



## OPEN ACCESS

## EDITED BY

Jernej Kovac,  
University Medical Center Ljubljana, Slovenia

## REVIEWED BY

Christopher McGraw,  
Massachusetts General Hospital and Harvard  
Medical School, United States  
Gasper Klancar,  
Institute of Oncology Ljubljana, Slovenia

## \*CORRESPONDENCE

Bo Tan

✉ tanbo@hospital.cqmu.edu.cn

Sanchuan Luo

✉ sanchuanluo@gmail.com

RECEIVED 10 August 2023

ACCEPTED 06 November 2023

PUBLISHED 24 November 2023

## CITATION

Wu Y, Li W, Tan B and Luo S (2023)

Identification of novel *SHANK2* variants in two  
Chinese families via exome and RNA  
sequencing.

*Front. Neurosci.* 17:1275421.

doi: 10.3389/fnins.2023.1275421

## COPYRIGHT

© 2023 Wu, Li, Tan and Luo. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Identification of novel *SHANK2* variants in two Chinese families via exome and RNA sequencing

Yong Wu<sup>1</sup>, Wenzhou Li<sup>1</sup>, Bo Tan<sup>2\*</sup> and Sanchuan Luo<sup>1\*</sup>

<sup>1</sup>Medical Research Institute, Shenzhen Baoan Women's and Children's Hospital, Shenzhen, China,

<sup>2</sup>Department of Gynecology and Obstetrics, The Second Affiliated Hospital of Chongqing Medical University, Chongqing, China

**Background:** *SHANK2* encodes a postsynaptic scaffolding protein involved in synapse formation, stabilization and homeostasis. Variations or microdeletions in the *SHANK2* gene have been linked to a variety of neurodevelopmental disorders, including autism spectrum disorders (ASD) and mild to moderate intellectual disability (ID) in human. However, the number of reported cases with *SHANK2* defects remains limited, with only 14 unrelated patients documented worldwide.

**Methods:** In this study, we investigated four patients from three families with ID. Whole-exome sequencing (WES) was performed to explore the genetic causes, while Sanger sequencing was used to confirm the identified variants. Furthermore, RNA sequencing and functional enrichment analysis were performed on patients with likely pathogenic variants to gain further insights into the molecular landscape associated with these variants.

**Results:** Two novel variants in the *SHANK2* gene: a heterozygous splicing substitution (NM\_012309.5:c.2198-1G>A p.Pro734Glyfs\*22) in Family 1, and a heterozygous nonsense variant [NM\_012309.5:c.2310dupT p.(Lys771\*)] in Family 2 were identified by WES and confirmed by Sanger sequencing. RNA sequencing and cohort analysis identified a total of 1,196 genes exhibiting aberrant expression in three patients. Functional enrichment analysis revealed the involvement of these genes in protein binding and synaptic functions.

**Conclusion:** We identified two novel loss of function variants that broadens the spectrum of *SHANK2* variants. Furthermore, this study enhances our understanding of the molecular mechanisms underlying *SHANK2*-related disorders.

## KEYWORDS

*SHANK2*, neurodevelopmental disorder, whole-exome sequencing, RNA sequencing, cohort analysis

## Introduction

Neurodevelopmental disorders (NDDs) are a group of mental health disorders resulting from the disruptions in crucial neurodevelopmental processes, leading to abnormal brain function that can affect emotions, cognition, learning, self-regulation, and memory (Morris-Rosendahl and Crocq, 2020). The severity and behavioral phenotypes observed in NDD patients vary widely, with diagnoses commonly including autism spectrum disorder (ASD), intellectual disability (ID), developmental delay (DD), and epilepsy (Zablotsky et al., 2019). Previous research has implicated various genetic variants in NDDs, including chromosomal

rearrangements, copy number variants (CNVs), and coding-sequence variants. Although numerous genes have been associated with these disorders, each gene or genomic alteration typically accounts for less than 1% of cases. Many of the genes implicated in NDDs play a role in the development or functioning of neuronal circuits. Among the most extensively studied biological pathways in NDDs are those involving synaptic genes (Toro et al., 2010; Guilmatre et al., 2014; Hu et al., 2014; Leblond et al., 2014; Parenti et al., 2020).

The SH3 and multiple ankyrin repeat domains2 (*SHANK2*) gene is located on chromosome 11q13.3 and belongs to the *SHANK* gene family. *SHANK2* encode a pivotal scaffold protein in the postsynaptic density (PSD) complexes of glutamatergic synapses. The PSD is a specialized structure of the postsynaptic membrane that plays a critical role in neuronal signaling. The *SHANK2* protein contains multiple domains facilitating protein–protein interactions and is vital for organizing the PSD through a complex network of molecular interactions (Sheng and Kim, 2000; Sasaki et al., 2020). In *Shank2* knock-out mice, both the ionotropic glutamate receptors at the synapse and the level of *Shank3* are upregulated. The mutant mice exhibit reduced dendritic spines and basal synaptic transmission. Moreover, they display remarkably hyperactive behavior and manifest significant autistic-like behavioral alterations, including repetitive grooming and deviations in vocal and social behaviors (Schmeisser et al., 2012; Won et al., 2012; Yoo et al., 2014).

Variants in the *SHANK2* gene have been implicated in individuals with ASD and ID. The initial discovery of *de novo* CNVs in the *SHANK2* gene in two unrelated patients was reported by Berkel et al. (2010), using microarray analysis. Subsequent investigations involved sequencing the *SHANK2* gene in a larger cohort of individuals, including 396 ASD cases, 184 cases of ID, and 659 unaffected individuals, leading to the identification of additional variants specific to ASD and ID (Berkel et al., 2010). In a study by Leblond et al., *SHANK2* was sequenced in 455 patients with ASD and 431 controls, and the findings were integrated with the previous research. A notable finding was the significant enrichment of variants affecting conserved amino acids in affected patients compared to controls. Furthermore, functional studies demonstrated a reduction in synaptic density at dendrites when neuronal cells were transfected with the variants identified in patients, as opposed to those exclusively detected in controls. These extensive investigations provide compelling evidence that certain *SHANK2* variants may confer an increased risk of ASD (Leblond et al., 2012).

Recently, there has been growing interest in utilizing total RNA sequencing in conjunction with whole-genome sequencing (WGS) or whole-exome sequencing (WES) to enhance our understanding of variant pathogenicity. This integrated approach enables the detection of outliers in both expressions and splicing, facilitating the interpretation of functional consequences (Kremer et al., 2017; Liu et al., 2022; Pan et al., 2022; Peymani et al., 2022). Moreover, it provides a valuable opportunity to investigate the molecular mechanisms underlying loss of function (LOF) variants in the *SHANK2* gene. In the present study, we investigated two novel *SHANK2* variants identified in three patient with ID from two families. Both the variants are LOF variants. Additionally, RNA sequencing and cohort analysis were performed on these patients to gain further insights into the impact of these LOF variants on gene expression. Through comprehensive analysis, we identified numerous genes with aberrant expression, which significantly contributed to our

understanding of the molecular mechanisms associated with LOF variants in the *SHANK2* gene. These findings provide valuable insights into the pathogenicity of *SHANK2* variants and shed light on the underlying molecular processes involved in ID.

## Materials and methods

### Ethical compliance

Prior to their participation in this study, informed consent was obtained from all patients or their legal guardians. This research was conducted in accordance with ethical guidelines and regulations established by the ethics committee of the Second Affiliated Hospital of Chongqing Medical University (Approval No. 2022-549, dated 7 March 2022).

### DNA isolation, whole-exome sequencing, and variant analysis

Peripheral blood samples were collected from the patients using EDTA tubes, and genomic DNA was isolated using DNeasy Blood & Tissue kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. A total of three microgram of genomic DNA was randomly fragmented and captured using the Agilent SureSelectXT V5 capture kit (Agilent Technologies, Santa Clara, CA). Sequencing was performed on an Illumina HiSeq2000 (Illumina, San Diego, CA) with 100-bp paired-end reads, following the recommended protocols. To ensure data quality, the raw sequencing reads underwent filtering using Fastp (Chen et al., 2018b) to obtain clean reads. FastQC was employed to evaluate the quality of the sequencing data in each sample (Trivedi et al., 2014). The clean DNA sequencing reads were aligned to the human reference genome hg19 (GRCh37) using the BWA-MEM algorithm (Li and Durbin, 2009). Ambiguously mapped reads (MAPQ < 10) and duplicated reads were removed using SAMtools (Li et al., 2009) and PicardTools,<sup>1</sup> respectively. Single nucleotide polymorphisms (SNPs) and small insertions and deletions (INDELs) were identified following the best practices recommended by the Genome Analysis Toolkit software (McKenna et al., 2010). Variants were annotated using the Ensembl Variant Effect Predictor (McLaren et al., 2016). The ACGS Best Practice Guidelines for Variant Classification in Rare Disease 2020 were followed (Ellard et al., 2020). Classification of the variants into pathogenic (P), likely pathogenic (LP), benign (B), likely benign (LB), or variants of uncertain significance (VUS) was performed in accordance with the ACMG/AMP and ACGS guidelines (Richards et al., 2015; Ellard et al., 2020). All identified variants were further validated by Sanger sequencing.

### Minigene splicing assay

Functional analysis of the NM\_012309.5:c.2198-1G>A variant was performed using an *in vitro* minigene assay. Genomic regions

<sup>1</sup> <http://broadinstitute.github.io/picard/>

containing exons 18, 19, 20, 21, and intron sequences (2,355 bp) of *SHANK2* from proband 1 (individual II-1 from Family 1) were amplified using primers (SH2-F: 5'-tagtccagtgtggtgaattcATGTGTGCAAAGTCGGCCACA-3', SH2-R: 5'-gcccttagactcgagcggccgcCCTTCTTCTTCCGGACCGAG-3'). The amplified fragments were then cloned into the pcDNA3.1(+) plasmid. Since proband 1 is heterozygous for this variant, clones of mutant-type hybrid minigene (minigene-MT) and wild-type hybrid minigene (minigene-WT) were obtained, respectively, and confirmed through Sanger sequencing. Subsequently, the hybrid minigenes were transduced into HEK293T cells with Lipofectamine 2000 reagent (Invitrogen). RNA was extracted 48 h after transfection followed by RT-PCR using the same upper primers. The resulting products were subjected to electrophoresis analysis and Sanger sequencing for further characterization.

## RNA isolation, sequencing and data preprocessing

Peripheral blood samples from the patients were collected using EDTA tubes. Subsequently, the red blood cells were removed through centrifugation after incubation with a red blood cell lysis solution. Total RNA sample was isolated within 24 h of collection and enriched using oligo-dT bead capture. Complementary DNA synthesis was performed following the manufacturer's instructions, and libraries were prepared using the Illumina TrueSeq stranded mRNA sample prep kit (Illumina, San Diego, CA). Subsequently, sequencing of the pooled samples was conducted on a NovaSeq 6000 sequencing system. To obtain high-quality data, the raw sequencing reads underwent processing using Fastp to obtain clean reads (Chen et al., 2018a). Quality assessment of the sequencing data was performed using FastQC and mulitQC, evaluating factor such as sequence quality per base, sequence duplication level, and quality score distribution for each sample. The average quality score for the RNA sequences exceeded 30, indicating that substantial portion of high-quality sequences (Ewels et al., 2016). The clean RNA-sequencing reads were then aligned to the human reference genome (hg19) using STAR (2.4.2a) in conjunction with the Gencode v19 annotation (Dobin et al., 2013). Mapping evaluation metrics, including sequencing depth, percentage of mapped reads, and the number of expressed genes, were computed using DROP v1.21 (Yepez et al., 2021). Furthermore, the match between the RNA sequencing sample and its annotated DNA sample was assessed using DROP v1.21, with a cutoff of 0.8. Aberrant gene expressing was detected using DROP v1.21 (Yepez et al., 2021).

The clean RNA-sequencing reads were aligned to the human reference genome (hg19) using STAR (2.7.8a) along with the Gencode v29 annotation (Dobin et al., 2013). The Genomic Alignments R package's "summarizeOverlaps" function was used for read counting. To enhance statistical power, we performed aberrant expression analysis by combining our data with 367 blood samples from GTEx data. Genes with a 95th percentile Fragments Per Kilobase of transcript per Million mapped reads (FPKM) less than 1 were considered as lowly expressed and were excluded from downstream analysis. OUTRIDER was employed to identify expression outliers (Brechtmann et al., 2018). Technical and biological covariates such as sex, age, and sequencing batch were automatically controlled by OUTRIDER, which utilized an autoencoder implementation. Genes were considered to have aberrant expression if they had an adjusted value of  $p < 0.05$ .

## Pathway enrichment analysis

To further explore the functional implications of the identified aberrations, we performed functional enrichment analysis using the KOBAS-I service (Bu et al., 2021). This comprehensive tool provides pathway enrichment analysis by leveraging various databases including GO, KEGG, Reactome, and GWAS catalogs. Pathways with an adjusted value of  $p < 0.05$  were considered as significant, providing valuable insights into the biological relevance of the aberrant gene and their involvement in key pathways and biological processes.

## Results

### Clinical presentation

This study included three unrelated Chinese families (Figure 1; Supplementary Figure S1). Proband 1 (individual II-1 from Family 1), a 27-year-old male, was the second child of non-consanguineous healthy parents. He had an uneventful full-term birth, walked at 3 years old, and began speaking at 7 years old. Proband 1 exhibited poor learning ability, limited mathematical skills, and discontinued education after the first grade of elementary school (Figure 1). Proband 2 (individual II-1 from Family 2), a 10-year-old female, was the first child of unrelated parents. The pregnancy and delivery were normal. The primary phenotype observed in this patient was mild ID. At the time of diagnosis, she was attending a regular primary school. Her brother was unaffected, but her mother had a diagnosis of mild ID (Figure 1).

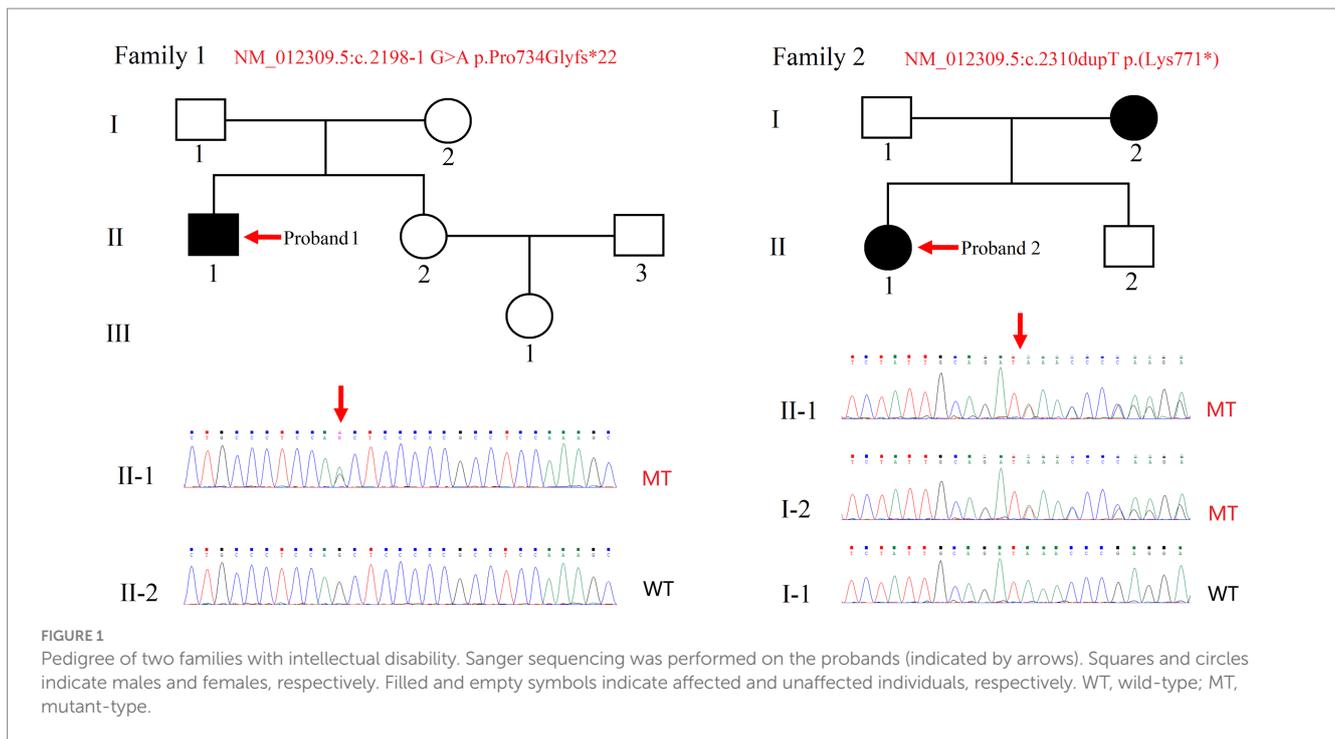
Proband 3 (individual II-1 from Family 3) is a 4-year-old boy born to non-consanguineous healthy parents as their only child. He presented with globally development delay, ID and exhibited tendencies toward ASD features including impaired social interactions, repetitive behaviors, and delayed speech development. There was no reported family history of similar conditions or disorders (Supplementary Figure S1).

### WES analysis

Due to proband 1 (individual II-1 from Family 1) being raised by grandparents while parents worked in another city, only singleton WES was performed, along with collection of peripheral blood from the proband's sister. Through the analysis of WES data and variant pathogenicity classification following ACMG guidelines, only one variant in the *SHANK2* gene (NM\_012309.5:c.2198-1G>A) were identified in the proband (Figure 1). Then, the variant was classified with criteria PVS1 + PM2 + PP3 and annotated as "LP" (Table 1). We have submitted this variant to ClinVar. It can be referenced under Submission Number: SUB13920791.<sup>2</sup>

In Family 2, trio WES was conducted for proband 2 (individual II-1) and her parents, since the mother was also affected. This revealed only one variant in the *SHANK2* gene in the proband—NM\_012309.5:c.2310dupT p.(Lys771\*). Individual II-2, who is the younger brother of Proband 2, resides in a different region, and we were unable to obtain a sample from him (Figure 1). Then, the variant was classified with criteria PVS1 + PP1\_moderate + PM2 and annotated as "LP" (Table 1). We have

<sup>2</sup> <https://ncbi.nlm.nih.gov/clinvar/>



submitted this variant to ClinVar. It can be referenced under Submission Number: SUB13920833 (see footnote 2).

Then for proband 3 (individual II-1 from family 3), singleton WES was carried out due to cost considerations. Blood was collected to validate inheritance. Following ACMG guidelines, this analysis uncovered a variant in the *SHANK2* gene [NM\_012309.5:c.178C>T p.(Arg60Cys)], which was classified based on criterion PM2 and annotated as “VUS” (Table 1). Another variant identified was in the *GLTSC1* gene [NM\_015711.3:c.1324G>C p.(Ala442Pro)], also classified under criterion PM2 and labeled as “VUS.” Moreover, two variants were found in the *NBAS* gene [NM\_015909.3:c.392C>T p.(Pro131Leu) and NM\_015909.3:c.3674A>C p.(Glu1225Ala)]. Of these, the variant NM\_015909.3:c.392C>T p.(Pro131Leu) was classified with criterion PM2 and annotated as “VUS,” while the variant NM\_015909.3:c.3674A>C p.(Glu1225Ala) was classified using criteria PM2 + PP3 and also annotated as “VUS.” Diseases associated with *NBAS* predominantly affect the liver, which does not align with the ID phenotype observed in proband 3 (Staufner et al., 2020). Conversely, the *GLTSC1* gene’s variants are implicated in the dominantly inherited NDD, Coffin-Siris syndrome-12 (Barish et al., 2020). Given our study’s emphasis on unveiling novel pathogenic variants in *SHANK2* and their potential mechanisms, our subsequent research will not encompass the less substantiated variant NM\_012309.5:c.178C>T p.(Arg60Cys) identified in proband 3.

Sanger sequencing was used to confirm the presence of these variants identified in the *SHANK2* gene (Figure 1; Supplementary Figure S1). The variants identified in proband 1 (individual II-1 from Family 1) and proband 2 (individual II-1 from Family 2) were novel. The NM\_012309.5:c.2198-1G>A variant is a classical splice site variant. On the other hand, the NM\_012309.5:c.2310dupT p.(Lys771\*) variant immediately introduces a premature termination codon (TAA) due to the presence of a downstream AAA codon (Lys). Both of these LOF variants are located in the proline-rich region of the SHANK2 protein (Figure 2).

Both LOF variants are classified as “LP,” while the NM\_012309.5:c.178C>T p.(Arg60Cys) is interpreted as VUS based on the available evidence (Table 1).

## Minigene splicing assay

*SHANK2* is primarily expressed in the nervous system. However, due to ethical considerations, we were limited to obtaining peripheral blood samples from the patients. Given the low expression of *SHANK2* in peripheral blood, we resorted to the minigene assay to uncover the true impact of the NM\_012309.5:c.2198-1G>A variant on pre-mRNA splicing. Subsequently, RT-PCR was employed to analyze the splicing products. Upon agarose gel electrophoresis, it was observed that cells transduced with minigene-WT produced a 240 bp band, whereas cells transduced with minigene-MT generated a 223 bp band. Subsequent Sanger sequencing verified that the minigene-WT product aligned with the reference sequence. Conversely, the minigene-MT product exhibited a skipping event of the first 17 bp of exon 19, leading to a frameshift and the generation of a premature stop codon (Figure 3).

## Transcriptome and cohort analysis

In order to investigate the potential molecular mechanism underlying these novel likely pathogenic variants, we performed RNA sequencing. Cohort analysis was conducted on RNA sequencing data from three patients of family 1 and 2, comparing them to publicly available databases as controls. A total of 1,196 genes were identified to exhibit aberrant expression patterns (Figure 4; Supplementary Table S1). Notably, among these genes, several were found to be associated with *SHANK2* and synapse function. One such gene is Glutamate receptor, ionotropic, N-methyl-D-aspartate associated protein 1 (*GRINA*). *GRINA* encodes a postsynaptic density

TABLE 1 Variants identified in the SHANK2 gene in probands in this study.

Proband	Localization	Nucleotide substitutions	AA change	SIFT	Polyphen2	CADD	SpliceAI	REVEL	Novel or ClinVar	Novel or HGMD	Population frequency [ExAC/gnomAD (non-neuro)]	Variant classification	ACMG classification
1	Intron 18	NM_012309.5:c.2198-1G>A	p.Pro734Glyfs*22	N/A	N/A	33	Splice-Altering/strong	N/A	Novel	Novel	0/0	PVS1 + PM2 + PP3	LP
2	Exon 22	NM_012309.5:c.2310dupT	p.(Lys771*)	N/A	N/A	N/A	N/A	N/A	Novel	Novel	0/0	PVS1 + PPI_ moderate+PM2	LP
3	Exon 3	NM_012309.5:c.178C>T	p.(Arg60Cys)	D	PD	24.4	N/A	Uncertain (0.36)	Novel	Novel	0.000104/0.0000385	PM2	VUS

Proband 1, individual II-1 from family 1; Proband 2, individual II-1 from family 2; Proband 3, individual II-1 from family 3. An asterisk (\*) denotes variants resulting in premature termination codons. ExAC, Exome Aggregation Consortium projects; gnomAD (non-neuro), Genome Aggregation Database; REVEL, rare exome variant ensemble learner; N/A, not applicable; D, deleterious; PD, possibly damaging.

protein involved in anchoring glutamate receptors (Schmeisser et al., 2012). Another gene of interest is *CTTN*, which encodes cortactin, an actin regulatory protein enriched at excitatory synapses (Mac Gillavry et al., 2016).

### Pathway enrichment analysis

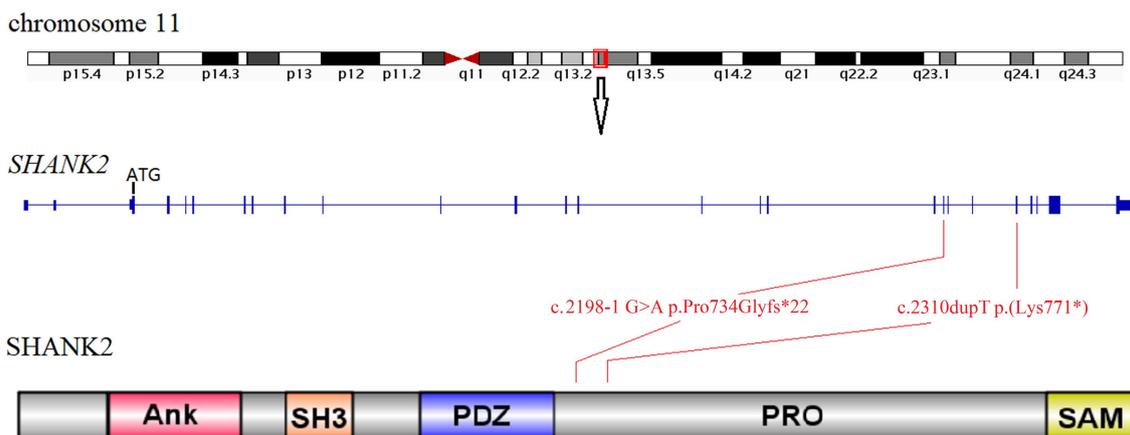
In order to further understand the functional implications of the aberrantly expressed genes, we performed pathway enrichment analysis. The functional annotation of these genes revealed their involvement in various biological pathways. Notably, a significant proportion of the genes were found to be associated with protein binding, indicating their participation in protein–protein interaction and molecular processes. Additionally, a subset of aberrantly expressed gene was found to be associated with the activation of NMDA receptors and postsynaptic events, further supporting their involvement in synaptic function and neuronal signaling. These findings are consistent with the known role of *SHANK2* as a postsynaptic scaffolding protein, highlighting the potential impact of the identified variants on synaptic organization and function (Figure 5; Supplementary Table S2).

### Discussion

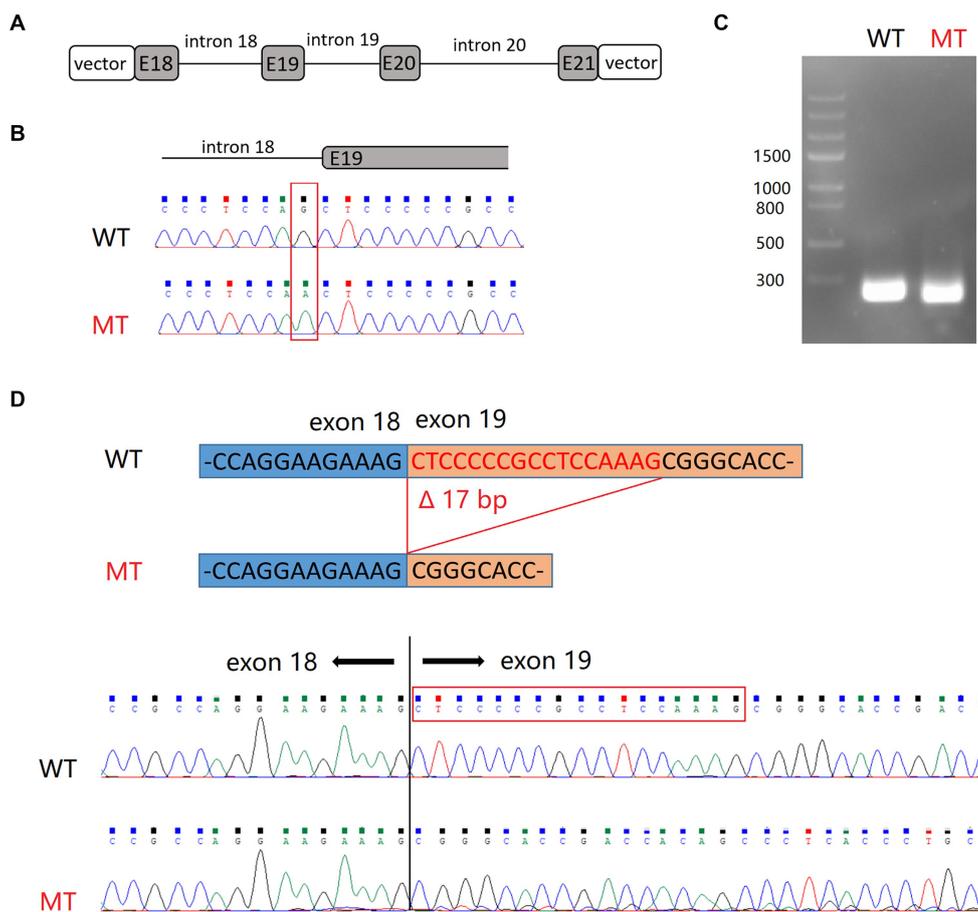
The *SHANK* gene family, consisting of *SHANK1*, *SHANK2*, and *SHANK3*, encodes multi-domain master scaffold proteins that play critical roles in the organization and function of the postsynaptic density (PSD) complexes at glutamatergic synapses. *SHANK* proteins participate in various synaptic functions by interacting with many synaptic proteins (Guilmatre et al., 2014; Monteiro and Feng, 2017). Variants in *SHANK* genes have been repeatedly reported in individuals with a range of NDDs (Leblond et al., 2014; Doddato et al., 2022).

Among the *SHANK* gene family members, *SHANK2* is the largest gene and is located on chromosome 11q13.3 (Figure 2). Only 14 cases with *SHANK2* variants have been documented before. In this study, we present the identification of two novel *SHANK2* variants [NM\_012309.5:c.2198-1G>A p.Pro734Glyfs\*22 and NM\_012309.5:c.2310dupT p.(Lys771\*)] in two unrelated Chinese families. Both variants are located within the proline-rich region (PRO) of *SHANK2* peptide. Out of the total 17 cases, seven individuals carried microdeletions encompassing *SHANK2* gene, while nine cases resulted in premature stop codons. Interestingly, the NM\_012309.5:c.2198-1G>A p.Pro734Glyfs\*22 variant reported in this study represents the first splicing variant identified in *SHANK2*, and is also considered a LOF variant (Table 2).

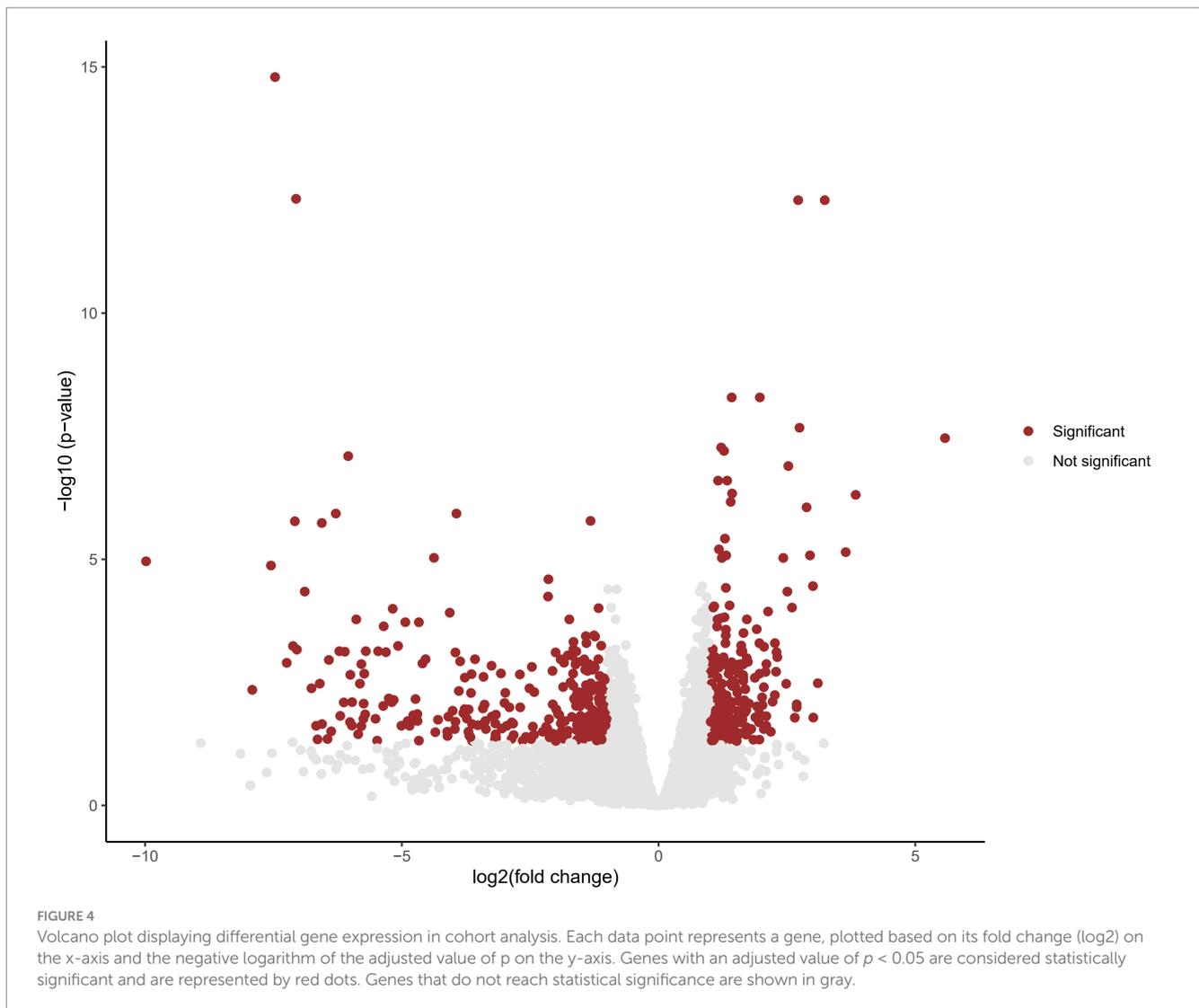
Through cohort analysis on the transcriptomic data from three patients carrying the identified novel LOF variants in *SHANK2*, a total of 1,196 genes exhibiting aberrant expression were identified. This dataset, derived from patient samples, presents a valuable resource providing insights into the molecular landscape associated with the disorder. *GRINA* belongs to the NMDA receptors (NMDARs) family. Studies conducted on mice lacking exons 6–7 of *Shank2* have demonstrated autistic-like behavioral abnormalities, which have been linked to altered N-methyl-D-aspartate receptor (NMDAR) function. Furthermore, upregulation of *GRINA* has been consistently observed in various psychiatric diseases in human subjects (Schmeisser et al., 2012). Cortactin, encoded by the other noteworthy gene, is a



**FIGURE 2** Location diagram of *SHANK2* variants identified in this study. The *SHANK2* gene is located on chromosome 11q13.3. The genomic structure of *SHANK2* is outlined in the middle diagram. The bottom cartoon shows the domains of human *SHANK2* peptide. Variants identified in this study are mapped onto the gene and protein domains. Ank, ankyrin repeats; SH3, Src homology 3; PDZ, PSD95/DLG/ZO1; PRO, proline-rich region; SAM, sterile alpha motif.



**FIGURE 3** Minigene splicing assay for NM\_012309.5:c.2198-1G>A variant in *SHANK2*. (A) Schematic representation of hybrid minigenes used in the assay. (B) The plasmids used in this assay were verified by Sanger sequencing. (C) Gel electrophoresis of RT-PCR products. (D) Sanger sequencing revealed that the product of minigene-MT exhibited skipping of the first 17 bp of exon 19. E, exon.

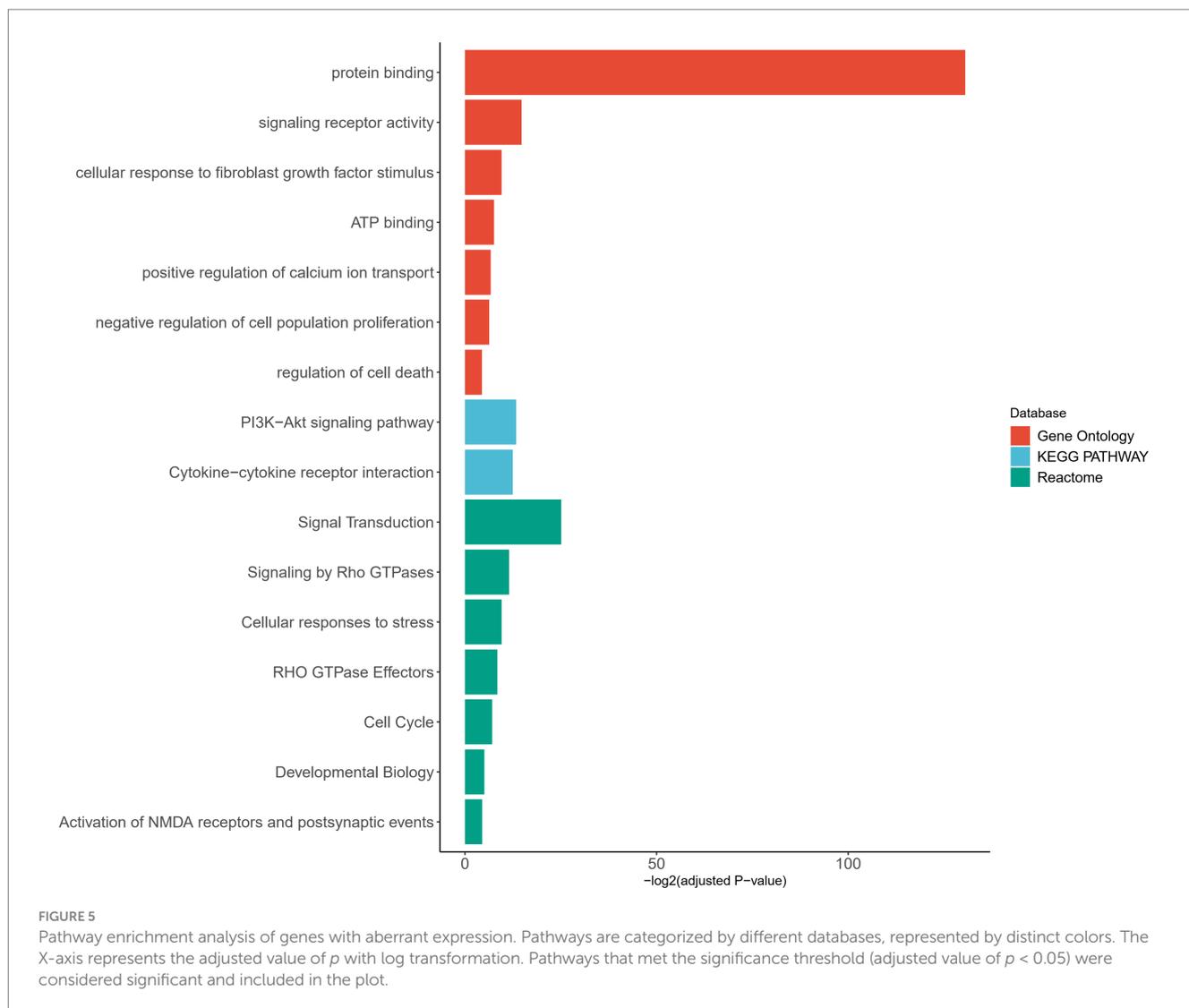


monomeric protein located in the cytoplasm. Previous research in mice has revealed that Shank proteins play a crucial role in maintaining the stability of the spine actin cytoskeleton through their C-terminal cortactin binding site. Knockdown of *Shank* has been shown to significantly reduce the levels of cortactin in spines and increase the mobility of spine cortactin, as measure by single-molecule tracking photoactivated localization microscopy. These findings suggest that Shank proteins are involved in recruiting and stabilizing cortactin at the synapse, thus contributing to synaptic structure and function (Mac Gillavry et al., 2016). Our findings provides further evidence that *SHANK2* disruption can lead to molecular changes related to glutamate signaling and cytoskeletal dynamics, which may contribute to the neurodevelopmental phenotypes. However, it is important to acknowledge that expanding the sample size by including more patients would be highly beneficial. This approach would lead to a more comprehensive understanding of the spectrum of gene expression abnormalities related to this disorder. Additionally, it would also facilitate the identification of additional *SHANK2* variants.

While NDDs caused by *SHANK2* variants exhibit autosomal dominant inheritance, the severity and specific behavioral phenotypes observed in individuals display a high degree of variability, including

possible incomplete penetrance. In our clinical cohort, proband 3 (individual II-1 from family 3) who carried a NM\_012309.5:c.178C>T p.(Arg60Cys) variant in *SHANK2* (Table 1; Supplementary Figure S1). Interestingly, this specific variant in *SHANK2* corresponds to the R12C alteration in the *SHANK3* SPN domain, which has been previously implicated as a potential pathogenic variant in AD patients (Leblond et al., 2014; Sasaki et al., 2020). However, it is important to note that this variant we discovered was inherited from his father, who exhibits no clinical phenotype. The combined Annotation Dependent Depletion (CADD) score is 24.4 and multiple in-silico programs consistently predicted the deleterious effect (Table 1). However, its REVEL score is only 0.36, which categorizes it as “Uncertain” (Ioannidis et al., 2016). Therefore, this variant was classified as VUS according to the ACMG criteria (Table 1; Supplementary Figure S1).

This study represents a preliminary investigation into the transcriptional changes associated with *SHANK2* variants and NDDs. A clear limitation is that the ideal neural tissues were not examined due to clinical inaccessibility and ethical considerations. Instead, we analyzed the more readily accessible peripheral blood nucleated cells, hoping to uncover valuable insights. Pathway enrichment analysis of the differentially expressed genes did reveal associations



with synaptic function and neuronal signaling. While we cannot explain the underlying mechanisms for these findings, we speculate that they may be related to mechanisms such as neuroendocrine signaling, albeit without direct evidence. Additionally, our study is limited by the small cohort size, comprising only three patients harboring two LOF variants. Moreover, patient II-1 from family 2 in this study inherited the variant from his affected mother. Considering the substantial phenotypic variability associated with such disorders, the inclusion of only three patients in our study is insufficient to draw definitive conclusions. Therefore further expansion of the patient cohort is warranted to obtain more robust and comprehensive data.

In summary, our study contributes to the current understanding of the genetics and clinical manifestations associated with *SHANK2* variants. Through the integration of WES and RNA-seq analyses, we have gained significant insights into the pathogenicity of two novel *SHANK2* variants identified in three patients diagnosed with ID. These findings enhance our understanding of the molecular mechanisms underlying *SHANK2*-related disorders and have implications for clinical diagnosis and management in affected individuals.

## Data availability statement

The raw sequence data reported in this study have been deposited in the Genome Sequence Archive (Chen et al., 2021) in National Genomics Data Center (CNCB-NGDC Members and Partners, 2022), China National Center for Bioinformatics/Beijing Institute of Genomics, Chinese Academy of Sciences (GSA-Human: HRA005493) that are publicly accessible at <https://ngdc.cncb.ac.cn/gsa-human>.

## Ethics statement

The studies involving humans were approved by the Ethics Committee of the Second Affiliated Hospital of Chongqing Medical University. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

TABLE 2 Pathogenic variants identified in the *SHANK2* gene.

Patient index	cDNA change	AA change	Function	Phenotype	References
1	arr[GRCh37]11q13.3q13.4(70388251_70506824) × 1		LOF	ID, ASD	Berkel et al. (2010)
2	arr[GRCh37]11q13.4(chr11:70476560_70543234) × 1		LOF	ID, ASD	Berkel et al. (2010)
3	NM_012309.5: c.2521C>T	p.(Arg841*)	LOF	ID, ASD	Berkel et al. (2010)
4	arr[GRCh37]11q13.4(70476810_70542984) × 1		LOF	ID, ASD	Pinto et al. (2010)
5	arr[GRCh37]11q13.2q13.4(67768754_71286835) × 1		LOF	ID, ASD	Wischmeijer et al. (2011)
6	arr[GRCh37]11q13.3q13.4(70399609_70868619) × 1		LOF	ID, ASD	Leblond et al. (2012)
7	arr[GRCh37]11q13.3q13.4(70253389_72098888) × 1		LOF	ID, ASD	Leblond et al. (2014)
8	arr[GRCh37]11q13.2q13.4(67799160_71304541) × 1		LOF	ID	Marcou et al. (2017)
9	NM_012309.5: c.1896dupA	p.(Asp633Argfs*3)	LOF	ID	Bowling et al. (2017)
10	NM_133266.5: c.87C>G	p.(Tyr29*)	LOF	ID, ASD	Guo et al. (2018)
11	NM_133266.5: c.2540_2541del	p.(Ser847*)	LOF	ASD	Zhou et al. (2019)
12	NM_133266.5: c.1322delT	p.(Ile441Thsfs*8)	LOF	ID, ASD	Caumes et al., 2020
13	NM_133266.5: c.132dupA	p.(Asp45Argfs*3)	LOF	ID, ASD	Caumes et al. (2020)
14	NM_133266.5: c.334C>T	p.(Gln112*)	LOF	ID	Doddato et al. (2022)
15	NM_012309.5: c.2198-1G>A	p.Pro734Glyfs*22	LOF	ID	This study
16	NM_012309.5: c.2310dupT	p.(Lys771*)	LOF	ID	This study
17	NM_012309.5: c.2310dupT	p.(Lys771*)	LOF	ID	This study

Sixteen variants were collected from a total of 17 patients diagnosed with ID and/or ASD, including data from published literature and the current study. LOF, loss-of-function variants. An asterisk (\*) denotes variants resulting in premature termination codons.

## Author contributions

YW: Funding acquisition, Investigation, Methodology, Writing – original draft. WL: Investigation, Methodology, Writing – original draft. BT: Conceptualization, Resources, Software, Supervision, Writing – review & editing. SL: Conceptualization, Supervision, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Scientific and Technological Research Program of Chongqing 277 Municipal Education Commission (grant no. KJQN202200420) and Shenzhen science and technology innovation committee foundation (JCYJ20180305164623090).

## Acknowledgments

The authors thank the patients and their family members for their participation in this study.

## References

- Barish, S., Barakat, T. S., Michel, B. C., Mashtalir, N., Phillips, J. B., Valencia, A. M., et al. (2020). BICRA, a SWI/SNF complex member, is associated with BAF-disorder related phenotypes in humans and model organisms. *Am. J. Hum. Genet.* 107, 1096–1112. doi: 10.1016/j.ajhg.2020.11.003
- Berkel, S., Marshall, C. R., Weiss, B., Howe, J., Roeth, R., Moog, U., et al. (2010). Mutations in the *SHANK2* synaptic scaffolding gene in autism spectrum disorder and mental retardation. *Nat. Genet.* 42, 489–491. doi: 10.1038/ng.589
- Bowling, K. M., Thompson, M. L., Amaral, M. D., Finnila, C. R., Hiatt, S. M., Engel, K. L., et al. (2017). Genomic diagnosis for children with intellectual disability and/or developmental delay. *Genome Med.* 9:43. doi: 10.1186/s13073-017-0433-1
- Brechtmann, F., Mertes, C., Matuszewska, A., Yopez, V. A., Avsec, Z., Herzog, M., et al. (2018). OUTRIDER: a statistical method for detecting aberrantly expressed genes in RNA sequencing data. *Am. J. Hum. Genet.* 103, 907–917. doi: 10.1016/j.ajhg.2018.10.025

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2023.1275421/full#supplementary-material>

- Bu, D., Luo, H., Huo, P., Wang, Z., Zhang, S., He, Z., et al. (2021). KOBAS-i: intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. *Nucleic Acids Res.* 49, W317–W325. doi: 10.1093/nar/gkab447
- Caumes, R., Smol, T., Thuillier, C., Balerdi, M., Lestienne-Roche, C., Manouvrier-Hanu, S., et al. (2020). Phenotypic spectrum of SHANK2-related neurodevelopmental disorder. *Eur. J. Med. Genet.* 63:104072. doi: 10.1016/j.ejmg.2020.104072
- Chen, T., Chen, X., Zhang, S., Zhu, J., Tang, B., Wang, A., et al. (2021). The genome sequence archive family: toward explosive data growth and diverse data types. *Genomics Proteomics Bioinformatics* 19, 578–583. doi: 10.1016/j.gpb.2021.08.001
- Chen, S., Fragoza, R., Klei, L., Liu, Y., Wang, J., Roeder, K., et al. (2018a). An interactome perturbation framework prioritizes damaging missense mutations for developmental disorders. *Nat. Genet.* 50, 1032–1040. doi: 10.1038/s41588-018-0130-z
- Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018b). Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560
- CNCB-NGDC Members and Partners (2022). Database resources of the National Genomics Data Center, China National Center for bioinformatics in 2022. *Nucleic Acids Res.* 50, D27–D38. doi: 10.1093/nar/gkab951
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics* 29, 15–21. doi: 10.1093/bioinformatics/bts635
- Doddato, G., Fabbiani, A., Scandurra, V., Canitano, R., Mencarelli, M. A., Renieri, A., et al. (2022). Identification of a novel SHANK2 pathogenic variant in a patient with a neurodevelopmental disorder. *Genes* 13:688. doi: 10.3390/genes13040688
- Ellard, S., Baple, E. L., Callaway, A., Berry, I., Forrester, N., Turnbull, C., et al. (2020). ACGS best practice guidelines for variant classification in rare disease 2020: Association for Clinical Genomic Science (ACGS), 2020. Available at: <https://www.acgs.uk.com/quality/best-practice-guidelines/#VariantGuidelines>
- Ewels, P., Magnusson, M., Lundin, S., and Kaller, M. (2016). Multi QC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32, 3047–3048. doi: 10.1093/bioinformatics/btw354
- Guilmatre, A., Huguet, G., Delorme, R., and Bourgeron, T. (2014). The emerging role of SHANK genes in neuropsychiatric disorders. *Dev. Neurobiol.* 74, 113–122. doi: 10.1002/dneu.22128
- Guo, H., Wang, T., Wu, H., Long, M., Coe, B. P., Li, H., et al. (2018). Inherited and multiple de novo mutations in autism/developmental delay risk genes suggest a multifactorial model. *Mol. Autism* 9:64. doi: 10.1186/s13229-018-0247-z
- Hu, W. F., Chahrouh, M. H., and Walsh, C. A. (2014). The diverse genetic landscape of neurodevelopmental disorders. *Annu. Rev. Genomics Hum. Genet.* 15, 195–213. doi: 10.1146/annurev-genom-090413-025600
- Ioannidis, N. M., Rothstein, J. H., Pejaver, V., Middha, S., McDonnell, S. K., Baheti, S., et al. (2016). REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. *Am. J. Hum. Genet.* 99, 877–885. doi: 10.1016/j.ajhg.2016.08.016
- Kremer, L. S., Bader, D. M., Mertes, C., Kopajtic, R., Pichler, G., Iuso, A., et al. (2017). Genetic diagnosis of Mendelian disorders via RNA sequencing. *Nat. Commun.* 8:15824. doi: 10.1038/ncomms15824
- Leblond, C. S., Heinrich, J., Delorme, R., Proepper, C., Betancur, C., Huguet, G., et al. (2012). Genetic and functional analyses of SHANK2 mutations suggest a multiple hit model of autism spectrum disorders. *PLoS Genet.* 8:e1002521. doi: 10.1371/journal.pgen.1002521
- Leblond, C. S., Nava, C., Polge, A., Gauthier, J., Huguet, G., Lumbroso, S., et al. (2014). Meta-analysis of SHANK mutations in autism Spectrum disorders: a gradient of severity in cognitive impairments. *PLoS Genet.* 10:e1004580. doi: 10.1371/journal.pgen.1004580
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Liu, L., Feng, X., Liu, S., Zhou, Y., Dong, X., Yao, H., et al. (2022). Whole-genome sequencing combined RNA-sequencing analysis of patients with mutations in SET binding protein 1. *Front. Neurosci.* 16:980000. doi: 10.3389/fnins.2022.980000
- Mac Gillavry, H. D., Kerr, J. M., Kassner, J., Frost, N. A., and Blanpied, T. A. (2016). Shank-cortactin interactions control actin dynamics to maintain flexibility of neuronal spines and synapses. *Eur. J. Neurosci.* 43, 179–193. doi: 10.1111/ejn.13129
- Marcou, C. A., Studinski, J., Murphree, M. I., Kirmani, S., and Hoppman, N. L. (2017). De novo 11q deletion including SHANK2 in a patient with global developmental delay. *Am. J. Med. Genet. A* 173, 801–805. doi: 10.1002/ajmg.a.38075
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernysky, A., et al. (2010). The genome analysis toolkit: a map reduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi: 10.1101/gr.107524.110
- McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R. S., Thormann, A., et al. (2016). The Ensembl variant effect predictor. *Genome Biol.* 17:122. doi: 10.1186/s13059-016-0974-4
- Monteiro, P., and Feng, G. (2017). SHANK proteins: roles at the synapse and in autism spectrum disorder. *Nat. Rev. Neurosci.* 18, 147–157. doi: 10.1038/nrn.2016.183
- Morris-Rosendahl, D. J., and Crocq, M. A. (2020). Neurodevelopmental disorders—the history and future of a diagnostic concept. *Dialogues Clin. Neurosci.* 22, 65–72. doi: 10.31887/DCNS.2020.22.1/macrocq
- Pan, X., Liu, S., Liu, L., Zhang, X., Yao, H., and Tan, B. (2022). Case report: exome and RNA sequencing identify a novel de novo missense variant in HNRNP1 in a Chinese patient with au-Kline syndrome. *Front. Genet.* 13:853028. doi: 10.3389/fgene.2022.853028
- Parenti, I., Rabaneda, L. G., Schoen, H., and Novarino, G. (2020). Neurodevelopmental disorders: from genetics to functional pathways. *Trends Neurosci.* 43, 608–621. doi: 10.1016/j.tins.2020.05.004
- Peymani, F., Farzeen, A., and Prokisch, H. (2022). RNA sequencing role and application in clinical diagnostic. *Pediatr. Investig.* 6, 29–35. doi: 10.1002/ped4.12314
- Pinto, D., Pagnamenta, A. T., Klei, L., Anney, R., Merico, D., Regan, R., et al. (2010). Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* 466, 368–372. doi: 10.1038/nature09146
- Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. ACMG laboratory quality assurance committee. *Genet. Med.* 17, 405–424. doi: 10.1038/gim.2015.30
- Sasaki, K., Kojitani, N., Hirose, H., Yoshihama, Y., Suzuki, H., Shimada, M., et al. (2020). Shank 2 binds to aPKC and controls tight junction formation with rap 1 signaling during establishment of epithelial cell polarity. *Cell Rep.* 31:107407. doi: 10.1016/j.celrep.2020.02.088
- Schmeisser, M. J., Ey, E., Wegener, S., Bockmann, J., Stempel, A. V., Kuebler, A., et al. (2012). Autistic-like behaviours and hyperactivity in mice lacking pro SAP1/Shank 2. *Nature* 486, 256–260. doi: 10.1038/nature11015
- Sheng, M., and Kim, E. (2000). The Shank family of scaffold proteins. *J. Cell Sci.* 113, 1851–1856. doi: 10.1242/jcs.113.11.1851
- Staufner, C., Peters, B., Wagner, M., Alameer, S., Barić, I., Broué, P., et al. (2020). Defining clinical subgroups and genotype-phenotype correlations in NBAS-associated disease across 110 patients. *Genet. Med.* 22, 610–621. doi: 10.1038/s41436-019-0698-4
- Toro, R., Konyukh, M., Delorme, R., Leblond, C., Chaste, P., Fauchereau, F., et al. (2012). Key role for gene dosage and synaptic homeostasis in autism spectrum disorders. *Trends Genet.* 26, 363–372. doi: 10.1016/j.tig.2010.05.007
- Trivedi, U. H., Cézard, T. E., Bridgett, S., Montazam, A., Nichols, J., Blaxter, M., et al. (2014). Quality control of next-generation sequencing data without a reference. *Front. Genet.* 5:111. doi: 10.3389/fgene.2014.00111
- Wischmeijer, A., Magini, P., Giorda, R., Gnoli, M., Ciccone, R., Cecconi, I., et al. (2011). Olfactory receptor-related Duplicons mediate a microdeletion at 11q13.2q13.4 associated with a syndromic phenotype. *Mol. Syndromol.* 1, 176–184. doi: 10.1159/000322054
- Won, H., Lee, H. R., Gee, H. Y., Mah, W., Kim, J. I., Lee, J., et al. (2012). Autistic-like social behaviour in Shank 2-mutant mice improved by restoring NMDA receptor function. *Nature* 486, 261–265. doi: 10.1038/nature11208
- Yepez, V. A., Mertes, C., Muller, M. F., Klaproth-Andrade, D., Wachutka, L., Fresard, L., et al. (2021). Detection of aberrant gene expression events in RNA sequencing data. *Nat. Protoc.* 16, 1276–1296. doi: 10.1038/s41596-020-00462-5
- Yoo, J., Bakes, J., Bradley, C., Collingridge, G. L., and Kaang, B. K. (2014). Shank mutant mice as an animal model of autism. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369:20130143. doi: 10.1098/rstb.2013.0143
- Zablotsky, B., Black, L. I., Maenner, M. J., Schieve, L. A., Danielson, M. L., Bitsko, R. H., et al. (2019). Prevalence and trends of developmental disabilities among children in the United States: 2009–2017. *Pediatrics* 144:e20190811. doi: 10.1542/peds.2019-0811
- Zhou, W. Z., Zhang, J., Li, Z., Lin, X., Li, J., Wang, S., et al. (2019). 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. doi: 10.1002/humu.23724