmental delay risk genes suggest a multifactorial model. Mol. Autism. 9, 64.

Hamdan, F.F., Srour, M., Capo-Chichi, J.M., Daoud, H., Nassif, C., Patry, L., Massicotte, C., Ambalavanan, A., Spiegelman, D., Diallo, O., et al., 2014. De novo mutations in moderate or severe intellectual disability. PLoS Genet. 10, e1004772.

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., et al., 2013. Integrated model of de novo and inherited genetic variants yields greater power to identify risk genes. PLoS Genet. 9, e1003671.

Hegarty, J.J., Lazzeroni, L.C., Raman, M.M., Hallmayer, J.F., Cleveland, S.C., Wolke, O.N., Phillips, J.M., Reiss, A.L., Hardan, A.Y., 2019. Genetic and environmental influences on corticostriatal circuits in twins with autism. J. Psychiatry Neurosci. 44, 190030.

Heyne, H.O., Artomov, M., Battke, F., Bianchini, C., Smith, D.R., Liebmann, N., Tadigotla, V., Stanley, C.M., Lal, D., Rehm, H., et al., 2019. Targeted gene sequencing in 6,994 individuals with neurodevelopmental disorder with epilepsy. Genet. Med. 21, 2496–2503.

Hoischen, A., Krumm, N., Eichler, E.E., 2014. Prioritization of neurodevelopmental disease genes by discovery of new mutations. Nat. Neurosci. 17, 764–772.

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., et al., 2014. The contribution of de novo coding mutations to autism spectrum disorder. Nature 515, 216–221.

Iossifov, I., Ronemus, M., Levy, D., Wang, Z., Hakker, I., Rosenbaum, J., Yamrom, B., Lee, Y.H., Narzisi, G., Leotta, A., et al., 2012. De novo gene disruptions in children on the autistic spectrum. Neuron 74, 285–299.

Jakob, B., Kochlamazashvili, G., Japel, M., Gauhar, A., Bock, H.H., Maritzen, T., Haucke, V., 2017. Intersectin 1 is a component of the Reelin pathway to regulate neuronal migration and synaptic plasticity in the hippocampus. Proc. Natl. Acad. Sci. U. S. A. 114, 5533–5538.

Ji, Y., Azuine, R.E., Zhang, Y., Hou, W., Hong, X., Wang, G., Riley, A., Pearson, C., Zuckerman, B., Wang, X., 2020. Association of cord plasma biomarkers of in utero acetaminophen exposure with risk of attention-deficit/hyperactivity disorder and autism spectrum disorder in childhood. JAMA Psychiatr. 77, 180–189.

Jiang, Y.H., Yuen, R.K., Jin, X., Wang, M., Chen, N., Wu, X., Ju, J., Mei, J., Shi, Y., He, M., et al., 2013. Detection of clinically relevant genetic variants in autism spectrum disorder by whole-genome sequencing. Am. J. Hum. Genet. 93, 249–263.

Jin, S.C., Homsy, J., Zaidi, S., Lu, Q., Morton, S., DePalma, S.R., Zeng, X., Qi, H., Chang, W., Sierant, M.C., et al., 2017. Contribution of rare inherited and de novo variants in 2,871 congenital heart disease probands. Nat. Genet. 49, 1593–1601.

Kanner, A.M., 2016. Management of psychiatric and neurological comorbidities in epilepsy. Nat. Rev. Neurol. 12, 106–116.

Kearney, H., Byrne, S., Cavalleri, G.L., Delanty, N., 2019. Tackling epilepsy with high-definition precision medicine: a review. JAMA Neurol. 76, 1109–1116.
Krumm, N., Turner, T.N., Baker, C., Vives, L., Mohajeri, K., Witherspoon, K., Raja, A., Coe, B.P., Stessman, H.A., He, Z.X., et al., 2015. Excess of rare, inherited truncating mutations in autism. Nat. Genet. 47, 582–588.
Lek, M., Karczewski, K.J., Minikel, E. v, Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al., 2016. Analysis of protein-coding genetic variation in 60,706 humans. Nature 536, 285–291.
Li, J., Cai, T., Jiang, Y., Chen, H., He, X., Chen, C., Li, X., Shao, Q., Ran, X., Li, Z., et al., 2016. Genes with de novo mutations are shared by four neuropsychiatric disorders discovered from NPdenovo database. Mol. Psychiatr. 21, 298.
Li, H., Durbin, R., 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26, 589–595.
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.
Lelieveld, S.H., Reijnders, M.R., Pfundt, R., Yntema, H.G., Kamsteeg, E.J., de Vries, P., de Vries, B.B., Willemsen, M.H., Kleefstra, T., Lohner, K., et al., 2016. Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. Nat. Neurosci. 19, 1194–1196.
Li, J., Hu, S., Zhang, K., Shi, L., Zhang, Y., Zhao, T., Wang, L., He, X., Xia, K., Liu, C., et al., 2019. A comparative study of the genetic components of three subcategories of autism spectrum disorder. Mol. Psychiatr. 24, 1720–1731.
Li, J., Wang, L., Guo, H., Shi, L., Zhang, K., Tang, M., Hu, S., Dong, S., Liu, Y., Wang, T., et al., 2017a. Targeted sequencing and functional analysis reveal brain-size-related genes and their networks in autism spectrum disorders. Mol. Psychiatr. 22, 1282–1290.
Li, J., Wang, L., Yu, P., Shi, L., Zhang, K., Sun, Z.S., Xia, K., 2017b. Vitamin D-related genes are subjected to significant de novo mutation burdens in autism spectrum disorder. Am. J. Med. Genet. B. Neuropsychiatr. Genet. 174, 568–577.
Li, J., Zhao, T., Zhang, Y., Zhang, K., Shi, L., Chen, Y., Wang, X., Sun, Z., 2018. Performance evaluation of pathogenicity-computation methods for missense variants. Nucleic Acids Res. 46, 7793–7804.
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., et al., 2013. Rare complete knockouts in humans: population distribution and significant role in autism spectrum disorders. Neuron 77, 235–242.
Lim, E.T., Uddin, M., de Rubeis, S., Chan, Y., Kamumbu, A.S., Zhang, X., D'Gama, A.M., Kim, S.N., Hill, R.S., Goldberg, A.P., et al., 2017. Rates, distribution and implications of postzygotic mosaic mutations in autism spectrum disorder. Nat. Neurosci. 20, 1217–1224.
Liu, L., Sabo, A., Neale, B.M., Nagaswamy, U., Stevens, C., Lim, E., Bodea, C.A., Muzny, D., Reid, J.G., Banks, E., et al., 2013. Analysis of rare, exonic variation amongst subjects with autism spectrum disorders and population controls. PLoS Genet. 9, e1003443.
Lugo-Marín, J., Magán-Maganto, M., Rivero-Santana, A., Cuellar-Pompa, L., Alviani, M., Jenaro-Rio, C., Díez, E., Canal-Bedia, R., 2019. Prevalence of psychiatric disorders in adults with autism spectrum disorder: a systematic review and meta-analysis. Res. Autism Spectr. Disord. 59, 22–33.
Matson, J.L., Shoemaker, M., 2009. Intellectual disability and its relationship to autism spectrum disorders. Res. Dev. Disabil. 30, 1107–1114.
Neale, B.M., Kou, Y., Liu, L., Ma'ayan, A., Samocha, K.E., Sabo, A., Lin, C.F., Stevens, C., Wang, L.S., Makarov, V., et al., 2012. Patterns and rates of exonic de novo mutations in autism spectrum disorders. Nature 485, 242–245.
Nguyen, H.T., Bryois, J., Kim, A., Dobbyn, A., Huckins, L.M., Munoz-Manchado, A.B., Ruderfer, D.M., Genovese, G., Fromer, M., Xu, X., et al., 2017. Integrated Bayesian analysis of rare exonic variants to identify risk genes for schizophrenia and neurodevelopmental disorders. Genome Med. 9, 114.
Notwell, J.H., Heavner, W.E., Darbandi, S.F., Katzman, S., McKenna, W.L., Ortiz-Londono, C.F., Tastad, D., Eckler, M.J., Rubenstein, J.L., McConnell, S.K., et al., 2016. TBR1 regulates autism risk genes in the developing neocortex. Genome Res. 26, 1013–1022.
O'Roak, B.J., Vives, L., Girirajan, S., Karakoc, E., Krumm, N., Coe, B.P., Levy, R., Ko, A., Lee, C., Smith, J.D., et al., 2012. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. Nature 485, 246–250.
Petrovski, S., Wang, Q., Heinzen, E.L., Allen, A.S., Goldstein, D.B., 2013. Genic intolerance to functional variation and the interpretation of personal genomes. PLoS Genet. 9, e1003709.
Sabus, A., Feinstein, J., Romani, P., Goldson, E., Blackmer, A., 2019. Management of self-injurious behaviors in children with neurodevelopmental disorders: a pharmacotherapy overview. Pharmacotherapy 39, 645–664.
Sanders, S.J., He, X., Willsey, A.J., Ercan-Sencicek, A.G., Samocha, K.E., Cicek, A.E., Murtha, M.T., Bal, V.H., Bishop, S.L., Dong, S., et al., 2015. Insights into autism spectrum disorder genomic architecture and biology from 71 risk loci. Neuron 87, 1215–1233.
Sanders, S.J., Murtha, M.T., Gupta, A.R., Murdoch, J.D., Raubeson, M.J., Willsey, A.J., Ercan-Sencicek, A.G., DiLullo, N.M., Parikshak, N.N., Stein, J.L., et al., 2012. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. Nature 485, 237–241.
Satterstrom, F.K., Kosmicki, J.A., Wang, J., Breen, M.S., de Rubeis, S., An, J.Y., Peng, M., Collins, R., Grove, J., Klei, L., et al., 2020. Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism. Cell 180, 568–584 e23.
Sierra-Arregui, T., Llorente, J., Gimenez Minguez, P., Tonnesen, J., Penagarikano, O., 2020. Neurobiological mechanisms of autism spectrum disorder and epilepsy, insights from animal models. Neuroscience 445, 69–82.
Stessman, H.A., Xiong, B., Coe, B.P., Wang, T., Hoekzema, K., Fenckova, M., Kvarnung, M., Gerdts, J., Trinh, S., Cosemans, N., et al., 2017. Targeted sequencing identifies 91 neurodevelopmental-disorder risk genes with autism and developmental-disability biases. Nat. Genet. 49, 515–526.
Strasser, L., Downes, M., Kung, J., Cross, J.H., de Haan, M., 2018. Prevalence and risk factors for autism spectrum disorder in epilepsy: a systematic review and meta-analysis. Dev. Med. Child Neurol. 60, 19–29.
Sullivan, P.F., Geschwind, D.H., 2019. Defining the genetic, genomic, cellular, and diagnostic architectures of psychiatric disorders. Cell 177, 162–183.
Takata, A., Miyake, N., Tsurusaki, Y., Fukai, R., Miyatake, S., Koshimizu, E., Kushima, I., Okada, T., Morikawa, M., Uno, Y., et al., 2018. Integrative analyses of de novo mutations provide deeper biological insights into autism spectrum disorder. Cell Rep. 22, 734–747.
Talkowski, M.E., Rosenfeld, J.A., Blumenthal, I., Pillalamarri, V., Chiang, C., Heilbut, A., Ernst, C., Hanscom, C., Rossin, E., Lindgren, A.M., et al., 2012. Sequencing chromosomal abnormalities reveals neurodevelopmental loci that confer risk across diagnostic boundaries. Cell 149, 525–537.
Taylor, J.L., Debost, J.P.G., Morton, S.U., Wigdor, E.M., Heyne, H.O., Lal, D., Howrigan, D.P., Bloemendal, A., Larsen, J.T., Kosmicki, J.A., et al., 2019. Paternal-age-related de novo mutations and risk for five disorders. Nat. Commun. 10, 3043.
Toma, C., Torrico, B., Hervas, A., Valdes-Mas, R., Tristan-Noguero, A., Padillo, V., Maristany, M., Salgado, M., Arenas, C., Puente, X.S., et al., 2014. Exome sequencing in multiplex autism families suggests a major role for heterozygous truncating mutations. Mol. Psychiatr. 19, 784–790.
Wang, T., Guo, H., Xiong, B., Stessman, H.A., Wu, H., Coe, B.P., Turner, T.N., Liu, Y., Zhao, W., Hoekzema, K., et al., 2016. De novo genic mutations among a Chinese autism spectrum disorder cohort. Nat. Commun. 7, 13316.
Wang, T., Hoekzema, K., Vecchio, D., Wu, H., Sulovari, A., Coe, B.P., Gillentine, M.A., Wilfert, A.B., Perez-Jurado, L.A., Kvarnung, M., et al., 2020. Large-scale targeted sequencing identifies risk genes for neurodevelopmental disorders. Nat. Commun. 11, 4932.
Wang, K., Li, M., Hakonarson, H., 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 38, e164.
Tripathi, S., Pohl, M.O., Zhou, Y., Rodriguez-Frandsen, A., Wang, G., Stein, D.A., Moulton, H.M., DeJesus, P., Che, J., Mulder, L.C., et al., 2015. Meta- and orthogonal integration of influenza "OMICs" data defines a role for UBR4 in virus budding. Cell Host Microbe 18, 723–735.
Wang, Y., Du, X., Bin, R., Yu, S., Xia, Z., Zheng, G., Zhong, J., Zhang, Y., Jiang, Y.H., Wang, Y., 2017. Genetic variants identified from epilepsy of unknown etiology in Chinese children by targeted exome sequencing. Sci. Rep. 7, 40319.
Wang, Y., Zeng, C., Li, J., Zhou, Z., Ju, X., Xia, S., Li, Y., Liu, A., Teng, H., Zhang, K., et al., 2018. PAK2 haploinsufficiency results in synaptic cytoskeleton impairment and autism-related behavior. Cell Rep. 24, 2029–2041.
Wu, J., Yu, P., Jin, X., Xu, X., Li, J., Li, Z., Wang, M., Wang, T., Wu, X., Jiang, Y., et al., 2018. Genomic landscapes of Chinese sporadic autism spectrum disorders revealed by whole-genome sequencing. J. Genet. Genomics 45, 527–538.
Xia, L., Ou, J., Li, K., Guo, H., Hu, Z., Bai, T., Zhao, J., Xia, K., Zhang, F., 2020. Genome-wide association analysis of autism identified multiple loci that have been reported as strong signals for neuropsychiatric disorders. Autism Res. 13, 382–396.
Xia, L., Xia, K., Weinberger, D.R., Zhang, F.Y., 2019. Common genetic variants shared among five major psychiatric disorders: a large-scale genome-wide combined analysis. Glob. Clin. Transl. Res. 1, 21–30.
Xu, B., Ionita-Laza, I., Roos, J.L., Boone, B., Woodrick, S., Sun, Y., Levy, S., Gogos, J.A., Karayiorgou, M., 2012. De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. Nat. Genet. 44, 1365–1369.
Xu, B., Roos, J.L., Dexheimer, P., Boone, B., Plummer, B., Levy, S., Gogos, J.A., Karayiorgou, M., 2011. Exome sequencing supports a de novo mutational paradigm for schizophrenia. Nat. Genet. 43, 864–868.
Xu, X., Wells, A.B., O'Brien, D.R., Nehorai, A., Dougherty, J.D., 2014. Cell type-specific expression analysis to identify putative cellular mechanisms for neurogenetic disorders. J. Neurosci. 34, 1420–1431.
Yan, H., Shi, Z., Wu, Y., Xiao, J., Gu, Q., Yang, Y., Li, M., Gao, K., Chen, Y., Yang, X., et al., 2019. Targeted next generation sequencing in 112 Chinese patients with intellectual disability/developmental delay: novel mutations and candidate gene. BMC Med. Genet. 20, 80.
Yuen, C.R.K., Merico, D., Bookman, M., Howe, L.J., Thiruvahindrapuram, B., Patel, R. v, Whitney, J., Deflaux, N., Bingham, J., Wang, Z., et al., 2017. Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. Nat. Neurosci. 20, 602–611.
Yuen, R.K., Merico, D., Cao, H., Pellecchia, G., Alipanahi, B., Thiruvahindrapuram, B., Tong, X., Sun, Y., Cao, D., Zhang, T., et al., 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., et al., 2015. Whole-genome sequencing of quartet families with autism spectrum disorder. Nat. Med. 21, 185–191.
Zhang, L., Huang, T., Teaw, S., Nguyen, L.H., Hsieh, L.S., Gong, X., Burns, L.H., Bordey, A., 2020a. Filamin A inhibition reduces seizure activity in a mouse model of focal cortical malformations. Sci. Transl. Med. 12, eaay0289.

Zhang, Y., Li, N., Li, C., Zhang, Z., Teng, H., Wang, Y., Zhao, T., Shi, L., Zhang, K., Xia, K., et al., 2020b. Genetic evidence of gender difference in autism spectrum disorder supports the female-protective effect. Transl. Psychiatry 10, 4.

Zhao, G., Li, K., Li, B., Wang, Z., Fang, Z., Wang, X., Zhang, Y., Luo, T., Zhou, Q., Wang, L., et al., 2020. Gene4Denovo: an integrated database and analytic platform for de novo mutations in humans. Nucleic Acids Res. 48, D913−D926.

Zheng, Z., Zheng, P., Zou, X., 2018. Association between schizophrenia and autism spectrum disorder: a systematic review and meta-analysis. Autism Res. 11, 1110−1119.

Zhou, H., Xu, X., Yan, W., Zou, X., Wu, L., Luo, X., Li, T., Huang, Y., Guan, H., Chen, X., et al., 2020. Prevalence of autism spectrum disorder in China: a nationwide multi-center population-based study among children aged 6 to 12 years. Neurosci. Bull. 36, 961−971.

Zhou, W.Z., Zhang, J., Li, Z., Lin, X., Li, J., Wang, S., Yang, C., Wu, Q., Ye, A.Y., Wang, M., et al., 2019a. Targeted resequencing of 358 candidate genes for autism spectrum disorder in a Chinese cohort reveals diagnostic potential and genotype-phenotype correlations. Hum. Mutat. 40, 801−815.

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., Chanda, S.K., 2019b. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat. Commun. 10, 1523.