An Intronic Variant of *CHD7* Identified in Autism Patients Interferes With Neuronal Differentiation and Development

Ran Zhang

Institute of Neuroscience Chinese Academy of Sciences

Hui He

Institute of Neuroscience Chinese Academy of Sciences

Bo Yuan

Institute of Neuroscience Chinese Academy of Sciences

Ziyan Wu

Institute of Neuroscience Chinese Academy of Sciences

Xiuzhen Wang

Institute of Neuroscience Chinese Academy of Sciences

Yasong Du

Shanghai Mental Health Center

Yuejun Chen

Institute of Neuroscience Chinese Academy of Sciences

Zilong Qiu (✉ zqiu@ion.ac.cn)

Institute of Neuroscience, Chinese Academy of Sciences <https://orcid.org/0000-0003-4286-3288>

Research

Keywords: Autism, CHD7, Intronic variant, Inherited variant

Posted Date: August 17th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-58118/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Neuroscience Bulletin on May 4th, 2021.
See the published version at <https://doi.org/10.1007/s12264-021-00685-w>.

Abstract

Background: Genetic compositions play critical roles in pathogenesis of autism spectrum disorders (ASD). Inherited and *de novo* intronic variants are often seen in patients with ASD. However, the biological significance of intronic variants are difficult to address. Here, we identified a recurrent inherited intronic variant in the *CHD7* gene among Chinese ASD cohort, which is specifically enriched in East Asian populations. *CHD7* has been implicated in numerous developmental disorders including CHARGE syndrome and ASD. Here, we use differentiated human embryonic stem cells to investigate whether the ASD-associated *CHD7* intronic variant may affect neural development.

Methods: 167 ASD probands and their parents performed whole-exome sequencing and a recurrent inherited intronic variant was selected for functional research. We established human embryonic stem cells carrying this intronic variant using CRISPR/Cas methods and differentiated to neural progenitor cells and neurons for analysis.

Results: The level of *CHD7* mRNA significantly decreased in cell lines carrying homozygous mutants compared to control. Upon differentiation towards the forebrain neuronal lineage, neural cells carrying the *CHD7* intronic variant exhibited developmental delay in neuronal differentiation and maturation. Importantly, *TBR1*, a gene also implicated in ASD, significantly increased in neurons carrying the *CHD7* intronic variant, suggesting the intrinsic relevance among ASD genes. Furthermore, the morphological defects found in neurons carrying *CHD7* intronic mutations could be rescued by knocking down *TBR1*, indicating that *TBR1* may be responsible for defects in *CHD7*-related disorders. Finally, the *CHD7* intronic variant generates three abnormal forms of transcripts through alternative splicing, which all exhibited *loss-of-function* manners in functional assays.

Limitations: The probands with *CHD7* intronic variant were heterozygous while both cell lines established were homozygous, so the phenotype of the disease could not be completely simulated.

Conclusions: Our study provides the crucial evidences for supporting the notion that intronic variant site of *CHD7* is a potentially autism susceptible site, shedding new light on identifying functions of intronic variants in autism genetic studies.

Background

Autism spectrum disorder (ASD) is known as a specific group of neurodevelopmental disorders characterized by impairments in social interaction and stereotyped behaviors (Kanner 1943). Although both genetic and environmental factors may contribute to the autistic symptoms, recent genome-wide analysis indicate that the genetic composition is the dominant cause for ASD (Ronald and Hoekstra 2011, Sandin, Lichtenstein et al. 2014, Tick, Bolton et al. 2016). The genetic architecture of ASD is highly heterogeneous as over a hundred of risk genes have been found (Huguet and Bourgeron 2016, Satterstrom, Kosmicki et al. 2020). Analysis of *de novo* and inherited rare variation linked to ASD has

identified convergent functional themes, such as neuronal development and axon guidance, signaling pathways, and chromatin and transcription regulation (Pinto, Delaby et al. 2014, Bourgeron 2015).

Due to the genetic heterogeneity, one strategy is to perform genome sequencing for a large number of ASD families and identify risk genes more comprehensively. Benefit from the increasing resolution and reducing cost of DNA sequencing technology such as karyotype, microarrays, whole-exome sequencing and whole-genome sequencing, nearly 20% of ASD cases can be explained by identifiable genetic causes, while the rest remained unclear. Till now, most efforts were focused on rare deleterious *de novo* single-nucleotide variants (likely gene-disrupting variants-LGD variants) on coding regions or splicing sites. However, multiple common inherited variants are clearly suspectable factors for ASD (Gaugler, Klei et al. 2014). Since inherited variants are also present in unaffected parents, identifying the causal relationship between inherited variations and ASD pathogenesis would require neurobiological evidences besides genetic evidences (Jiang, Yuen et al. 2013, Guo, Wang et al. 2018).

The current largest genetic sequencing projects for ASD cohorts are from Europe and the United States, which are mostly composed of Caucasians, Latino, Ashkenazi Jewish and African Americans populations etc. Due to the genetic geographical difference between Caucasian and Asians, the potential genetic causes for Asians population may be incomplete overlap with those in the western world. Considering the vast population of China, there is no doubt a large population of ASD in Chinese populations. Identification of the ASD risk genes based on Chinese population not only critical to provide precise genetic diagnosis and specific interventions for Chinese ASD patients, but also provides the important missing piece for a comprehensive and in-depth understanding of autism.

In this study, we performed the whole-exome sequencing for 167 Chinese ASD proband along with their parents. We identified an intronic inherited variation of chromodomain-helicase-DNA-binding protein 7 (CHD7) (ch8-61757392-C-T), which exclusively appears in East Asian populations by 0.39% in the gnomAD database, while enriched by 3.6% in our Chinese ASD cohort (6/167), suggesting that this intronic inherited variant may function as a susceptibility variant for autism. The CHD7 protein belongs to the CHD family of chromatin remodelers and catalyses the translocation of nucleosomes along DNA in chromatin (Petty and Pillus 2013). Mutations of the *CHD7* gene are the major cause of the CHARGE syndrome, which is characterized by coloboma of the eye, heart defects, atresia of the choanae, retardation of growth and development, genital abnormalities, and ear abnormalities (Vissers, van Ravenswaaij et al. 2004). Children with the CHARGE syndrome frequently exhibit the autistic-like deficits on vocalization, social responsiveness, and repetitive behaviors, suggesting that CHD7 has a more direct impact on autism (Hartshorne, Grialou et al. 2005).

To investigate whether this intronic variation may affect the proper expression of the CHD7, we constructed two homozygous point mutant cell lines using human H9 embryonic stem cells (hESCs). We found that the relative mRNA amount of CHD7 is downregulated after introducing of the intronic point mutation in both cell lines. The differentiation of the neural precursor cells (NPCs) carrying intronic mutations appears to delay, comparing to wild-type NPCs. Importantly, when NPCs carrying intronic

mutations differentiated into neurons, the dendrite length and branch number are significantly lower than neurons which do not carrying mutations, which could be rescued by knockdown *TBR1*. Interestingly, we found that the intronic mutation affect alternative splicing of *CHD7* mRNA by introducing three kinds of non-canonical transcripts. We further showed that the three non-canonical transcripts are *loss-of-functions* mutants in mouse culture neurons. In summary, we identify an intronic inherited variation in the *CHD7* gene enrich in the Chinese ASD cohorts. The intronic variation leads to downregulation of *CHD7* mRNA and non-canonical alternative splicing process, affecting processes related to neurodevelopment which could be rescued by knockdown its target gene *TBR1*. This study demonstrates that the intronic inherited variation of *CHD7* is a potential suspectable factor for autism specifically in Chinese ASD cohorts, opening up a new venue for studying the complex genetic causes of autism.

Methods

Subjects

A total of 168 core families with probands diagnosed with ASD were recruited from the outpatient of Department of the Child and Adolescent Psychiatry, Shanghai Mental Health Center. All patients were diagnosed on the basis of the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-V). All patients were Han Chinese and their age was ranged from 2 to 18 years.

Human H9 embryonic stem cells (hESCs) and intronic mutation (8-61757392-C-T) hESCs

hESCs (H9, line WA09 (WiCell), passages 20-40) were cultured on a feeder layer of irradiated mouse embryonic fibroblasts (MEFs) in a humidified incubator with 5%CO₂ at 37°C. The hESC medium containing DMEF/F12, 20% Knock Serum Replacement (Life Technologies), 0.1mM beta-mercaptoethanol (Sigma), 1% (v/v) Non-Essential Amino Acids (NEAA, Life Technologies), 0.5% (v/v) GlutaMAX (Life Technologies). Medium was changed every day with 10ng/ml of bFGF (PeproTech). hESCs were passaged every 7 days. Intronic mutation (8-61757392-C-T) hESCs were transformation from H9 hESCs by introducing a intronic mutation and two synonymous mutations by homology-mediated end joining-based targeted integration using CRISPR/Cas9. The editor vector (addgene #48138) linked with sgRNA targeting sequence and the donor vector homology with gene locus about 1.6kb exclude a intronic mutation and two synonymous mutations and flanked with sgRNA targeting sequence linked in pUC57 were liposome transfection by 1:1 ratio into H9 hESCs. Single GFP⁺ cells were isolated using flow cytometry and cultured. The intronic point mutation cells were identified by DNA sequencing.

HEK293 cells and HEK293 intronic mutation (8-61757392-C-T) cells

HEK293 cells were purchased from Cell Bank of the Chinese Academy of Sciences (Shanghai, China). HEK293 intronic mutation (8-61757392-C-T) cells were transformation from HEK293 cells by introducing a intronic mutation and two synonymous mutations by homology-mediated end joining-based targeted integration using CRISPR/Cas9. The editor vector (addgene #48138) linked with sgRNA targeting sequence and the donor vector homology with gene locus about 1.6kb exclude a intronic mutation and two

synonymous mutations and flanked with sgRNA targeting sequence linked in pUC57 were liposome transfection by 1:1 ratio into HEK293 cells. Single GFP+ cells were isolated using flow cytometry and cultured. The intronic point mutation cells were identified by DNA sequencing. HEK293 synonymous control cells were similar to HEK293 intronic mutation cells with two synonymous mutations but not intronic mutation. All the HEK293 cells were cultured in DMEM (Gibco/Life Technologies) with 10% FBS (Gibco/Life Technologies) at 37 °C in 5% CO2 incubator (Thermo Scientific Heraeus).

Primers for cell genotypes as follows:

CHD7-forward: GAAGTTCACAGGAGCCAGAG

CHD7-reverse: CAGAAAGTAGAATGGTGATTGCCAG

Primary cultures of cortical neurons

Mouse cortical neurons were cultured from E14.5 C57BL/6 J of either sex. Cerebral cortices were dissected, dissociated, and cultured in 0.5 mL/well Neurobasal medium (Gibco, 21103-049) with 2% B27 (Gibco, 17504-044) and 2 mM Glutamax-I (Gibco, 35050-061) on Lab-Tek II Chamber Slide (Thermo Fisher Scientific, 154,941) at 100,000 cells/cm². For the axon and dendrite experiment, the neurons were transfected by Lipo3000 (Invitrogen cat.L3000075) with totally 0.9 µg vector following LipofectamineTM 3000 Reagent Protocol 24 h after planting. After transfection, feed the cultures with new medium every 2 days.

Plasmid construct

The gene editor vector was sgRNA targeting sequence cloned in Cas9 vector (addgene #48138). The donor vector was homologous arm cloned in pUC57. The vector that expressed GFP was FUGW (addgene #14883). The control expression vector was GFP-removed FUGW. The CHD7 expressing vector was a gift from WeiJun Feng, Institutes of Biomedical Sciences Fudan University, Shanghai, China. The CHD7 (del and dup) expressing vector were modified by enzyme ligation and homologous recombination from CHD7 expressing vector. shRNA for mouse *Chd7*, human CHD7 and human TBR1 was cloned into FUGW-H1 vector (addgene #25870). shRNA for control was DsRed.

shRNA sequence as follows:

shRNA for mouse *Chd7*: GCAGCAGCCTCGTTCGTTTAT

shRNA for human CHD7: GCAGCAGTCTCGTCCATTTAT

shRNA for human TBR1: GCCTTTCTCCTTCTATCATGC

shRNA for DsRed: AGTTCCAGTACGGCTCCAA

Whole-exome sequencing

DNA was extracted from peripheral blood of patients and their parents using the DNeasy Blood & Tissue Kit (Cat no. 69506, QIAGEN, GmbH, Germany), following the manufacturer's instructions. For each sample to be sequenced, individual library preparations, hybridizations, and captures were performed following the protocol of Agilent SureSelect capture kit (V5) or IDT XGen Exome Research Panel. Sequencing was performed on an Illumina HiSeq X-10 instrument (Illumina) following the manufacturer's protocol (HiSeq X-10 System User Guide).

Variation identify by Sanger sequencing

Based on the data from WES, all families within probands carrying *CHD7* variations were selected for Sanger sequencing to validate whether variations are de novo or inherited from parents. The primers for Sanger sequencing as follows:

CHD7-forward: CCAGGGTTAGCTTTGTGGGT

CHD7-reverse: TGGCTTTGTGACCCTGTAGC

RNA isolation and reverse transcription

Each group of cells were dissociation in 1 mL Trizol (Invitrogen, 15,596,018). Total RNA was isolated using the method of user guide of TRIzolTM Reagent. Reverse transcribed using the Reverse Transcriptase M-MLV kit (TaKaRa, D2639B). 1 µg total mRNA and 50 nmol oligo dT as primer were used in the reverse transcription.

Quantitative real-time RT-PCR (qPCR)

For qPCR analysis, the gene expression of cDNA sample was analyzed using SYBR green (Toyobo, QPK-201). The qPCR program was three steps with melt as follows: 95 °C denaturation for 10 min, followed by 40 cycles of 95 °C for 10 s, 60 °C for 15 s, and 72 °C for 20 s. RNA level was calculated and standardized by using the Δ Ct method and GAPDH expression level as control.

Primers for qPCR as follows:

CHD7 exon19-forward: ACGAAAAGGGGCCTATGGTG

CHD7 exon20-reverse: TTCAGCCTTCTTAGCCCACT

CHD7 exon25-forward: TCCCTGAACCTTTCCATGCT

CHD7 exon26-reverse: TCCCTGAACCTTTCCATGCT

CHD7 exon35-forward: ATGGCTGAAGCTGCACCCTA

CHD7 exon38-reverse: AGGCGGTCAAACATCGACTC

CHD7 intronic retention-forward: TGGAGAAGAATCTGCTTGTCTATGGG

CHD7 intronic retention-reverse: TCTGGGCTTTCACCTTCTTT

CHD7 exon22-23 deletion-forward: AATCTGCTTGTCTATGGGGTCC

CHD7 exon22-23 deletion-reverse: TCCCTGAACCTTTCCATGCT

CHD7 exon22-23 duplication-forward: CAACCATTCCGGTTTGTTCAGC

CHD7 exon22-23 duplication-reverse: TCTGGGCTTTCACCTTCTTT

TBR1-forward: GACTCAGTTCATCGCCGTCA

TBR1-reverse: TGCTCACGAACTGGTCCTG

GAPDH-forward: CATCGCTCAGACACCATGGG

GAPDH-reverse: CCTTGACGGTGCCATGGAAT

Chd7-forward: TCCACATTTGCTAAGGCCAG

Chd7-reverse: TTCAGCCTTCTTAGCCCACT

Gapdh-forward: GTGAAGGTCGGTGTGAACGG

Gapdh-reverse: CGCTCCTGGAAGATGGTGAT

Differentiation of dorsal forebrain glutamate neurons

HESC colonies were cultured with daily medium change until they reached about 80% confluence. Then, hESC colonies were detached from the feeder layer by digestion with dispase (Life Technologies), and then resuspended in hESC medium for 4 days to form embryoid bodies (EBs). For neural induction, EBs were cultured in neural induction media (DMEM/F12, 1% (v/v) N2 supplement, 5% (v/v) B27 without RA, 1% (v/v) NEAA, all from Life Technologies) (NIM) supplemented with SB-431542 (2uM, Stemgent) and DMH-1 (2uM, Tocris) for 3 days. EBs were then attached to six-well plate in NIM supplemented with 5% fetal bovine serum. The cells were fed with NIM every other day till neural tube-like rosette formation at around day 16. Then, the rosettes were blown off by a 1-ml pipette and cultured in suspension. After two days, the cell clusters would form neurospheres, and then the medium was changed every other day. On day 26, neurospheres were digested into single cell by accutase, and seeded at a density of ~40,000 cells/cm² on coverslips pre-coated with matrigel. After 5-6 hours, neuronal differentiation media (neural basal media, 1%(v/v) N2, BDNF (10ng/ml, PeproTech), GDNF (10ng/ml, PeproTech), cAMP (1uM, Sigma), IGF-I (10ng/ml, PeproTech) and AA (200uM, Sigma)) (NDM) was added to the wells, media was changed weekly.

Immunohistochemical staining

Cultured cells were washed with PBS for 5 mins, then fixed in 4% PFA at room temperature for 30 min. Wash every 10 mins with PBS twice. Block the cells with 5% BSA and 0.3% TritonX-100 in PBS at room temperature for 2 h. Incubate overnight at 4 °C with primary antibody in 3% BSA and 0.1% TritonX-100 in PBS. Wash with PBS every 10 mins for 3 times. Incubate at room temperature with secondary antibody and DAPI in PBS for 2 h. Wash with PBS every 10 mins for 3 times.

Primary antibody used and concentrations as follows:

Anti-SOX2 (R&D, AF2018, 1:500)

Anti-PAX6 (DSHB, AB-528427, 1:10)

Anti-Tuj1 (Sigma, T8660, 1:5000)

Anti-Ki67 (abcam, ab15580, 1:800)

Anti-CHD7 (CST, #6505, 1:1000)

Anti-GFP (abcam, ab6673, 1:400)

Anti-MAP2 (Millipore, MAB3418, 1:1000)

Anti-SMI312 (Biolegend, 837904, 1:1000)

Anti-TBR1 (abcam, ab31940, 1:1000)

Anti-CTIP2 (abcam, ab18465, 1:200)

Hoechst (Life-tech/3570, 1:2000)

DAPI (Sigma, D9542, 1:1000)

Western blot

4-20% SDS-polyacrylamide gradient gel was used in the western blot. Electrophoresis program was 80 V 30mins and then 120 V 180 mins. Proteins were transferred onto the Immobilon polyvinylidene difluoride membrane (Millipore) for 210mins at 200 mA. The membrane was blocked by TBST solution (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, and 0.1% Tween 20) with 5% BSA for 2 h at room temperature and incubated overnight at 4 °C with primary antibody solution in 3% BSA. Wash with TBST every 10 mins for 3 times. The membrane was treated with secondary antibody for 2 h at room temperature. Wash with TBST every 10 mins for 3 times. The reaction was analyzed by using imaging film.

Primary antibody used and concentrations as follows:

Anti-CHD7 (CST, #6505, 1:1000)

Anti-GAPDH (ab8245, 1:5000)

Analysis of dendrites and axons

About 40–50 GFP-positive neurons were picked up randomly from each group. The searcher was blinded until statistical analysis was completed. The images were analyzed using Fiji software according to the standard: all of dendritic branches and secondary branch, the longest axon and secondary branch, and the total length of all neurites were taken into account. At least three independent experiments were performed,

Alternative splicing analysis

cDNA was segmental amplification by PCR. The products were separated by agarose gel electrophoresis and retrieved. Then ligated with pGEM-T Easy Vector. The ligation products were transformation into Escherichia coli Top10 and monoclonal culture. At least 40 monoclonal were sequenced per sample.

Transcriptome analysis—RNAseq

For RNA-sequencing, total RNA was extracted and subsequently a sequencing library was prepared by using the Illumina TrueSeq Total RNA Sample Prep Kit and sequenced on the Illumina Hi-Seq 2000. The clean reads were aligned to 9606—NCBI Taxonomy ID—genome (version: GRCh38) using the Hisat2. We applied HTseq to calculate the counts of the genes. RPKM/FPKM—Reads/Fragments Per Kilobase Million Reads—was used to standardize the expression data. We applied DEseq2 algorithm to filter the differentially expressed genes, then we filtered fold change and FDR under the following criteria: i) $\log_2FC > 0.585$ or < -0.585 ; ii) ,FDR < 0.05. For Gene ontology (GO) analysis, we downloaded the GO annotations from NCBI (<http://www.ncbi.nlm.nih.gov/>), UniProt (<http://www.uniprot.org/>) and the Gene Ontology (<http://www.geneontology.org/>). Fisher's exact test was applied to identify the significant GO categories and FDR was used to correct the p-values.

Statistical analysis

Statistical tests were carried out using GraphPad Prism 6 (Graphpad Software, Inc., RRID:SCR_002798). Two-tailed Student's t-test was used for sample pairs, one-way ANOVAs followed by Tukey's multiple comparison tests was used for 3 or more groups. Data distribution was test by Kolmogorov- Smirnov test and Shapiro-Wilk test by SPSS software (IBM, RRID:SCR_002865). The data distribution was normal distribution. Results were shown as mean $\pm$ SEM, and “n” represented to either the number of neurons (for morphological analysis) or the repeat experiment (for qPCR and RNAseq). Mouse cortical neurons were at least 3 times independently from 3 different litters. Stem cells differentiation were at least 3 batches independently. All data analyses were performed blinded to the experimental condition. All conditions statistically different from control are indicated. *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001.

Exclusion criteria: If the data was not in the 95% confidence interval of the group, the data would be excluded.

Results

An intronic variant of the *CHD7* gene found in Chinese ASD leads to downregulation of CHD7 mRNA

After performing whole-exome sequencing on 167 ASD probands and their parents collected from Shanghai Mental Health Center, we identified an intronic variation in the *CHD7* gene (NM_017780.3:c.4851-31C>T, het) from six probands (Fig. 1A). Variants of five probands were inherited paternally, whereas one variant was inherited maternally and his non-carrier brother was un-affected (Fig. 1B). Sanger sequencing was carried out to confirm the existence of the variant (Fig. 1C).

When we examined the frequency of this variant (rs149348445) in various population on the gnomAD database, we surprisingly found that this variant only existed in East Asia population as a common variant with relatively low frequency (0.39%), instead of Caucasian or other populations. However, this variant exhibits a nearly 10 times enriched frequency (3.6%, 6/167) in our ASD cohorts, strongly suggesting that this variant may be implicated in ASD.

Due to the high variability of non-coding region between rodent and human, the mouse *Chd7* gene does not contain similar sites with which we could make mouse models to mimic. Therefore, to investigate whether this intronic variant may affect the expression of the *CHD7* gene, we set out to perform the mutagenesis in human embryonic stem cells.

The point mutation was introduced to human embryonic stem cells (H9) by homology-mediated end joining-based targeted integration with CRISPR/Cas9 technology (Yao, Wang et al. 2017). Two synonymous mutations were also introduced into the small guiding RNA (sgRNA) target region within exon 22 to avoid the unwanted digestion by Cas9 after recombination occurred (Fig. 1D). Two subclones (hom1-H9-hESCs and hom2-H9-hESCs) were successfully established, which carried the intronic mutations in a homozygous manner (Fig. S1). Unfortunately, heterozygous mutations were failed to establish.

To examine whether point mutation affect the expression of CHD7, we analyzed the mRNA expression using Real-time Quantitative PCR with multiple primer pairs amplifying exons adjacent to the intronic variants site and mRNA terminus. We found that the relative mRNA amount of CHD7 is downregulated in both cell lines carrying homozygous mutants (hom1-H9-hESCs and hom2-H9-hESCs) compared to H9 wild-type cells, using primer pairs amplifying upstream (exon19-20) or downstream exons (exon25-26) of the point mutation, as well as the primer pair amplifying the 3'end of mRNA (exon35-38) (Fig. 1E-1G). These evidences indicated that the intronic variant (ch8-61757392-C-T) in the *CHD7* gene may affect the expression of CHD7 in human cells.

Intronic variation of *CHD7* delays neuronal differentiation of human ESCs

The role of Chd7 previously was found to be implicated in adult neurogenesis and neural differentiation of cerebellar granule cells in mouse (Feng, Khan et al. 2013, Feng, Kawauchi et al. 2017). In order to study whether the intronic variation of *CHD7* found in autistic patients may affect neuronal differentiation, we differentiated hESCs into dorsal forebrain glutamate neurons according to established protocol (Fig. S2A).

At the 28th day (D28) of neural differentiation, we first examine whether proliferation of neural progenitor cells (NPCs) may be affected by the intronic variant of *CHD7*. We performed immunostaining using antibodies against the proliferation marker Ki67 and neural stem cell marker SOX2. We found that the percentage of Ki67 positive cells among SOX2 positive cells are similar between NPCs derived from wild-type H9 cells and two cell lines (hom1-H9-hESCs and hom2-H9-hESCs) carrying homozygous variants (Fig. 2A and 2B), suggesting that the intronic variation has no effect on the proliferation capacity of human NPCs.

In order to further examine the process of neural differentiation, we performed immunostaining for SOX2 (early neural stem cell marker), PAX6 (intermediate neural stem cell marker) and Tuj1 (neuronal marker) at D28 (Fig. 2C). We first found that most of the neuronal cells derived from wild-type hESCs or mutant hESCs expressed dorsal forebrain progenitor marker PAX6 in a similar proportion, suggesting that the intronic mutation of *CHD7* did not affect the dorsal forebrain progenitor differentiation of hESCs (Fig. 2D). Interestingly, there were more SOX2 positive cells in the culture derived from both mutant hESC lines compared with those from wild-type hESCs (Fig. 2E). In contrast, the proportion of Tuj1 positive cells derived from mutant hESCs were much lower than those from wild-type hESCs (Fig. 2F). These results are consistent with the previous report that a lack of Chd7 in mouse neural stem cells caused delayed neural differentiation (Feng, Khan et al. 2013).

Taken together, these evidences strongly suggest that the increase of SOX2-positive cells in hESCs-derived neurons carrying intronic variant is most likely due to the delayed differentiation of forebrain stem cells, rather than increased proliferative capacity.

Since CHD7 plays a critical role in chromatin remodeling, it has been shown in previous work that loss of Chd7 lead to dramatic changes in gene expression (Feng, Kawauchi et al. 2017). To further determinate the molecular mechanism by which CHD7 regulates neural differentiation, we performed RNA-seq with RNA collected from neural precursor cells at D26 after initiation of neural differentiation. We found that expression of numerous genes changed significantly in NPCs carrying *CHD7* intronic variants comparing to wild-type NPCs (Fig. 2G and 2H). Altered genes were involved in the biological processes including dopaminergic neuron differentiation, dentate gyrus development and Wnt signaling pathway (Fig. S2B). Genes which are closely related to neural differentiation and development are of great interest, marking with * (Fig. 2G and 2H). Among them, LMX1A is required for proper ear histogenesis and morphogenesis (Nichols, Pauley et al. 2008), as loss of hearing is an pivotal defect of the CHARGE syndrome. WNT5A, a critical gene in WNT signaling pathway, favors bone marrow MSC differentiation into osteoblasts by inhibits the function of activated PPAR γ through complex formation between NLK, SETDB1, CHD7

(Takada, Kouzmenko et al. 2009). LAMB1 is one of the risk genes for autism spectrum disorder (Takata, Miyake et al. 2018). These evidences indicate that the intronic variant in the *CHD7* gene lead dysregulation of a series of critical gene which implicated in neural developmental disorders, the effect is similar to that caused by gene deletion.

Intronic variation of *CHD7* impairs neurite development and dendritic morphology

Cortical glutamatergic neurons comprise the major excitatory network in the central nervous system (Molyneaux, Arlotta et al. 2007). Glutamatergic neurons play critical roles in controlling cognitive, emotion, language, and motor function. Dysfunction of cortical glutamatergic neurons may be relevant to autism (DeVito, Drost et al. 2007). Several of the known genetic disorders associated with autism have important implications for glutamatergic deficits in the disorder.

It was reported that the development of new born neurons in SGZ (subgranular zone) of adult *Chd7* null mice was severely compromised, showing less complex dendritic morphology, comparing to wild-type new born neurons (Feng, Khan et al. 2013). In order to study whether the intronic variation of *CHD7* found in autistic patients may affect neuronal development, we differentiated human embryonic stem cells into dorsal forebrain glutamate neurons according to established protocol (Fig. S2A). The cortical deeper-layer markers of glutamatergic neurons, TBR1 and CTIP2, were observed on D40 (Fig. S3A and S3B).

During the differentiation of NPCs towards mature neurons, neurites including axon and dendrite will start to develop and form functional synapses. Thus, to examine whether neuronal development may be affected by intronic variant of *CHD7*, we measured the neurites growth of differentiated neurons derived from human embryonic stem cells. To illustrate the morphology of neurons, we transfected GFP-expressing plasmids at D38 after neural differentiation. At D43, we performed immunostaining using antibodies against GFP and MAP2 (protein marker for dendrites) and found that the signals of GFP and MAP2 are fully overlapped (Fig. 3A), suggesting that the polarity of neurons has not been established as axonal differentiation has not accomplished. Although neuronal differentiation is ongoing at D43, we still found that total neurite length and branch number decreased in neurons carrying homozygous mutations in comparison with wild-type neurons derived from H9 ES cells, although the difference was not significant (Fig. 3B and 3C).

To examine whether intronic variation of *CHD7* may affect development of mature neurons, we transfected GFP-expressing plasmid into mature neurons derived from human embryonic stem cells at D68 from initiation of neural differentiation. At D72, we performed immunostaining on neurons with intronic variant and wild-type neurons, we found that the long axons labeled by GFP cannot be marked by MAP2 (Fig. 3D), suggesting the axonal differentiation is finished and mature dendrites were labeled by MAP2. We found that both dendritic length and branch number exhibited significant decrease in neurons carrying intronic variant comparing to the wild-type neurons derived from H9 ES cells (Fig. 3E and 3F), suggesting the intronic variation of *CHD7* impairs proper formation of dendritic morphology.

The morphological defects caused by intronic variation could be rescued by knocking down its up-regulated gene *TBR1*

To investigate whether the gene expression profile may be interfered by intronic variation of *CHD7* during neuronal development, we performed RNA-seq from RNA collected from differentiated neurons at D40 after initiation of neural differentiation. Interestingly, we found that the genes, which are upregulated in both neurons derived from hom1-H9-hESCs and hom2-H9-hESCs comparing to wild-type neurons, are much more than downregulated genes (Fig. 4A and 4B, Fig. S4A), strongly suggesting that *CHD7* may play a negative role in regulating gene expression in neurons. Genes which are closely related to neural development are marking with * (Fig. 4A and 4B). We found that *TBR1*, a critical gene implicated in autism, were strongly upregulated in neurons carrying intronic variation, comparing to wild-type neurons (Fig. 4A). We further validated this finding by using real-time quantitative PCR in neurons at D40 (Fig. 4C).

The *TBR1* gene has been reported to be implicated in amygdala development, laminar patterning of retinal ganglion cells, as well as cortical neurons (Huang, Chuang et al. 2014, Darbandi, Schwartz et al. 2018, Liu, Reggiani et al. 2018). Interestingly, *TBR1* is a putative transcription factor that is highly expressed in glutamatergic early-born cortical neurons and regulates differentiation of the preplate and layer 6 neurons (Hevner, Shi et al. 2001). Another up-regulated gene *EOMES* (*TBR2*), and *TBR1* are expressed sequentially by intermediate progenitor cells and postmitotic neurons in developing neocortex (Englund, Fink et al. 2005). Due to the function of *TBR1* in regulating neuron projection development, as well as several of the regulated genes are associated with *TBR1*, we wondered whether the morphological defects could be improved by returning the expression of *TBR1*.

We transfected shRNA for *TBR1* and GFP-expressing plasmid into differentiated neurons with the *CHD7* intronic variant at D60. As controls, shRNA for DsRed and GFP-expressing plasmid were transfected into D60 differentiated wild-type neurons. The vectors amount of shRNA was 3 times of GFP to ensure the GFP+ neurons were also be transfected with shRNA. At D72, we performed immunostaining on neurons with anti-GFP (Fig. 4D). From the results, both dendritic length and branch number in neurons with intronic variant transfected with shRNA for *TBR1* were improved from those transfected with shRNA for DsRed, and had no difference with wild type (Fig. 4E-4H). These evidences indicated that *TBR1* is an important downstream gene of *CHD7*.

Intronic variation of *CHD7* affects alternative splicing, resulting in three abnormal transcripts that are functionally deficient

Precise excision of introns is catalyzed by a sophisticated ribonucleoprotein machinery called the spliceosome (Papasaikas and Valcárcel 2016). Intron–exon boundaries are delimited by short consensus sequences at the 5' (donor) and 3' (acceptor) splicing sites that are recognized by the spliceosome. In addition, the spliceosome interacts with a catalytic adenosine (the branch point) and a polypyrimidine tract (PyT) located between the branch point adenosine and the 3' splicing sites (Vaz-Drago, Custódio et al. 2017).

Since the intronic variation of the *CHD7* gene locates adjacent to the branch point and may form a new splicing site, it would be intriguing to determine whether alternative splicing processes may be altered in cells carrying the intronic variant.

Since the intronic variant (ch8-61757392-C-T) locates in the intron between exon 21 and exon 22 (Fig. 1A), in order to examine potential transcripts, we amplified mRNA segments of *CHD7* from exon 21 to exon 24 with PCR in wild-type hESCs and cells carrying homozygous intronic variants (Fig. 5A). The amplification products were ligated with the T-vector followed by monoclonal sequencing. Interestingly, we found that, besides wild-type transcripts found in H9 ES cells, there are three novel transcripts identified in ES cells carrying intronic variants, including the transcript that deletion of exon 22-23 (del), the transcript with duplication of exon 22-23 (dup) and the transcript that retention 32bp of the intronic fragment that seamless connection upstream to exon 22 (Fig. 5B).

To further determine the expression level of each novel transcripts in wild-type ES cells and ES cells carrying intronic variants, we examine levels of each transcripts by performing quantitative PCR with specific primers in RNA samples collected from corresponding cell lines (arrows shown in Fig. 5B). Importantly, we found that the expression levels of three novel transcripts significantly increased in ES cells with intronic variant, in comparison with that in H9 ES cells (Fig. 5C-5E), suggesting the intronic variant compromised alternative splicing of the *CHD7* mRNA.

In order to further verify whether intronic variant may affect alternative splicing of *CHD7* in other cell lines, we also constructed the intronic point mutation cell lines using HEK293 cells, and successfully screened heterozygotes and homozygotes including intronic and synonymous mutations (3PM het and 3PM hom), as well as heterozygotes and homozygotes with only synonymous mutations (2PM het and 2PM hom) as the control. Synonymous mutations were introduced in the exonic sites using the same strategy in ES cells, to avoid unwanted digest of Cas9 after homologous recombination occurred. We found that intronic variants led to increase of three abnormal transcripts in HEK293 cells as well, in consistent with findings in human ES cells (Fig. S5A-S5C).

To further examine whether these abnormal types of alternative splicing exhibits in differentiated neural precursors and neurons, we performed Q-PCR experiments in NPCs and neurons derived from hES cells. Briefly, we amplified mRNA segments of *CHD7* from exon 21 to exon 24 by PCR on day D26 and D40 of differentiated wild-type H9 cells and cells carrying homozygous intronic variants. We found that three types of abnormal transcripts were all present as ES cells and consistently, the expression levels of three abnormal transcripts significantly increased in cells with intronic variant on D26 and D40, in comparison with that in H9 wild-type cells at same stage (Fig. S5D-S5I). Based on these results, we were very curious about the existence of three novel transcripts in the mRNA of variant carriers. Unfortunately, it is difficult to obtain samples from patients.

We would like to further address whether three abnormal transcripts of *CHD7* caused by intronic variant are functional transcripts. The transcript containing extra 32bp of the intronic fragment would lead to frameshift during translation thus is functionally deficient. Since exon 22 and exon 23 together contain

360bp, the del transcript which exon 22 and 23 were absent and the dup transcript which contains two copies of exon 22 and exon 23 are able to be translated to full length proteins with less or more amino acids coded by exon 22 and 23, respectively. To determine whether the protein products translated from the del or dup transcripts may function differentially as the wild-type CHD7 protein, we constructed the del and dup CHD7 expression constructs and examine the expression of wild-type, del and dup CHD7 in HEK293 by western blot. We found that the protein products generated from the del and dup transcripts are comparable to the wild-type CHD7 protein, which is around 336kd (Fig. S5J).

To determine whether the proteins carrying duplication of exon 22, 23 or without exon 22, 23 have normal functions as wild-type CHD7, we knockdown endogenous Chd7 in cortical culture neurons taken from E14.5 mice using shRNA which only target mouse Chd7, but not human CHD7 (Fig. S5K and S5L). Meanwhile, we performed the rescue experiment, by co-expression wild-type CHD7, del and dup CHD7 constructs respectively, along with shRNA against Chd7.

We first measured axonal growth at 3 day in vitro (DIV) and found that the total axonal length decreased significantly after Chd7 knockdown, which was fully rescued by wild-type human CHD7, but not either del or dup forms of CHD7, suggesting that del and dup forms of CHD7 are *loss-of-function* transcripts (Fig. S5M and S5N).

At 12 DIV, we found that the dendritic length and branch number significantly decreased after Chd7 knockdown (Fig. 5F-5H), indicating that Chd7 plays a critical role in proper development of cortical neurons. Importantly, the defects of dendritic growth was fully rescued by wild-type human CHD7, but not the del or dup transcripts (Fig. 5F-5H). Together, these evidences strongly indicate that the del and dup CHD7 transcripts caused by intronic variant function as *loss-of-function* manner.

In the previous results, we found that the intronic variant led to the increase of TBR1 expression (Fig. 4C) and defect in neuronal morphology (Fig. 4E-4H). In order to investigate whether this was caused by the decrease of CHD7, we analyzed the expression of TBR1 5 days after knockdown of CHD7 by shRNA in the differentiated neurons at D40. Results showed that CHD7 mRNA was knockdown by 40% (Fig. S5O) and TBR1 was significant up-regulated (Fig. 5I).

Based on the above results, we found that the intronic variant produced three novel transcripts, especially the huge increase of del transcript in 3PM het of 293 cells (Fig. S5A), dup transcript in 3PM hom of 293 cells (Fig. S5B) and in hom2 of H9 cells (Fig. 5D, S5E and S5H), which indicated that the intronic variant led to the instability of exon 22-23 splicing. Moreover, both transcripts did not show normal functions, it can be considered that this further reduced the content of normal transcripts in varying degrees due to alternative splicing. This may partly explain the phenotypic difference of the intronic variant carriers.

Thus, the intronic variant identified in autism patients indeed plays a critical role in regulating the function of CHD7 by downregulating the mRNA level and disrupted alternative splicing patterns. The proper neural differentiation and development of human NPCs and neurons were severely affected by

intronic variant, providing the crucial evidences for supporting the notion that intronic variant site of *CHD7* is a potentially autism susceptible site.

Discussion

Genes implicated in autism have provided critical insights into the pathogenesis of the disease. Some of well-documented autism risk factors include genes associated with rare syndromic forms of ASD (MECP2, FMR1, PTEN), synaptic cell adhesion and scaffolding molecules (NLGN3, NLGN4, NRXN1, CNTNAP2, SHANK3) and genes with de novo mutations (CHD8, SCN2A, DYRK1A among others) identified in whole-exome sequencing studies (Huguet and Bourgeron 2016). Due to the high heterogeneity of autism, the phenotype of patients varies from mild to severe and is affected by the genetic background of the family. Therefore, according to the unified criteria of autism diagnosis, large-scale family analysis with family members as control is very important for screening of risk mutations. Deleterious mutations are usually screened for coding regions and splicing sites, because such mutations can cause a serious impact on the loss of gene function. Even so, at most 25% of the cases can be detected with a genetic cause for ASD. The understanding of the genetic basis of autism has evolved from high-load disruption caused by a single mutation to a complex of multi-genes variation with low-load effect. These mutations alone is not enough to cause phenotype, but it can acts as a helper. For this, some carriers will get sick while others will be asymptomatic, making it difficult to present the comprehensive picture of genetic variation of ASD patients.

Many introns contain highly conserved sequence elements, including the consensus splice site sequences, the binding sites for regulatory proteins, as well as the sequences of non-coding RNA genes (Kelly, Georgomanolis et al. 2015). Alternative splicing increases the diversity of transcriptome by generating multiple mRNA isoforms from a single gene. A pre-mRNA molecule can be alternatively spliced through exon skipping, alternative splice site selection, and intron retention (Chen and Manley 2009). Mutations in intronic regions have been documented in various diseases. For example, a mutation that creates a novel donor splice site (ss) leads to inclusion of a 95 nucleotide intronic sequence in the BRCA2 mRNA (Anczuków, Buisson et al. 2012, Vaz-Drago, Custódio et al. 2017). A mutation that creates a novel binding site for SRSF1 activating a splicing enhancer element thus leads to inclusion of a 147 nucleotide pseudo-exon in the COL4A5 mRNA (King, Flinter et al. 2002, Vaz-Drago, Custódio et al. 2017). In addition, genetic variants have been reported to cause disease through inactivation of intron-encoded RNA genes (He, Liyanarachchi et al. 2011). In our study, intronic variation creates a novel acceptor splice site and leads to inclusion of a 32 nucleotide intronic sequence in the CHD7 mRNA (Fig. 5B and 5E). Interestingly, two abnormal transcripts that deletion of exon 22–23 and duplication of exon 22–23 were also found (Fig. 5B–5D). We have not further explored the mechanism of mutation disturbing the splicing stability of exon 22–23. Nevertheless, the results provide a new form of alternative splicing. If exon deletion is caused by the selection of splicing sites and lead to exon skipping, the mechanism of exon duplication is much more complicated, such as the process of sequence repeat may be combined of transcription and splicing.

As evidences we present in this study, we found that intronic variation of *CHD7* has critical effects on neural differentiation and morphology development. In our study, transcriptome sequencing in three stages of differentiation was performed (ESC, NPC, neuron). The genes with significant differences between the wild-type H9 cells and the cells carrying *CHD7* intronic variants on D26 and D40 were listed for analysis (Fig. 2G and 2H, 4A and 4B). While there were few regulated genes during the ESC period (data not show). Moreover, previous results showed that *CHD7* mRNA expressed in ESC while the protein expression was scarce, suggesting that the function of *CHD7* in stem cells is limited. In neural precursor cells, the affected genes participate in the biological processes of dopaminergic neuron differentiation, dentate gyrus development and Wnt signaling pathway (Fig. 2G and 2H, S2B). In neurons, the affected genes are involved in cell metabolism, signal transduction pathway, synaptic transmission and axon guidance (Fig. 4A and 4B, S4A).

Comprehensive analysis of three stages, we found that the up-regulated genes in each period are highly specific and be regulated only in the specific period, suggesting that the inhibitory function of *CHD7* to the target genes is stage specific. Among the down-regulated genes, there are not only highly specific genes, but also genes that being continuously regulated. For example, some genes are down-regulated in stem cells and precursor cells such as *ZNF826P*; some genes are down-regulated in precursor cells and neurons such as *ELAVL4*; Some genes are down-regulated in all three stages such as *LINC01087*, *MED15P4*, *ZNF528-AS1*. In addition, although some genes are regulated only in a specific period, there are direct interactions between regulated genes in different periods. For example, *FEZF2* that up-regulated in precursor phase can be inhibited by *TBR1* that up-regulated in neurons during neocortical development (McKenna, Betancourt et al. 2011). Based on the above results, *CHD7* begins to accumulate in the stem cells, but plays a role mainly in the neural differentiation and development. Its function partly period-specific and partly continuous. It suggests that different treatments should be given at different stages of development in the defects caused by *CHD7* mutation.

Limitations

The intronic variant were carried in heterozygous manner in probands. While in this study, two subclones that carrying intronic variant were both homozygous (Fig. S1), which may magnify the defect caused by intronic variant. The blood sample from patients were failed to obtained, so we couldn't verify whether a intronic variant based on the patients' genetic background would produce the same defect.

Conclusions

In conclusion, the level of *CHD7* mRNA significantly decreased in cell lines carrying homozygous mutants compared to control. Upon differentiation towards the forebrain neuronal lineage, neural cells carrying the *CHD7* intronic variant exhibited developmental delay and maturity defects. *TBR1*, a gene also implicated in ASD, significantly increased in neurons carrying the *CHD7* intronic variant. Furthermore, the morphological defects found in neurons carrying *CHD7* intronic mutations could be rescued by knocking down *TBR1*, indicating that *TBR1* may be responsible for defects in *CHD7*-related disorders. Finally, the

CHD7 intronic variant generates three abnormal forms of transcripts through alternative splicing, which all exhibited *loss-of-function* manners in functional assays. Our study provides the crucial evidences for supporting the notion that intronic variant site of *CHD7* is a potentially autism susceptible site, shedding new light on identifying functions of intronic variants in autism genetic studies.

Abbreviations

ASD: Autism spectrum disorders

hESCs: Human H9 embryonic stem cells

NPCs: Neural precursor cells

Del: the *CHD7* transcript that deletion of exon 22-23

Dup: the *CHD7* transcript with duplication of exon 22-23

Het: Heterozygous

Hom: Homozygous

Pm: Point mutation

Declarations

Acknowledgments

We thank Weijun Feng for providing the *CHD7*-expressing plasmid, and all the members of Qiu laboratory for comments on the manuscript.

Competing interests: The authors declare no competing interests.

Author contributions

R.Z. and Z.L.Q. conceived the project. Y.J.C and Z.L.Q supervised the project. Y.S.D provided samples of autistic patients. B.Y. performed genetic analysis of autistic susceptible genes. R.Z. and Z.Y.W constructed of point mutation cell lines. H.H. and X.Z.W. differentiation of hESCs. R.Z. performed all functional experiment. R.Z. and Z.L.Q. wrote and edited the manuscript.

Funding: This work was supported by National Key Research and Development Program of China (2018YFA0108000), the National Nature Science Foundation of China (NSFC) Grants (#31625013, #81941405, #31771137, #31722024, #91732302), Shanghai Brain-Intelligence Project from STCSM (#16JC1420501), the Strategic Priority Research Program of the Chinese Academy of Sciences (#XDBS01060200, XDA16010310) and the Shanghai Municipal Science and Technology Major Project

(#2018SHZDZX05). Shanghai Pujiang Program (17PJ1410200). The research is supported by the Open Large Infrastructure Research of Chinese Academy of Sciences.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Consent for publication

Not applicable

Ethics statement

We obtained assent from the Institutional Review Board (IRB), Shanghai Mental Health Center of Shanghai Jiao Tong University (FWA Number: 00003065; IROG Number: 0002202). Dr. Yi-Feng Xu approved and signed our study with ethical review number 2016–4. Written informed consent was obtained from parents in consideration of the fact that all patients were minors. All participants were screened using the appropriate protocol approved by the IRB of Shanghai Mental Health Center, Shanghai Jiao Tong University School of Medicine.

References

1. Anczuków, O., et al. (2012). "BRCA2 deep intronic mutation causing activation of a cryptic exon: opening toward a new preventive therapeutic strategy." Clinical Cancer Research**18**(18): 4903-4909.
2. Bourgeron, T. (2015). "From the genetic architecture to synaptic plasticity in autism spectrum disorder." Nature Reviews Neuroscience**16**(9): 551-563.
3. Chen, M. and J. L. Manley (2009). "Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches." Nature reviews Molecular cell biology**10**(11): 741-754.
4. Darbandi, S. F., et al. (2018). "Neonatal Tbr1 dosage controls cortical layer 6 connectivity." Neuron**100**(4): 831-845. e837.
5. DeVito, T. J., et al. (2007). "Evidence for cortical dysfunction in autism: a proton magnetic resonance spectroscopic imaging study." Biological psychiatry**61**(4): 465-473.
6. Englund, C., et al. (2005). "Pax6, Tbr2, and Tbr1 Are Expressed Sequentially by Radial Glia, Intermediate Progenitor Cells, and Postmitotic Neurons in Developing Neocortex." The Journal of Neuroscience**25**(1): 247-251.
7. Feng, W., et al. (2017). "Chd7 is indispensable for mammalian brain development through activation of a neuronal differentiation programme." Nature Communications**8**: 14758.
8. Feng, W., et al. (2013). "The chromatin remodeler CHD7 regulates adult neurogenesis via activation of SoxC transcription factors." Cell stem cell**13**(1): 62-72.

9. Gaugler, T., et al. (2014). "Most genetic risk for autism resides with common variation." Nature genetics**46**(8): 881.
10. Guo, H., et al. (2018). "Inherited and multiple de novo mutations in autism/developmental delay risk genes suggest a multifactorial model." Molecular autism**9**(1): 64.
11. Hartshorne, T. S., et al. (2005). "Autistic-like behavior in CHARGE syndrome." American Journal of Medical Genetics Part A**133**(3): 257-261.
12. He, H., et al. (2011). "Mutations in U4atac snRNA, a component of the minor spliceosome, in the developmental disorder MOPD I." science**332**(6026): 238-240.
13. Hevner, R. F., et al. (2001). "Tbr1 Regulates Differentiation of the Preplate and Layer 6." Neuron**29**(2): 353-366.
14. Huang, T.-N., et al. (2014). "Tbr1 haploinsufficiency impairs amygdalar axonal projections and results in cognitive abnormality." Nature Neuroscience**17**(2): 240-247.
15. Huguet, G. and T. Bourgeron (2016). Genetic causes of autism spectrum disorders. Neuronal and synaptic dysfunction in autism spectrum disorder and intellectual disability, Elsevier: 13-24.
16. Jiang, Y.-h., et al. (2013). "Detection of clinically relevant genetic variants in autism spectrum disorder by whole-genome sequencing." The American Journal of Human Genetics**93**(2): 249-263.
17. Kanner, L. (1943). "Autistic disturbances of affective contact." Nervous child**2**(3): 217-250.
18. Kelly, S., et al. (2015). "Splicing of many human genes involves sites embedded within introns." Nucleic acids research**43**(9): 4721-4732.
19. King, K., et al. (2002). "Unusual deep intronic mutations in the COL4A5 gene cause X linked Alport syndrome." Human genetics**111**(6): 548-554.
20. Liu, J., et al. (2018). "Tbr1 instructs laminar patterning of retinal ganglion cell dendrites." Nature Neuroscience**21**(5): 659-670.
21. McKenna, W. L., et al. (2011). "Tbr1 and Fezf2 regulate alternate corticofugal neuronal identities during neocortical development." Journal of Neuroscience**31**(2): 549-564.
22. Molyneaux, B. J., et al. (2007). "Neuronal subtype specification in the cerebral cortex." Nature Reviews Neuroscience**8**(6): 427-437.
23. Nichols, D. H., et al. (2008). "Lmx1a is required for segregation of sensory epithelia and normal ear histogenesis and morphogenesis." Cell and tissue research**334**(3): 339-358.
24. Papasaikas, P. and J. Valcárcel (2016). "The spliceosome: the ultimate RNA chaperone and sculptor." Trends in biochemical sciences**41**(1): 33-45.
25. Petty, E. and L. Pillus (2013). "Balancing chromatin remodeling and histone modifications in transcription." Trends in Genetics**29**(11): 621-629.
26. Pinto, D., et al. (2014). "Convergence of genes and cellular pathways dysregulated in autism spectrum disorders." The American Journal of Human Genetics**94**(5): 677-694.
27. Ronald, A. and R. A. Hoekstra (2011). "Autism spectrum disorders and autistic traits: a decade of new twin studies." American Journal of Medical Genetics Part B: Neuropsychiatric Genetics**156**(3): 255-

274.

28. Sandin, S., et al. (2014). "The familial risk of autism." Jama**311**(17): 1770-1777.
29. Satterstrom, F. K., et al. (2020). "Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism." Cell.
30. Takada, I., et al. (2009). "Wnt and PPAR γ signaling in osteoblastogenesis and adipogenesis." Nature Reviews Rheumatology**5**(8): 442.
31. Takata, A., et al. (2018). "Integrative analyses of de novo mutations provide deeper biological insights into autism spectrum disorder." Cell reports**22**(3): 734-747.
32. Tick, B., et al. (2016). "Heritability of autism spectrum disorders: a meta-analysis of twin studies." Journal of Child Psychology and Psychiatry**57**(5): 585-595.
33. Vaz-Drago, R., et al. (2017). "Deep intronic mutations and human disease." Human genetics**136**(9): 1093-1111.
34. Vissers, L. E., et al. (2004). "Mutations in a new member of the chromodomain gene family cause CHARGE syndrome." Nature genetics**36**(9): 955-957.
35. Yao, X., et al. (2017). "Homology-mediated end joining-based targeted integration using CRISPR/Cas9." Cell research**27**(6): 801-814.

Figures

Figure 1

An intronic variation of CHD7 found in Chinese ASD leads to downregulation of CHD7 mRNA. (A) The position of intronic variation in the genomic structure of CHD7 gene. (B) Genogram of the six families with proband carrying CHD7 variation. The blank graph indicates non-carriers; the blank graph with a center indicate asymptomatic carriers; the solid graph indicates sick carriers. (C) Sanger sequencing to verify the intronic variation in CHD7 gene. The red arrowheads indicate the variant base. (D) The diagram of gene editing. The purple arrowhead indicates normal base, the red arrowheads indicate introduced mutant base. HAL / HAR, left / right homology arm. (E-G) mRNA levels as detected by RT-PCR for CHD7 in cell lines carrying homozygous mutants with primers located in (E) exon 19-20; (F) exon 25-26; (G) exon 35-38. Value represents mean $\pm$ SEM ($n = 3$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$; one-way ANOVAs). See also Figure S1.

Figure 2

Intronic variation of CHD7 delays neuronal differentiation of hESCs. (A) Confocal images of differentiated human neural precursor cells (H9 control and hom1, hom2) on D28 coimmunostained for Ki67 and SOX2. The scale bar represents 50 μ m. (B) Proportion of Ki67 positive cells among SOX2 positive cells. At least 1,000 cells of 8 fields were analyzed for each group, respectively. (C) Confocal images of differentiated human neural precursor cells (H9 control and hom1, hom2) on D28 coimmunostained for SOX2, PAX6, Tuj1 and DAPI. The scale bar represents 50 μ m. (D-F) Statistical ratio

of (D) PAX6, (E) SOX2 and (F) Tuj1 to DAPI. At least 1,000 cells of 8 fields were analyzed for each group, respectively. (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$; one-way ANOVAs). (G and H) Heatmap exhibits significantly (G) upregulated genes and (H) downregulated genes (Fold change >1.5 , FDR < 0.05) in differentiated neural precursor cells hom1 and hom2 compared to H9 control on D26. The colour scale is shown on the right. * indicate genes that are involved in neural differentiation and development. See also Figure S2.

Figure 3

Intronic variation of CHD7 impairs neurite development and dendritic morphology. (A) Representative images of differentiated neurons (H9 control and hom1, hom2) on D43 transfected with GFP, coimmunostained for GFP, MAP2, CHD7 and DAPI. The scale bar represents 50 μm . (B and C) Quantification of the (B) total neurite length and (C) branch number of differentiated neurons (H9 control and hom1, hom2) on D43. 25-30 neurons were analyzed for each group, respectively. (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$; one-way ANOVAs). (D) Representative images of differentiated neurons (H9 control and hom1, hom2) on D72, coimmunostained for GFP, MAP2, CHD7 and DAPI. Yellow arrowheads indicate axons labeled by GFP but not MAP2. The scale bar represents 50 μm . (E and F) Quantification of the (E) total dendrite length and (F) branch number of differentiated neurons (H9 control and hom1, hom2) on D72. 25-30 neurons were analyzed for each group, respectively. (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$; one-way ANOVAs). See also Figure S3.

Figure 4

The morphological defects caused by intronic variation could be rescued by knocking down its up-regulated gene TBR1. (A and B) Heatmaps exhibit significantly (A) upregulated genes and (B) downregulated genes (Fold change >1.5, FDR < 0.05) in differentiated neurons hom1 and hom2 compared to H9 control on D40. The color scale is shown on the right. * indicate genes that are involved in neural differentiation and development. (C) mRNA levels as detected by RT-PCR for TBR1 in H9 and cell

lines carrying homozygous mutants. Value represents mean $\pm$ SEM ($n = 3$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$; t-test). (D) Representative images of differentiated neurons (H9 control and hom1, hom2) on day D72 transfected with GFP and shRNA, coimmunostained for GFP, MAP2, TBR1 and DAPI. The scale bar represents 50 μm . (E-H) Quantification of the (E and G) total dendrite length and (F and H) branch number of differentiated neurons (H9 control and hom1, hom2) on D72. At least 30 neurons were analyzed for each group, respectively. (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$; one-way ANOVAs). See also Figure S4.

Figure 5

Intronic variation of CHD7 affects alternative splicing, resulting in three abnormal transcripts that are functionally deficient. (A) Agarose gel electrophoresis illustrating PCR products of exon 21-24 from cDNA of H9 control, hom1 and hom2. (B) Schematic representation of alternative splicing around the intronic mutation. (C-E) mRNA levels as detected by RT-PCR for three kinds of abnormal transcripts (C) exon22-23 deletion (D) exon22-23 duplication and (E) intron retention in H9 control and two point-mutant cell lines. Value represents mean $\pm$ SEM (n = 3, *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001; one-way ANOVAs). (F) Representative images of E14.5 mouse primary cortical neurons transfected with GFP, together with shRNA for DsRed as control or shRNA for mouse Chd7, and with human CHD7 transcript wild-type, del or dup at DIV 12d, coimmunostained for GFP, CHD7, MAP2 and DAPI. The scale bar represents 50 μ m. (G and H) Quantification of the (G) total dendrite length and (H) branch number of DIV 12d neurons. At least 30 neurons were analyzed for each group, respectively. (*p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001; one-way ANOVAs). (I) mRNA levels as detected by RT-PCR for TBR1 in the differentiated neurons after transfected with shRNA for CHD7 at D40 for 5d. Value represents mean $\pm$ SEM (n=8, *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001; t-test). See also Figure S5.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigureSupp.docx](#)
- [figureS1.jpg](#)
- [figureS2.jpg](#)
- [figureS3.jpg](#)
- [figureS5.jpg](#)
- [figureS4.jpg](#)