



## OPEN ACCESS

## EDITED BY

Nahid Tayebi,  
Genetics and Molecular Biology Branch  
(NHGRI), United States

## REVIEWED BY

Jens Lichtenberg,  
National Institutes of Health (NIH), United States  
Grisel Lopez,  
National Human Genome Research Institute  
(NIH), United States

## \*CORRESPONDENCE

Naveed Wasif

✉ naveedwasif@gmail.com

Saadullah Khan

✉ saadkhanwazir@gmail.com;

✉ saad@kust.edu.pk

<sup>†</sup>These authors have contributed equally to this work

RECEIVED 17 February 2023

ACCEPTED 28 April 2023

PUBLISHED 25 May 2023


## CITATION

Muhammad N, Hussain SI, Rehman ZU, Khan SA, Jan S, Khan N, Muzammal M, Abbasi SW, Kakar N, Rehman ZU, Khan MA, Mirza MU, Muhammad N, Khan S and Wasif N (2023) Autosomal recessive variants c.953A>C and c.97-1G>C in NSUN2 causing intellectual disability: a molecular dynamics simulation study of loss-of-function mechanisms. *Front. Neurol.* 14:1168307. doi: 10.3389/fneur.2023.1168307

## COPYRIGHT

© 2023 Muhammad, Hussain, Rehman, Khan, Jan, Khan, Muzammal, Abbasi, Kakar, Rehman, Khan, Mirza, Muhammad, Khan and Wasif. 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.

# Autosomal recessive variants c.953A>C and c.97-1G>C in NSUN2 causing intellectual disability: a molecular dynamics simulation study of loss-of-function mechanisms

Nazif Muhammad<sup>1†</sup>, Syeda Iqra Hussain<sup>1†</sup>, Zia Ur Rehman<sup>2</sup>, Sher Alam Khan<sup>1</sup>, Samin Jan<sup>1</sup>, Niamatullah Khan<sup>1</sup>, Muhammad Muzammal<sup>3</sup>, Sumra Wajid Abbasi<sup>4</sup>, Naseebullah Kakar<sup>5,6</sup>, Zia Ur Rehman<sup>1</sup>, Muzammil Ahmad Khan<sup>3</sup>, Muhammad Usman Mirza<sup>7</sup>, Noor Muhammad<sup>1</sup>, Saadullah Khan <sup>1\*</sup> and Naveed Wasif<sup>8,9\*</sup>

<sup>1</sup>Department of Biotechnology and Genetic Engineering, Kohat University of Science and Technology (KUST), Kohat, Khyber Pakhtunkhwa, Pakistan, <sup>2</sup>Department of General Medicine, Northwest General Hospital & Research Center, Peshawar, Khyber Pakhtunkhwa, Pakistan, <sup>3</sup>Gomal Center of Biochemistry and Biotechnology, Gomal University, D.I.Khan, Khyber Pakhtunkhwa, Pakistan, <sup>4</sup>NUMS Department of Biological Sciences, National University of Medical Sciences, The Mall, Rawalpindi, Punjab, Pakistan, <sup>5</sup>Department of Biotechnology, Faculty of Life Sciences and Informatics, BUITEMS, Quetta, Pakistan, <sup>6</sup>Institute of Human Genetics, Universitätsklinikum Schleswig-Holstein, Lübeck, Germany, <sup>7</sup>Department of Chemistry and Biochemistry, University of Windsor, Windsor, ON, Canada, <sup>8</sup>Institute of Human Genetics, Ulm University, and Ulm University Medical Center, Ulm, Germany, <sup>9</sup>Institute of Human Genetics, University Hospital Schleswig-Holstein, Kiel, Germany

**Introduction:** Intellectual disability (ID) is a clinically and genetically heterogeneous disorder. It drastically affects the learning capabilities of patients and eventually reduces their IQ level below 70.

**Methods:** The current genetic study ascertained two consanguineous Pakistani families suffering from autosomal recessive intellectual developmental disorder-5 (MRT5). We have used exome sequencing followed by Sanger sequencing to identify the disease-causing variants.

**Results and discussion:** Genetic analysis using whole exome sequencing in these families identified two novel mutations in the *NSUN2* (NM\_017755.5). Family-A segregated a novel missense variant c.953A>C; p.Tyr318Ser in exon-9 of the *NSUN2*. The variant substituted an amino acid Tyr318, highly conserved among different animal species and located in the functional domain of *NSUN2* known as "SAM-dependent methyltransferase RsmB/NOP2-type". Whereas in family B, we identified a novel splice site variant c.97-1G>C that affects the splice acceptor site of *NSUN2*. The identified splice variant (c.97-1G>C) was predicted to result in the skipping of exon-2, which would lead to a frameshift followed by a premature stop codon (p. His86Profs\*16). Furthermore, it could result in the termination of translation and synthesis of dysfunctional protein, most likely leading to nonsense-mediated decay. The dynamic consequences of *NSUN2* missense variant was further explored together with wildtype through molecular dynamic simulations, which uncovered the disruption of *NSUN2* function due to a gain in structural

flexibility. The present molecular genetic study further extends the mutational spectrum of *NSUN2* to be involved in ID and its genetic heterogeneity in the Pakistani population.

#### KEYWORDS

consanguinity, *NSUN2* gene, intellectual disability, novel variants, molecular dynamics simulation

## 1. Introduction

Intellectual disability (ID) is a severe manifestation of the central nervous system leading to drastically reduced general mental abilities, intelligent working, and adaptive behavior compared with people of the same age group, gender, and sociocultural setup. The primary distinguishing characteristics of intellectual disability involve IQ (intelligence quotient) <70, impairment in at least two adaptive abilities, and disease onset before 18 years of age (1, 2). The epidemiological estimate indicates that 1–3% of the general population is affected by intellectual disability. In addition to consanguinity, developing countries are also facing the issues of malnutrition, lack of health facilities, unhygienic environment, and cultural deprivation as contributing risk factors of intellectual disability. Furthermore, the recent literature survey has reported over 700 genetic entities involved in syndromic and non-syndromic ID (3) and are transmitted in an autosomal dominant, recessive, X-Linked, or mitochondrial fashion (Orphanet report). The autosomal recessive inheritance of intellectual disability is relatively rare and accounts for <12% of cases of intellectual disabilities (4). Physiologically, the reported ID genes are involved in various cellular signaling cascades, inter-neuronal connectivity, proliferation, and migration, and the genetic or epigenetic control over transcription and translation of the genes (5).

The *NSUN2* (NOP2/Sun RNA methyltransferase 2) gene is located at 5p15, comprising 19 exons. The gene transcribes a 2.1-kb mRNA molecule expressed in brain tissue and generates a methyltransferase enzyme. The tRNA precursors' cytosine at positions 34 and 48 is methylated to 5-methylcytosine (m<sup>5</sup>C) (6). This modification might promote tRNA stability and mRNA export (7–9), and enzymatic failure is associated with many diseases, such as autism spectrum disorder (10), depression (11), Dubowitz syndrome (12, 13), and non-syndromic autosomal recessive intellectual disability (ARID) (14, 15), and is differentially expressed in various cancers (16–19). Evidence for the significance of *NSUN2*-mediated m<sup>5</sup>C in ID has been recently reported in Pakistani, Iranian, Qatari, and Lebanese families (9, 13, 20–22). *NSUN2* is shown to be localized to the nucleoli of Purkinje cells, and it has been found that these nucleoli are often situated between or in close proximity to dense heterochromatic regions (9).

Over the past decade, advances in NGS techniques and machine learning development have sped up gene identification and increased the diagnostic efficiency of rare genetic disorders, revolutionizing the healthcare systems (23). This genetic study aimed at identifying disease-causing variants in two unrelated Pashtun families of Pakistani origin containing five affected

members with impairments in adaptive behavior and cognitive functioning. We have used exome sequencing and Sanger sequencing to identify the genetic variations.

With the advancement in molecular modeling campaigns, molecular dynamics simulations are considered an adequate method to explore the dynamic effects on protein structure due to mutations (16, 24–27), which consented to investigate the conformational characteristics (28, 29). Our findings after molecular dynamics (MD) simulations delineate the structural insights of novel disease-causing variant and provide a molecular illustration of plausible aberrations caused by mutation.

## 2. Materials and methods

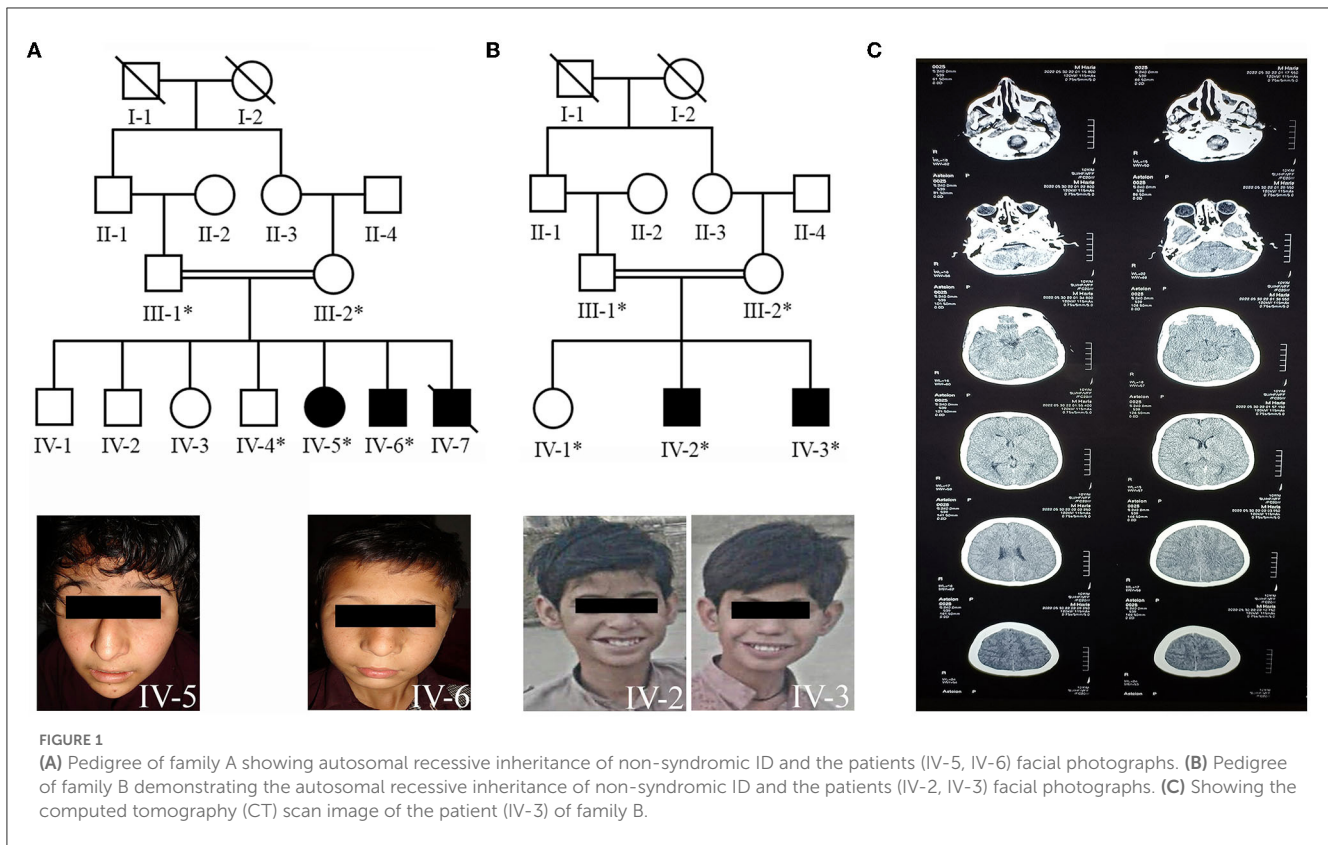
### 2.1. Sample collection

In this study, two consanguineous families (family A and family B), segregating autosomal recessive intellectual disability, were ascertained from Upper Dir and district Karak of Khyber Pakhtunkhwa (KP) province in Pakistan. Familial-based sampling approach recruited two individuals affected with ID, and three unaffected members, including parents from each family (Figures 1A, B). The institutional ethical review board of Kohat University of Science and Technology, Kohat, approved this study, and patients were enrolled after written informed consent. Blood samples were collected in EDTA Vacutainer tubes, and DNA was extracted using the standard phenol–chloroform method.

### 2.2. Exome sequencing and variant calling in family A

Exome sequencing was performed for a paired-end library using the Nextera DNA Exome kit (Illumina Inc., USA) as per the manufacturer's instructions. Approximately 100 ng genomic DNA of an index patient (IV-5) was used for the enzymatic fragmentation, subsequently performing unique adaptors ligation (dual indexes) and exome enrichment. Exome sequencing (75 × 2 paired-end) was performed using NextSeq<sup>®</sup> 550/500 High Output Reagent Cartridge v2 kit on Illumina NextSeq500 instrument (Illumina, San Diego, CA, USA).

The filtration of primary data was carried out by the Illumina real-time analysis (RTA) software v1.8. Afterward, the BWA-SW alignment algorithm (30) was used to map the reads to the human reference genome build GRCh37/hg19 (<http://www.genome.ucsc.edu/>). Using Samtools v.0.0.1.9 (31), the subsequent



sequence alignment map file (SAM) was converted into a binary alignment map (BAM) format. Picard tools (32) were used to improve the read quality, and Genome Analysis Toolkit (GATK) was used for realignment and base quality score recalibration (33). The calling of single nucleotide polymorphisms (SNPs) and short insertions/deletions (INDELS) was performed by Platypus (34), Haplotype Caller (35), and Mpileup programs (31), and further filtration was carried out through variant quality score calibration (VQSR) using GATK (33). For the coverage analysis of CNV detection, CNMOPS and ExomeDepth algorithms were utilized (36, 37). In addition, the COMBINE and FUNC algorithms (varpipe\_v2.26, <https://varbank.ccg.uni-koeln.de/>) combined the data and the annotation of functional variants.

### 2.3. Data analysis and variant filtration

Exome sequencing data analysis and filtration of variants were performed using the exome analysis pipeline “Varvis” (Limbus Medical Technologies GmbH). The mean coverage of the data at 20X and 10X of the targeted bases was >92 and 96.6%, respectively. In addition, we performed exclusion mapping of homozygous or compound heterozygous variants in OMIM genes associated with autosomal recessive intellectual disability. We analyzed the exome sequencing data for variants with coverage of >6 reads and an allele frequency ≥75% in the affected individual. Furthermore, a minor allele frequency (MAF) <0.5% in the gnomAD database was observed. The filtration criteria identified 10 homozygous variants [*RPL5*: c.891C>A,

p.(Ser297Arg), *DYNC2II*: c.845C>T, p.(Ser282Leu), *HEMK1*: c.157G>C, p.(Glu53Gln), *ITIH1*: c.2515G>A, p.(Gly839Ser), *ITIH1*: c.2510\_2511insTAGAAGTTTCTGA, p.(Ile838Argfs\*4), *PFKP*: c.1490C>T, p.(Thr497Met), *USP19*: c.1447G>T, p.(Gly483Cys), *VAVI*: c.2411A>C, p.(Lys804Thr), *MBTPS1*: c.493G>C, p.(Ala165Pro)], including a homozygous variant (c.953A>C, p.Tyr318Ser) in *NSUN2*. An in-house intellectual disability gene panel, which includes autosomal dominant, recessive, and X-linked genes involved in syndromic and non-syndromic intellectual disability, did not reveal any other rare variant except the *NSUN2* variant.

### 2.4. Exome sequencing and variant calling in family B

Genomic DNA was extracted from the patient’s blood samples. First, the exonic regions of all 22,000 human genes were captured by the xGen Exome Research Panel v2 (Integrated DNA Technologies, Coralville, Iowa, USA). After capture, NovaSeq 600 was used to sequence all captured regions (Illumina, San Diego, CA, USA). We acquired ≥20X coverage in >98.9% and ≥10X coverage in >99.1% of target sequences.

Following the sequencing, the data were analyzed using open-source bioinformatics tools and proprietary software. bcl2fastq v2.20.0.422 (<https://emea.support.illumina.com/downloads/bcl2fastq-conversion-software-v2-20.html>) was used to convert and demultiplex base call (BCL) sequence files to FASTQ files. Variant calling and annotation followed the alignment of the

sequencing data to the GRCh37/hg19 human reference genome was carried out using BWA-mem 0.7.17 (arXiv:1303.3997 [q-bio.GN]) to generate BAM files. BAM files were processed using GATK best practices (GATK v.3.8, [broadinstitute.org](https://broadinstitute.org)) for single nucleotide variants (SNV) and small insertions/deletions (indel) variant calling to generate VCF files (33, 38). For copy number variant (CNV) calling based on depth-of-coverage (DOC) data, Conifer (39) and 3bCNV (<https://3billion.io/resources>) are used. The region of homozygosity (ROH) was mapped from the VCF file using AutoMap v1.2 (40).

One of the in-house tools, EVIDENCE, was designed to select variants based on ACMG guidelines and each patient's phenotype. Variant filtration, categorization, and similarity score for the patient's phenotype are three significant steps in this approach. For allele frequency estimation, a genome aggregation database (gnomAD, <http://gnomad.broadinstitute.org/>) and a 3-billion genome database were utilized in the first step. According to ACMG guidelines, gene variations with more than 5% allele frequency were filtered out. Next, the VarSome (41), Human Gene Mutation Database (HGMD) Professional 2022.1, Database of Single Nucleotide Polymorphisms (dbSNP), and ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>) were utilized for the evaluation of variants. Then, the assessment of each variant concerning disease phenotype was carried out using the ACMG guidelines (42). Finally, in the third step, the patient's clinical phenotypes were converted to standardized human phenotype ontology terms (<https://hpo.jax.org/>) and retrieved to determine the degree of similarity (43, 44) with each of 7,000 rare genetic diseases (<https://omim.org/> and <https://www.orpha.net/consor/cgi-bin>). According to the ACMG guideline, the similarity score between each patient's phenotype and symptoms related to that disease caused by priority variations varied from 0 to 10. The filtration criteria revealed four rare variants (*RELN*: c.8843+3A>C, *PMM2*: c.348-5C>A, *MADCAM1*: c.784\_785insAGGAGCCTCCCGACACCACCTCCCAGGAGCCTCCCGACACCACCTCCC, p.Ser261\_Pro262insGlnGluProProAspThrThrSerGlnGluProProAspThrThrSer, *NSUN2*: c.97-1G>C).

## 2.5. Segregation analysis and sequence conservation alignment

The filtered variants were then subjected to Sanger sequencing to confirm variant segregation with disease phenotype in the whole family with the available DNA samples. Primers (flanking the variant regions) were designed using the online Primer3 software. We used the following primer sets for the missense variant (c.953A>C) “forward primer: 5'-GGACTGGAATGTATGATACCAA-3' and reverse primer: 5'-AATAGAACGGTGGTGTGAGG-3'” and, for splice acceptor site variant (c.97-1G>C) “forward primer: 5'-AGGTGTAGGGCTAGAGTTCTG-3' and reverse primer: 5'-CTCAATGCTTCCTGAATCC-3'”.

Multiple protein sequence alignment was performed through the Clustal Omega tool (Embl-Ebi Clustal Omega) to check the conservation of identified missense variant (<https://www.ebi.ac.uk/Tools/msa/clustalo/> (accessed August 30, 2022)). The pathogenicity

of the novel missense variant of *NSUN2* was predicted through SIFT (45), PolyPhen 2.0 (46), and I-Mutant 3.0 (47).

## 2.6. Molecular modeling and protein stability predictions

Although no x-ray-resolved crystal structure of *NSUN2* is available in the protein data bank, the homology modeling was performed using human-*NSUN6* complexed with tRNA and S-adenosylmethionine (SAM). Both share a strikingly similar architecture; therefore, the user-defined template method of SWISS-MODEL was executed to build the wild-type (*wt*) and mutant (*mt*) *NSUN2* model, which extracts initial structural information from the template structure (48). As the inherent structural elucidation relies on homology model, the reliability of the *NSUN2* model was evaluated using MolProbity (49) and ProSA (50) and further refined and optimized through MD simulations. In addition, protein function and the stability effect upon substitution ( $\Delta\Delta G$ ) were assessed using the DUET server (51, 52), a combined computational approach that combines two complementary techniques (mCSM and SDM). Moreover, the thermal stability ( $\Delta\Delta G$ ) rising from vibrational entropy changes ( $\Delta\Delta S$ ) was also predicted using the Elastic Network Contact Model (ENCoM) server (53). Finally, the effect of mutation on the overall structural dynamics of *wt-NSUN2* compared with its wild-type (*wt*) was analyzed through MD simulations, which were performed in two steps: 100 ns MD simulation to refine and optimize the models (*wt* and *mt*), while another 100 ns to analyze the residual fluctuations of *NSUN2* with or without reported mutation. All simulations were executed using AMBER 20 (54) using the same protocol we described elsewhere (25, 55). Furthermore, the figures were drawn using Chimera (56), while structure-based protein multiple sequence alignment of *NSUN* members (1–3, 5–8) was generated by ESPript (57).

## 3. Results

### 3.1. Clinical features of family A

Family A resides in the Upper Dir area of Khyber Pakhtunkhwa. The family was Pashtun by ethnicity, who usually preferred cousin marriages. The family's pedigree consists of four generations, in which the affected individuals (IV-5, IV-6, and IV-7, Figure 1A) born to first-cousin parents appeared in the fourth generation. Before our first sampling visit, patient IV-7 was deceased due to unknown reasons. There was no prior family history of any genetic disease.

Clinical features of all the affected individuals from family A had moderate-to-severe ID with speech cluttering and unable to understand the arithmetic calculations. The patients had relatively small stature, abnormal behavioral features, and developmental delay. However, no epileptic fits, spasticity, hypotonia, or deep tendon reflexes were observed. Both patients exhibited dysmorphic facial features with ptosis, shallow and flat philtrum grooves, thin upper and thick lower lips, high nasal bridges, and protruding teeth in common. A pointed chin, heart-shaped face, and snub nose with



droopy nasal tip were peculiar to the female patient (IV-5). The male patient (IV-6) also had additional dysmorphic facial features, including a square-shaped face, hypertelorism, sparse eyebrows, and a big nose (Figure 1A, IV-5, IV-6).

### 3.2. Clinical features of family B

Family B also belonged to Pashtun ethnic group and resided in District Karak. However, it is unrelated to family A. The presented family pedigree consisted of four generations with two affected individuals (IV-2 and IV-3, Figure 1B) also born to first consanguineous parents.

Clinical features of the affected individuals from family B include moderate-to-severe ID with speech cluttering and facial dysmorphism. The prominent facial dysmorphic features included square-shaped faces, thick scalp hair, ptosis, large ears, shallow philtrum grooves, thin upper and thick lower lips, and prominent teeth (Figure 1B, IV-2, IV-3). They had no counting or recognition abilities. They could not recognize and differentiate among different currency notes. They also had a lack of social interactions and often showed aggressive modes upon irritation. The patient's (IV-3) computed tomography (CT) scan image found evidence of fluid present bilaterally in the maxillary and ethmoid sinuses of the brain (Figure 1C). The additional clinical description of patients in both families is summarized in Supplementary Table 1.

### 3.3. Molecular findings

Genetic analysis using whole exome sequencing performed on the DNA of an affected individual from family A identified 10 homozygous variants, including a missense variant (c.953A>C, p.Tyr318Ser) in *NSUN2* (Figures 2A, B). Sanger sequencing validated and confirmed the segregation of only *NSUN2* variant with the disease phenotype in the family. Furthermore, the zygosity analysis determined that the affected individuals (IV-5 and IV-6) were homozygous (c.953C/C). At the same time, the unaffected parents (III-1 and III-2), as well as the unaffected sibling (IV-4), were heterozygous carriers (c.953A/C) for the identified missense variant (Figure 2C), consistent with an autosomal recessive mode of inheritance.

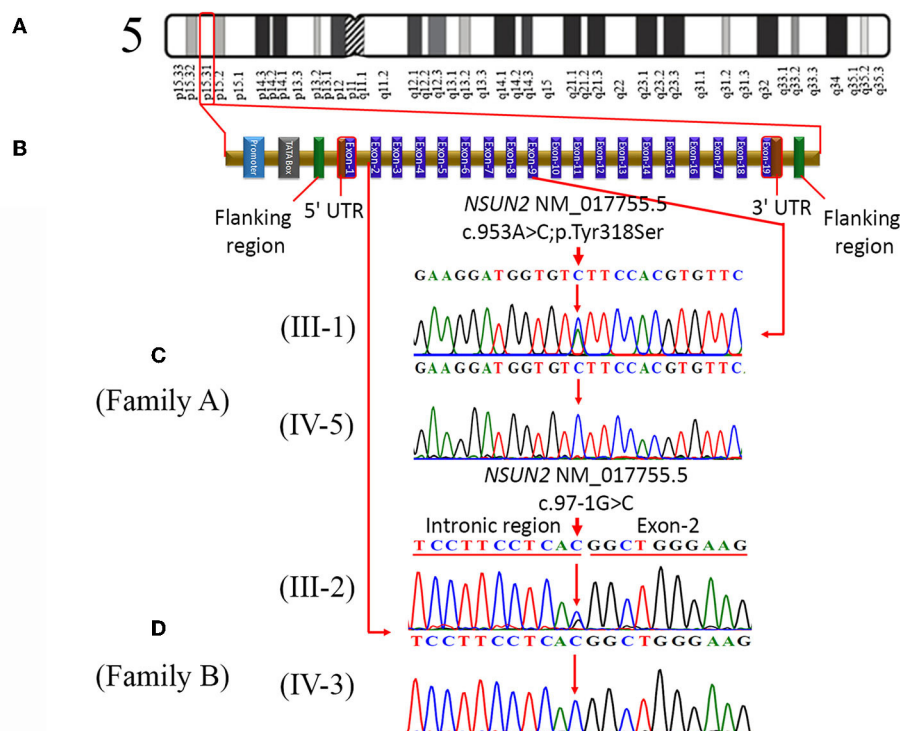
In family B, exome sequencing of a patient (IV-2) revealed a novel homozygous splice site variant (c.97-1G>C) in the *NSUN2* (Figure 2B), along with the other three variants on different chromosomal locations. Likewise, in segregation analysis in family A, the segregation of the *NSUN2* variant with disease phenotype was also confirmed by Sanger sequencing. Zygosity analysis found that the unaffected sibling (IV-1) and both unaffected parents (III-1 and III-2) were heterozygous c.97-1G/C, while both the affected individuals (IV-2 and IV-3) were homozygous c.97-1C/C for splice site variant (Figure 2D). The nomenclature of both variants is based on the transcript NM\_017755.5. The variants identified other than *NSUN2* variants were excluded during the segregation analysis.

### 3.4. Structural elucidation and MD interpretations

SWISS-MODEL generated the *NSUN2* models with high accuracy due to the template belonging to the human *NSUN6*, which revealed a sequence identity of 34.32 and query coverage of 0.61. The model is refined and optimized for 100 ns MD simulations. Later, the homology model evaluations determined 90.21% residues in Ramachandran favored regions, and ProSA determined the reliable z-score (-7.5) as the overall quality model (see Supplementary Figure S1). Like in *NSUN6*, the *NSUN2* revealed a similar Rossmann-fold catalytic core (residues 171–429) and PUA domain (residues 54–147) that accommodates the SAM cofactor when superimposed on *NSUN6/tRNA/SAM*. In addition, the *NSUN* proteins use two catalytic cysteines in the active site, present in conserved motif IV (Cys271 in *NSUN2*) and VI (Cys321 in *NSUN2*) (Figures 3A–C).

In *NSUN6/tRNA/SAM*, the methylation target C72 has profound interactions with *NSUN6* and the methylation factor, SAM. The side chains of Lys248 (190 in *NSUN2*), Asp323 (motif IV; 268 in *NSUN2*), Cys326 (motif IV; 271 in *NSUN2*), Cys373 (motif VI; 321 in *NSUN2*), and Phe458 (422 in *NSUN2*) and the main chain of residue Ser371 (319 in *NSUN2*) are all involved in recognition of C72. The conformations of these C72 and SAM interacting residues were strikingly similar in *NSUN2*. The Tyr318Ser mutation resided in the conserved antiparallel  $\beta$ -sheets structural fold of *NSUNs*, in the vicinity of active site (tRNA/SAM binding site) (Figure 3B; multiple sequence alignment). To examine the dynamic consequences of Tyr318Ser mutation in the active site, the protein structure stability changes ( $\Delta\Delta G$ ) upon mutation in the *mt-NSUN2* model were estimated from various programs. DUET predicted Tyr318Ser in *NSUN2* as destabilizing with a negative free energy change ( $\Delta\Delta G$ ) value of  $-3.694$  kcal/mol. Moreover, the estimated vibrational entropy energy change ( $\Delta\Delta S_{Vib}$ ) and thermal stability ( $\Delta\Delta G$ ), as accessed from ENCoM server (53), also indicated the increase in protein flexibility ( $\Delta\Delta S_{Vib}$  ENCoM:  $1.127$  kcal.mol $^{-1}$ .K $^{-1}$ ;  $\Delta\Delta G$ :  $-4.077$  kcal/mol as destabilizing). The calculated  $\Delta\Delta S$  per residue vibrational entropy changes indicated the region (residues 260–330), which gained flexibility upon Tyr318Ser substitution, while Ser318 revealed  $\Delta\Delta S$  of  $-0.254$  kcal.mol $^{-1}$ .K $^{-1}$  (see Supplementary Figure S2).

Tyr318 was present in critical central core region close to active site and found crucial in establishing significant non-covalent interactions and buttressed the configuration of conserved *NSUN* fold inside the catalytic core. The substitution to Ser318 abolished these crucial interactions, especially the amide-ring ( $\pi$ - $\pi$ ) interaction between the backbone amide of Val331 and benzene ring of Tyr318 (Figures 3D, E), which instigated and amplified the flexibility toward the active site of *NSUN2* as predicted from vibrational entropy energy. To prove the hypothesis of increased flexibility in *NSUN2* upon Tyr318Ser substitution at an atomistic level, 100 ns MD simulations were executed on *wt* and *mt-NSUN2* models. The structural stability of the Rossmann catalytic core and PUA domain was investigated during simulations by computing mainly the root mean square fluctuations (RMSF) concerning the wild-type structures. RMSF graph highlighted the flexible regions of *wt* and *mt-NSUN2*.



**FIGURE 2** (A) Represents cytogenic location (p15.31) of *NSUN2* on chromosome 5, (B) typical structure of *NSUN2* gene, consisting of 19 exons. It also indicates the location of a missense variant (c.953A>C; p.Tyr318Ser) in exon-9 and the location of a splice acceptor site variant (c.97-1G>C) in the intronic region. (C) Sequence chromatograms of a heterozygous father (III-1) and homozygous affected daughter (IV-5) in family A. (D) Partial sequence chromatograms of the heterozygous mother (III-2) and homozygous affected son (IV-3) in family B.

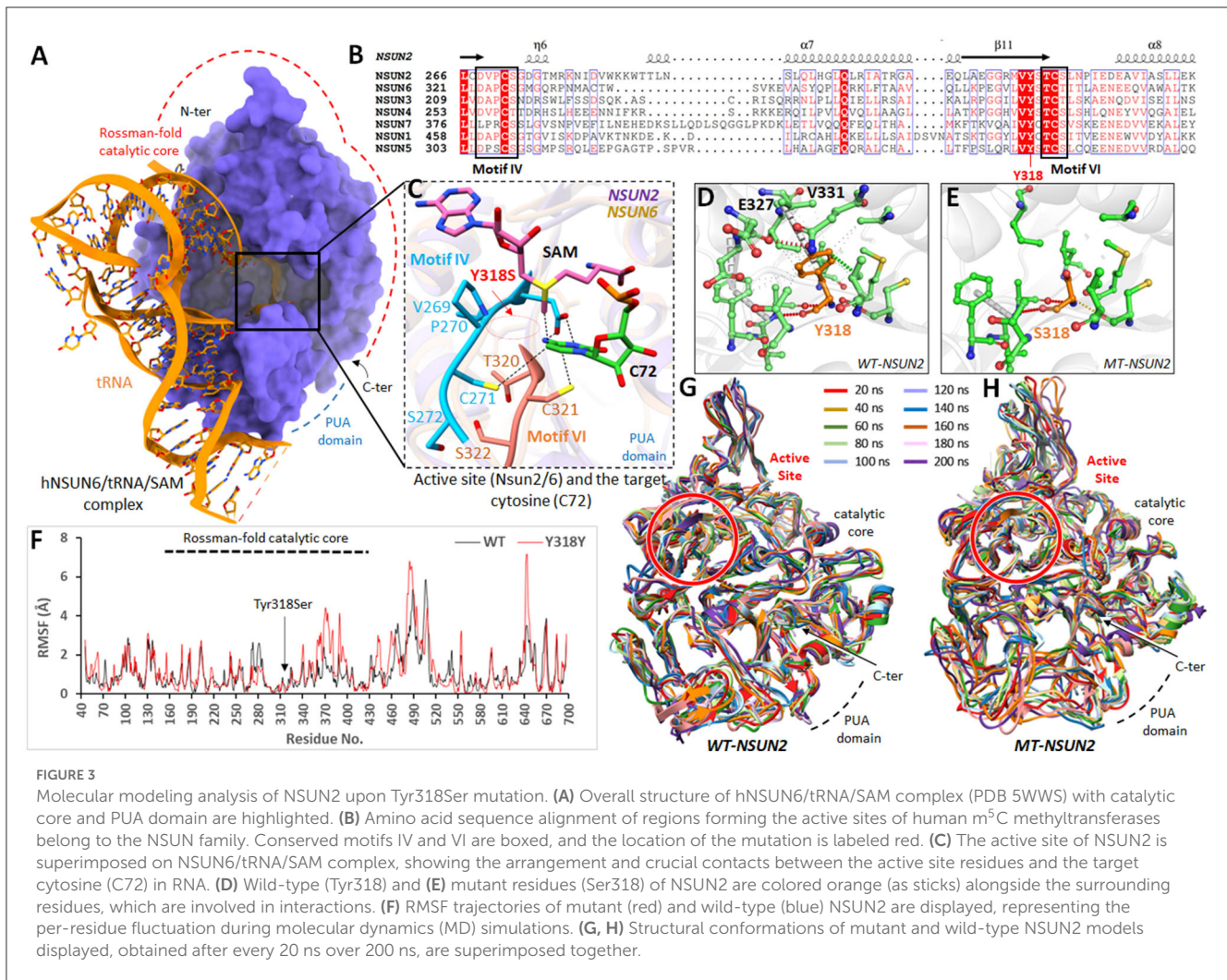
Over a simulation period of 100 ns, residual fluctuations of the catalytic core retained within 3Å in *wt-NSUN2*, and the structure remained converged (Figure 3F). In contrast, *mt-NSUN2* revealed evident Cα-RMSF differences in the catalytic core for residues 338–400 and amplified the fluctuations for residues 460–510. In addition, replacing Tyr318 with Ser318 elicited substantial mobility over simulation period, which retained its impact on the adjacent regions, including the compact active site core indicated to accommodate the methylation factor and facilitate the binding with C72. Overall, the fluctuations revealed the extent of the conformational shift of various regions, including the active site in the *mt-NSUN2* compared with its *wt*, which remained stable and converged throughout the simulation (Figures 3G, H).

### 4. Discussion

Subcellular, *NSUN2* is localized in the nucleolus of cells (9). Physiologically, *NSUN2* is involved in the post-transcriptional modification of tRNA and is a well-defined RNA methyltransferase responsible for cytosine-5 methylation (m<sup>5</sup>C) of tRNA-Leu (CAA) at position C34 (58). CCA is a conserved sequence at the 3' end of mature tRNA-Leu molecules to function as the site of leucine attachment. This sequence is added and maintained by tRNA nucleotidyltransferase (59). *NSUN2* was first reported in non-syndromic autosomal recessive ID by Khan et al. (9)

and Abbasi-Moheb et al. (22) in consanguineous Iranian and Pakistani families in 2012 (9, 22). The first SUN-domain-containing protein discovered in vertebrates was *NSUN2*, which is highly conserved from bacteria to humans. Frye and Watt identified it as a modulator of Myc-induced proliferation as it is the human homolog of the yeast *Trmt4* protein. They also discovered that it exhibits methyltransferase activity against rRNA, tRNA, and hemimethylated DNA *in vitro*. The structural integrity, translational efficiency, and fidelity of tRNAs are all affected by methylation and other modifications (60). The nucleotide in the wobble position of tRNA-leu is modified by *NSUN2* (CAA). The functional loss of *NSUN2* could result in the absence of tRNA-leu (CAA) and lead to changes in tissue-specific protein production, such as changes in codon usage. Codon usage is known to significantly impact translation rate in mammals, particularly during cell differentiation (61). *NSUN2* may have a significant role in translational control during synaptic plasticity; thus, its functional loss may affect learning and memory.

*Nsun2*<sup>-/-</sup>-knockout mice showed a significant reduction in body weight and body size in two *Nsun2*-knockout lines, as reported by Blanco et al. (62). These findings match the short stature observed in several affected individuals, including patients from the present study and previously reported families by Khan et al. (9). In addition, while working on a mouse model, Khan et al. said that *NSUN2* protein was localized to the nucleolus of Purkinje cells in the cerebellum of the



**FIGURE 3** Molecular modeling analysis of NSUN2 upon Tyr318Ser mutation. **(A)** Overall structure of hNSUN6/tRNA/SAM complex (PDB 5WWS) with catalytic core and PUA domain are highlighted. **(B)** Amino acid sequence alignment of regions forming the active sites of human m<sup>5</sup>C methyltransferases belong to the NSUN family. Conserved motifs IV and VI are boxed, and the location of the mutation is labeled red. **(C)** The active site of NSUN2 is superimposed on NSUN6/tRNA/SAM complex, showing the arrangement and crucial contacts between the active site residues and the target cytosine (C72) in RNA. **(D)** Wild-type (Tyr318) and **(E)** mutant residues (Ser318) of NSUN2 are colored orange (as sticks) alongside the surrounding residues, which are involved in interactions. **(F)** RMSF trajectories of mutant (red) and wild-type (blue) NSUN2 are displayed, representing the per-residue fluctuation during molecular dynamics (MD) simulations. **(G, H)** Structural conformations of mutant and wild-type NSUN2 models displayed, obtained after every 20 ns over 200 ns, are superimposed together.

mouse. The effects of the pathogenic variant (p.Gly679Arg) were confirmed by the transfection of wild-type and mutant constructs into cells and subsequent immunohistochemistry. In addition, p.Gly679Arg caused the failure of NSUN2 localization within the nucleolus (9).

NSUN2 genetic variations are reported in several consanguineous families, mainly from Pakistan or Iran. Primarily, Najamabadi and his colleagues in 2007 found the MRT5 locus at chromosome 5p15.31 associated with the ID in two Iranian families. Later, Abbasi-Moheb and co-workers in 2012 identified two nonsense variants (c.679C>T, p. Gln227Ter; c.1114C>T, p.Gln372Ter) in NSUN2 located within MRT5 region 16, 20, 21. Another group of researchers working on a Kurdish family also identified a homozygous transversion (g.6622224A>C) located 11-bp upstream from exon-6. This genomic variant predictably causes premature termination (Ile179ArgfsTer192) in exon-7 due to the skipping of exon-6 (22). Similarly, a Lebanese family reported a splice site mutation adjacent to exon-6. This splice site change (G>C transversion) occurred at 1-bp upstream of exon-6 of NSUN2 (13). Finally, in the same year, a missense variant c.2035G>A (p.Gly679Arg) in exon-19 of NSUN2 was reported in a Pakistani family, which resulted in a mutant protein that failed to localize appropriately in the nucleolus (9).

Apart from non-syndromic ID, NSUN2 variants are also involved in Noonan-like and Dubowitz-like syndromes. Nonetheless, NSUN2 variants in non-syndromic patients also exhibit some commonly occurring abnormal features, such as moderate to severe ID, short stature, facial dysmorphism with a long face, distinctive eyebrows, high nasal bridge or broad nasal base, drooping eyes, small pointed chin, and speech cluttering (62). The detailed description of all reported pathogenic variations and their associated clinical features are summarized in [Supplementary Table 2](#).

Our current genetic study enrolled two consanguineous Pakistani families with Pashtun ethnicity segregating non-syndromic ID. Genetic analysis in family A revealed a novel missense variant (c.953A>C; p.Tyr318Ser) in exon-9 of the NSUN2, while in family B, a novel splice site variant (NM\_017755.5, c.97-1G>C) in the same gene was identified. The identified novel splice variant (p.His86Profs\*16) introduced a premature termination codon, expected to cause a non-sense mediated decay (NMD), leading to the loss of function of NSUN2 protein. However, Hentze and Kulozik (63) reported that the imperfect messages initiated by truncating mutations are eliminated by NMD (63).

In this study, the impact of the Tyr318Ser novel missense variant on the stability of NSUN2 was accessed through a series of



computational methods to improve the predictions. The possible effects due to Tyr318Ser substitution include (i) the considerable decrease in the side chain volume, which could eliminate the electrostatic interactions with the neighboring side chains (64), (ii) shifts in the side-chain packing, and (iii) changes in solvent accessibility, which has a crucial role in secondary structure stabilization (65, 66), thus destabilizing the overall active site conformation. Furthermore, MD simulation revealed substantial fluctuations in NSUN2 in several regions upon Tyr318Ser substitution, especially in the  $\beta$ -sheet catalytic core region.

The fact that flexibility causes large fluctuations in *mt-NSUN2* suggests that a slight conformational change upon mutation stimulates a considerable structural change in the peripheral region of the protein (67). This inherent occurrence of flexibility might significantly destabilize the overall active site configuration in *mt-NSUN2*, which is necessary to recognize methylation target C72 and conformation of a cofactor for methylation activity. For example, Asp293 (268 in *NSUN2*) recognizes the N6 of the adenine ring, and its displacement could lead to the loss of SAM binding capability and a loss of methylation activity to tRNA<sup>Cys</sup>. Likewise, removing Ser371 (319 in *NSUN2*; one residue ahead to the mutation site), typically involved in recognizing C72, might disrupt this interaction. Additionally, these dynamic consequences were evident from increased vibrational entropy energy between *wt* and *mt-NSUN2* ( $\Delta\Delta S_{\text{vib}}$ : 1.127 kcal.mol<sup>-1</sup>.K<sup>-1</sup>), which worsened the flexibility of catalytic core [calculated by ENCoM (53)], as observed in the contiguous residues toward both directions from the site of substitution (Supplementary Figure S2). This vibrational entropy considerably influences the binding free energies of proteins (53, 68). Therefore, we speculated that the Tyr318Ser substitution of NSUN might be a loss-of-function mutation that eventually leads to the loss of SAM binding capability and methylation activity to tRNA<sup>Cys</sup>. Overall, the *NSUN2* structural analyses uncovered that this pathogenic variant alters the overall protein function and implicates the abrupt catalytic mechanism of the RNA:m5C methyltransferase, thus leading to non-syndromic ID.

## 5. Conclusion

The present molecular genetic study found two novel variants in *NSUN2* and, thus, further extended the mutational spectrum and allelic heterogeneity of *NSUN2* and its association with non-syndromic ID. Furthermore, the comprehensive genotype-linked phenotype analysis has found that the *NSUN2* variant usually leads to facial dysmorphism, moderate ID, short stature, and speech delay. The long-term ID conditions of the patients create numerous social barriers. Such genetic investigations not only increase general awareness about the disease but also help to improve personalized patient support and treatment.

## Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: National Center for

Biotechnology Information (NCBI) ClinVar, <https://www.ncbi.nlm.nih.gov/clinvar/>, SCV002599111 and SCV002599110.1.

## Ethics statement

The studies involving human participants were reviewed and approved by Institutional Ethical Review Board of Kohat University of Science and Technology, Kohat, Pakistan. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin. Written informed consent was obtained from the individual(s), and minor(s)' legal guardian/next of kin, for the publication of any potentially identifiable images or data included in this article.

## Author contributions

NaM and SH have performed experiments and data analysis. SK, SJ, and NKh recruited the families. ZR (3rd author) performed clinical analysis, while MMu, SA, and ZR (10th author) were involved in Sanger sequencing. MMi performed molecular modeling studies. NKa, MK, NoM, NW, and SK conceptualized and supervised the study and remained involved in data analysis, manuscript drafting, and fund acquisition. All authors have read, edited, and approved the final version of the manuscript.

## Funding

The current research is partially supported by HEC, Pakistan, through an NRPU grant (5564/KPK/NRPU/R&D/HEC/2016).

## Acknowledgments

We are indebted to all research participants for their volunteer contribution to the present genetic study. We are thankful to Prof. Reiner Siebert, Institute of Human Genetics, Ulm University, for his moral support. We are thankful to 3billion, Inc. (Seoul, South Korea), for exome analysis of family B.

## 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.

## References

- Khan MA, Khan S, Windpassinger C, Badar M, Nawaz Z, Mohammad RM. The molecular genetics of autosomal recessive nonsyndromic intellectual disability: a mutational continuum and future recommendations. *Ann Hum Genet.* (2016) 80:342–68. doi: 10.1111/ahg.12176
- American Psychiatric Association D, Association AP. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*. Washington, DC: American Psychiatric Association. (2013). doi: 10.1176/appi.books.9780890425596
- Moudi M, Mehrjardi MYV, Hozhabri H, Metanat Z, Kalantar SM, Taheri M, et al. Novel variants underlying autosomal recessive neurodevelopmental disorders with intellectual disability in Iranian consanguineous families. *J Clin Lab Anal.* (2022) 36:e24241. doi: 10.1002/jcla.24241
- Jamra R. Genetics of autosomal recessive intellectual disability. *Medizinische genetic.* (2018) 30:323–27.
- Srouf M, Shimokawa N, Hamdan FF, Nassif C, Poulin C, Gazali LA, et al. Dysfunction of the Cerebral Glucose Transporter SLC45A1 in Individuals with Intellectual Disability and Epilepsy. *Am J Hum Genet.* (2017) 100:824–30. doi: 10.1016/j.ajhg.2017.03.009
- Li M, Tao Z, Zhao Y, Li L, Zheng J, Li Z, et al. 5-methylcytosine RNA methyltransferases and their potential roles in cancer. *J Transl Med.* (2022) 20:1–16. doi: 10.1186/s12967-022-03427-2
- Shinoda S, Kitagawa S, Nakagawa S, Wei F-Y, Tomizawa K, Araki K, et al. Mammalian NSUN2 introduces 5-methylcytidines into mitochondrial tRNAs. *Nucleic Acids Res.* (2019) 47:8734–45. doi: 10.1093/nar/gkz575
- Yang X, Yang Y, Sun BF, Chen YS, Xu JW, Lai WY, et al. 5-methylcytosine promotes mRNA export—NSUN2 as the methyltransferase and ALYREF as an m5C reader. *Cell Res.* (2017) 27:606–25. doi: 10.1038/cr.2017.55
- Khan MA, Rafiq MA, Noor A, Hussain S, Flores JV, Rupp V, et al. Mutation in NSUN2, which encodes an RNA methyltransferase, causes autosomal-recessive intellectual disability. *Am J Hum Genet.* (2012) 90:856–63. doi: 10.1016/j.ajhg.2012.03.023
- Doan RN, Lim ET, De Rubeis S, Betancur C, Cutler DJ, Chiochetti AG, et al. Recessive gene disruptions in autism spectrum disorder. *Nat Genet.* (2019) 51:1092–8. doi: 10.1038/s41588-019-0433-8
- Blaze J, Navickas A, Phillips HL, Heissel S, Plaza-Jennings A, Miglani S, et al. Neuronal Nsun2 deficiency produces tRNA epitranscriptomic alterations and proteomic shifts impacting synaptic signaling and behavior. *Nat Commun.* (2021) 12:1–16. doi: 10.1038/s41467-021-24969-x
- Innes AM, McInnes BL, Dymont DA. Clinical and genetic heterogeneity in Dubowitz syndrome: Implications for diagnosis, management and further research. In: *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*. Hoboken, USA: John Wiley & Sons, Inc. (2018). p. 387–397. doi: 10.1002/ajmg.c.31661
- Martinez FJ, Lee JH, Lee JE, Blanco S, Nickerson E, Gabriel S, et al. Whole exome sequencing identifies a splicing mutation in NSUN2 as a cause of a Dubowitz-like syndrome. *J Med Genet.* (2012) 49:380–5. doi: 10.1136/jmedgenet-2011-100686
- Kato K, Mizuno S, Morton J, Toyama M, Hara Y, Wasmer E, et al. Expanding the phenotype of biallelic loss-of-function variants in the NSUN2 gene: Description of four individuals with juvenile cataract, chronic nephritis, or brain anomaly as novel complications. *Am J Med Genet Part A.* (2021) 185:282–5. doi: 10.1002/ajmg.a.61927
- Komara M, Al-Shamsi AM, Ben-Salem S, Ali BR, Al-Gazali L. A novel single-nucleotide deletion (c. 1020delA) in NSUN2 causes intellectual disability in an Emirati child. *J Molec Neurosci.* (2015) 57:393–9. doi: 10.1007/s12031-015-0592-8
- Sun Y-Z, Chen X-B, Wang R-R, Li W-Y, Ma Y. Exploring the effect of N308D mutation on protein tyrosine phosphatase-2 cause gain-of-function activity by a molecular dynamics study. *J Cell Biochem.* (2019) 120:5949–61. doi: 10.1002/jcb.27883
- Yang J-C, Risch E, Zhang M, Huang C, Huang H, Lu L. Association of tRNA methyltransferase NSUN2/IGF-II molecular signature with ovarian cancer survival. *Future Oncol.* (2017) 13:1981–90. doi: 10.2217/fo-2017-0084
- Frye M, Dragoni I, Chin S-F, Spiteri I, Kurowski A, Provenzano E, et al. Genomic gain of 5p15 leads to over-expression of Misu (NSUN2) in breast cancer. *Cancer Lett.* (2010) 289:71–80. doi: 10.1016/j.canlet.2009.08.004
- Yan J, Liu J, Huang Z, Huang W, Lv J. FOXC2-AS1 stabilizes FOXC2 mRNA via association with NSUN2 in gastric cancer cells. *Hum Cell.* (2021) 34:1755–64. doi: 10.1007/s13577-021-00583-3
- Kuss AW, Garshasbi M, Kahrizi K, Tzschach A, Behjati F, Darvish H, et al. Autosomal recessive mental retardation: homozygosity mapping identifies 27 single linkage intervals, at least 14 novel loci and several mutation hotspots. *Hum Genet.* (2011) 129:141–8. doi: 10.1007/s00439-010-0907-3
- Najmabadi H, Motazacker MM, Garshasbi M, Kahrizi K, Tzschach A, Chen W, et al. Homozygosity mapping in consanguineous families reveals extreme heterogeneity of non-syndromic autosomal recessive mental retardation and identifies 8 novel gene loci. *Hum Genet.* (2007) 121:43–8. doi: 10.1007/s00439-006-0292-0
- Abbasi-Moheb L, Mertel S, Gonsior M, Nouri-Vahid L, Kahrizi K, Cirak S, et al. Mutations in NSUN2 cause autosomal-recessive intellectual disability. *Am J Hum Genet.* (2012) 90:847–55. doi: 10.1016/j.ajhg.2012.03.021
- Vinkškel M, Writzl K, Maver A, Peterlin B. Improving diagnostics of rare genetic diseases with NGS approaches. *J Commun. Genet.* (2021) 12:247–56. doi: 10.1007/s12687-020-00500-5
- Parveen A, Mirza MU, Vanmeert M, Akhtar J, Bashir H, Khan S, et al. A novel pathogenic missense variant in CNM4 underlying Jalili syndrome: Insights from molecular dynamics simulations. *Molec Genet Genom Med.* (2019) 7:e902. doi: 10.1002/mgg3.902
- Saeed S, Bonnefond A, Tamanini F, Mirza MU, Manzoor J, Janjua QM, et al. Loss-of-function mutations in ADCY3 cause monogenic severe obesity. *Nat Genet.* (2018) 50:175. doi: 10.1038/s41588-017-0023-6
- Pirolli D, Sciandra F, Bozzi M, Giardina B, Brancaccio A, De Rosa MC. Insights from molecular dynamics simulations: structural basis for the V567D mutation-induced instability of zebrafish alpha-dystroglycan and comparison with the murine model. *PLoS ONE.* (2014) 9:e103866. doi: 10.1371/journal.pone.0103866
- Durrani FG, Gul R, Mirza MU, Kaderbhai NN, Froeyen M, Saleem M. Mutagenesis of DsbAss is Crucial for the Signal Recognition Particle Mechanism in Escherichia coli: Insights from Molecular Dynamics Simulations. *Biomolecules.* (2019) 9:133. doi: 10.3390/biom9040133
- Daggett V, Levitt M. Realistic simulations of native-protein dynamics in solution and beyond. *Annu Rev Biophys Biomol Struct.* (1993) 22:353–80. doi: 10.1146/annurev.bb.22.060193.002033
- Shaw DE, Maragakis P, Lindorff-Larsen K, Piana S, Dror RO, Eastwood MP, et al. Atomic-level characterization of the structural dynamics of proteins. *Science.* (2010) 330:341–6. doi: 10.1126/science.1187409
- Li H, Durbin R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics.* (2010) 26:589–95. doi: 10.1093/bioinformatics/btp698
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics.* (2009) 25:2078–9. doi: 10.1093/bioinformatics/btp352
- Toolkit P. *GitHub Repository*. Broad Institute. (2019).
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* (2010) 20:1297–303. doi: 10.1101/gr.107524.110
- Rimmer A, Phan H, Mathieson I, Iqbal Z, Twigg SRF, Consortium W, et al. Integrating mapping-, assembly-and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat Genet.* (2014) 46:912–8. doi: 10.1038/ng.3036
- Friedman S, Gauthier L, Farjoun Y, Banks E. Lean and deep models for more accurate filtering of SNP and INDEL variant calls. *Bioinformatics.* (2020) 36:2060–7. doi: 10.1093/bioinformatics/bt2901
- Klambauer G, Schwarzbauer K, Mayr A, Clevert DA, Mitterecker A, Bodenhofer U, et al. cn. MOPS: mixture of Poissons for discovering copy number variations in next-generation sequencing data with a low false discovery rate. *Nucl Acids Res.* (2012) 40:e69–e69. doi: 10.1093/nar/gks003
- Plagnol V, Curtis J, Epstein M, Mok KY, Stebbings E, Grigoriadou S, et al. A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. *Bioinformatics.* (2012) 28:2747–54. doi: 10.1093/bioinformatics/bts526

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fneur.2023.1168307/full#supplementary-material>

38. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* (2011) 43:491–8. doi: 10.1038/ng.806
39. Krumm N, Sudmant PH, Ko A, O’Roak BJ, Malig M, Coe BP, et al. Copy number variation detection and genotyping from exome sequence data. *Genome Res.* (2012) 22:1525–32. doi: 10.1101/gr.138115.112
40. Quinodoz M, Peter VG, Bedoni N, Bertrand BR, Cisarova K, Salmaninejad A, et al. AutoMap is a high performance homozygosity mapping tool using next-generation sequencing data. *Nat Commun.* (2021) 12:518. doi: 10.1038/s41467-020-20584-4
41. Kopanos C, Tsiolkas V, Kouris A, Chapple CE, Aguilera MA, Meyer R, et al. VarSome: the human genomic variant search engine. *Bioinformatics.* (2019) 35:1978–80. doi: 10.1093/bioinformatics/bty897
42. Kleinberger J, Maloney KA, Pollin TI, Jeng LJB. An openly available online tool for implementing the ACMG/AMP standards and guidelines for the interpretation of sequence variants. *Genet Med.* (2016) 18:1165. doi: 10.1038/gim.2016.13
43. Greene D, BioResource N, Richardson S, Turro E. Phenotype similarity regression for identifying the genetic determinants of rare diseases. *Am J Hum Genet.* (2016) 98:490–9. doi: 10.1016/j.ajhg.2016.01.008
44. Köhler S, Schulz MH, Krawitz P, Bauer S, Dölken S, Ott CE, et al. Clinical diagnostics in human genetics with semantic similarity searches in ontologies. *Am J Hum Genet.* (2009) 85:457–64. doi: 10.1016/j.ajhg.2009.09.003
45. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* (2003) 31:3812–4. doi: 10.1093/nar/gkg509
46. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Current Prot Human Genet.* (2013) 76:1–41. doi: 10.1002/0471142905.hg0720s76
47. Capriotti E, Fariselli P, Rossi I, Casadio R. A three-state prediction of single point mutations on protein stability changes. *BMC Bioinform.* (2008) 9:S6. doi: 10.1186/1471-2105-9-S2-S6
48. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. *Electrophoresis.* (1997) 18:2714–23. doi: 10.1002/elps.1150181505
49. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallograph. Section D.* (2010) 66:12–21. doi: 10.1107/S0907444909042073
50. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* (2007) 35:W407–10. doi: 10.1093/nar/gkm290
51. Pires DEV, Ascher DB, Blundell TL. DUET: a server for predicting effects of mutations on protein stability using an integrated computational approach. *Nucleic Acids Res.* (2014) 42:W314–9. doi: 10.1093/nar/gku411
52. Yan M, Zhang Z, Brady JR, Schillbach S, Fairbrother WJ, Dixit VM. Identification of a novel death domain-containing adaptor molecule for ectodysplasin-A receptor that is mutated in crinkled mice. *Curr Biol.* (2002) 12:409–13. doi: 10.1016/S0960-9822(02)00687-5
53. Frappier V, Chartier M, Najmanovich RJ. ENCoM server: exploring protein conformational space and the effect of mutations on protein function and stability. *Nucleic Acids Res.* (2015) 43:W395–400. doi: 10.1093/nar/gkv343
54. Case DA, Babin V, Berryman JT, Betz RM, Cai Q, Cerutti DS, et al. (2014). *The Amber Molecular Dynamics Package*. p. 14. Available online at: <https://ambermd.org/doc12/Amber14.pdf>
55. Parveen A, Khan SA, Mirza MU, Bashir H, Arshad F, Iqbal M, et al. Deleterious variants in WNT10A, EDAR, and EDA causing isolated and syndromic tooth agenesis: A structural perspective from molecular dynamics simulations. *Int J Mol Sci.* (2019) 20:5282. doi: 10.3390/ijms20215282
56. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem.* (2004) 25:1605–12. doi: 10.1002/jcc.20084
57. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDScript server. *Nucleic Acids Res.* (2014) 42:W320–4. doi: 10.1093/nar/gku316
58. Brzezicha B, Schmidt M, Makalowska I, Jarmolowski A, Pienkowska J, Szwejkowska-Kulinska Z. Identification of human tRNA:m5C methyltransferase catalysing intron-dependent m5C formation in the first position of the anticodon of the pre-tRNA Leu (CAA). *Nucleic Acids Res.* (2006) 34:6034–43. doi: 10.1093/nar/gkl765
59. Hou YM. CCA addition to tRNA: implications for tRNA quality control. *IUBMB Life.* (2010) 62:251–60. doi: 10.1002/iub.301
60. Frye M, Watt FM. The RNA methyltransferase Misu (NSun2) mediates Myc-induced proliferation and is upregulated in tumors. *Curr Biol.* (2006) 16:971–81. doi: 10.1016/j.cub.2006.04.027
61. Plotkin JB, Robins H, Levine AJ. Tissue-specific codon usage and the expression of human genes. *Proc Natl Acad Sci U S A.* (2004) 101:12588–91. doi: 10.1073/pnas.0404957101
62. Blanco S, Kurowski A, Nichols J, Watt FM, Benitah SA, Frye M. The RNA-methyltransferase Misu (NSun2) poises epidermal stem cells to differentiate. *PLoS Genet.* (2011) 7:e1002403. doi: 10.1371/journal.pgen.1002403
63. Hentze MW, Kulozik AE. A perfect message: RNA surveillance and nonsense-mediated decay. *Cell.* (1999) 96:307–10. doi: 10.1016/S0092-8674(00)80542-5
64. Ghosh T, Garde S, García AE. Role of backbone hydration and salt-bridge formation in stability of  $\alpha$ -helix in solution. *Biophys J.* (2003) 85:3187–93. doi: 10.1016/S0006-3495(03)74736-5
65. Vila JA, Ripoll DR, Scheraga HA. Physical reasons for the unusual  $\alpha$ -helix stabilization afforded by charged or neutral polar residues in alanine-rich peptides. *Proc Nat Acad Sci.* (2000) 97:13075–9. doi: 10.1073/pnas.240455797
66. Garcia AE, Sanbonmatsu KY.  $\alpha$ -Helical stabilization by side chain shielding of backbone hydrogen bonds. *Proc Nat Acad Sci.* (2002) 99:2782–7. doi: 10.1073/pnas.042496899
67. Gohlke H, Kuhn LA, Case DA. Change in protein flexibility upon complex formation: analysis of Ras-Raf using molecular dynamics and a molecular framework approach. *PROTEINS.* (2004) 56:322–37. doi: 10.1002/prot.20116
68. Goethe M, Fita I, Rubi JM. Vibrational entropy of a protein: large differences between distinct conformations. *J Chem Theory Comput.* (2014) 11:351–9. doi: 10.1021/ct500696p