UNIVERSITÀ DEL PIEMONTE ORIENTALE

School of Medicine

Master of Science in Medical Biotechnologies

Department of Health Science

Thesis Title

Assessment of *RNU4-2* Spliceosomal snRNA Gene Variants in an Italian Cohort of Patients with Neurodevelopmental Disorders

MENTOR

Prof.ssa Mara Giordano

CANDIDATE

Bulbiankova Darya

Matricula number: 20039630

Academic year 2025-2026

Table of Contents

SUMMARY 4

1. INTRODUCTION 5

1.1 Neurodevelopmental disorders: definition, classification and genetic background 5

1.2 The spliceosome: structure, function and biological significance 6

1.3 Spliceosomal snRNAs and the emergence of RNUopathies 7

1.4 *RNU4-2*: characteristics, function and association with neurodevelopmental disorders 8

1.4.1 Genomic characteristics and U4 snRNA biogenesis 8

1.4.2 Functional role of U4 snRNA in spliceosome activation 9

1.4.3 Pathogenic *RNU4-2* variants and dominant ReNU syndrome 11

1.4.4 Expansion of the *RNU4-2* disease spectrum: recessive *RNU4-2*-associated neurodevelopmental disorder 15

2. AIM OF THE STUDY 17

3. MATERIALS AND METHODS 18

3.1 Patient selection and clinical data review 18

3.2 Genomic DNA extraction from peripheral blood 18

3.3 DNA quantification and quality assessment 19

3.4 PCR amplification of RNU4-2 20

3.5 Agarose gel electrophoresis 21

3.6 PCR product purification 21

3.7 Sanger sequencing 22

3.8 Variant analysis, classification, and genotype-phenotype correlation 23

4. RESULTS 24

4.1 Patients cohort 24

4.2 *RNU4-2* variants identification 24

4.3 Segregation analysis 25

4.4 Structural localization of the identified variants 26

5. DISCUSSION 28

5.1 Structural distribution of the identified *RNU4-2* variants 28

5.2 Domain-based clinical interpretation 29

5.6 Limitations and future perspectives 31

6. CONCLUSION 33

BIBLIOGRAPHY 34

SUMMARY

Rationale of the study

Neurodevelopmental disorders are clinically and genetically heterogeneous conditions that may include developmental delay, intellectual disability, epilepsy, hypotonia, language impairment and behavioural abnormalities. Despite advances in genetic testing, many patients remain without a molecular diagnosis, partly because standard approaches mainly focus on protein-coding genes. Recent studies have identified pathogenic variants in non-coding spliceosomal small nuclear RNA genes as causes of NDDs. In particular, *RNU4-2*, encoding U4 snRNA of the major spliceosome, has been linked to dominant ReNU syndrome, mainly caused by monoallelic pathogenic variants in a central hotspot, and recently found to be associated to a recessive disorder.

Planning of the study

This thesis aimed to investigate *RNU4-2* variants in patients with neurodevelopmental disorders who remained genetically undiagnosed after previous testing. Genomic DNA was extracted from peripheral blood, quantified, amplified by PCR, checked by agarose gel electrophoresis, purified and analysed by Sanger sequencing. The identified variants were evaluated according to population frequency, structural localization within U4 snRNA and comparison with both dominant and recessive *RNU4-2*-associated disorders.

Results

A total of 300 patients with neurodevelopmental disorders were analysed. Eight heterozygous *RNU4-2* variants of uncertain significance were identified in seven patients. None of the variants mapped to the classical dominant ReNU syndrome hotspot. Some variants were located near regions implicated in the recessive *RNU4-2*-associated disorder; however, they were found at the heterozygous state, not fitting the expected biallelic inheritance model. The clinical presentation of the patients showed some features related to the *RNU4-2* phenotypes, albeit the genotype–phenotype correlation in this disorder is not yet well-defined.

Conclusions

This study expands the screening of the non-coding spliceosomal gene *RNU4-2* in a cohort of patients with neurodevelopmental disorders who remained without a molecular diagnosis after previous genetic testing. Since the identified variants do not overlap with the classical dominant ReNU syndrome hotspot and do not clearly fit the recessive biallelic model, their clinical relevance remains to be clarified. However, future segregation studies, analysis of larger cohorts and, particularly, functional studies will be essential to verify the biological impact of the identified VUS and to assess whether they may contribute to disease susceptibility or phenotype modulation.

1. INTRODUCTION

1.1 Neurodevelopmental disorders: definition, classification and genetic background

Patients with neurodevelopmental disorders (NDDs) form a clinically heterogeneous group characterized by impairments in motor and behavioral skills, as well as cognitive development[1]. Disease onset is typically observed in early childhood. There are marked variations in the severity and clinical presentation of these medical conditions, which may manifest through developmental delay, intellectual disability, autism spectrum disorders, epilepsy, and a wide range of syndromic presentations with congenital anomalies or growth abnormalities. NDDs interfere with key aspects of development and frequently impede normal functioning throughout the lifespan. These conditions profoundly affect patients, their families, and impose a substantial burden on healthcare systems.

The genetic architecture of NDDs is complex[2]. Disease risk may result from chromosomal abnormalities, copy-number variations, single-nucleotide variants, small insertions or deletions, repeat expansions, *de novo* variants, and inherited rare variants among others. Many established NDD genes encode proteins involved in chromatin regulation, transcription, RNA processing, synaptic function, neuronal migration, and cytoskeletal organization. Even though genomic diagnostics have advanced considerably and become more widely accessible in diagnostic practice, a large number of affected individuals still lack a conclusive molecular diagnosis even after undergoing conventional testing.

One of the causes of these unresolved cases is the fact that clinical interpretation has traditionally focused mostly on protein-coding regions, while pathogenic variants in the non-coding genome have been more challenging to identify and/or interpret. However, the increasing employment of large-scale genomic sequencing has revealed the substantial contribution of non-coding variants to the pathogenesis of rare developmental disorders. This finding has underscored a need to take functionally significant non-coding elements into account, in addition to coding genes, in diagnostic workflows for NDDs[3].

1.2 The spliceosome: structure, function and biological significance

Pre-mRNA splicing is a critical post-transcriptional process that produces mature transcripts by excising introns from pre-mRNA and then ligating the remaining exons. In humans this reaction is catalysed by the spliceosome, which is a large ribonucleoprotein complex composed of small nuclear RNAs (snRNAs) and numerous associated proteins. Accurate splicing is necessary to safeguard the coding potential of genes and to regulate the generation of alternative transcript isoforms. Disruption of splicing can have significant developmental effects in the nervous system, where fine-tuned temporal and spatial control of gene expression is necessary for neuronal differentiation, migration, synaptogenesis, and circuit maturation[4].

Two related spliceosomal systems operate in eukaryotic cells. The major spliceosome processes more than 99% of introns and is composed of U1, U2, U4, U5 and U6 snRNAs. The minor spliceosome processes an extremely small subset of introns and contains U11, U12, U4atac, U6atac and the shared U5 snRNA[5].

Major spliceosome assembly occurs through ordered interactions between snRNPs and the pre-mRNA. U1 snRNP recognizes the 5′ splice site, marking the beginning of the intron and helping commit the transcript to the splicing pathway, thus constituting the commitment, or early, complex (E complex). Next, U2 snRNP binds the branch-point sequence (BPS), forming the A complex. During this interaction, the adenosine of the BPS is left unpaired and is "bulging out". This adesonine will later act as the nucleophile in the first catalytic reaction. After that U4/U6.U5 tri-snRNP is recruited, generating the pre-catalytic, or B, complex (Figure 1). This means that the catalytic potential of the assembled spliceosome is restrained through U6 and U4 pairing, which avoids premature activation that would lead to an incorrect splicing event. Upon activation, U1 and U4 are released, U6 rearranges to form part of the catalytic center together with U2, and U5 aligns the exons for ligation[4].

These rearrangements are highly coordinated; therefore, even single-nucleotide variants in structured snRNA domains can affect any of the aforementioned processes[6]. This explains why non-coding snRNA genes, despite their small size, can have major consequences when mutated in functionally meaningful regions.

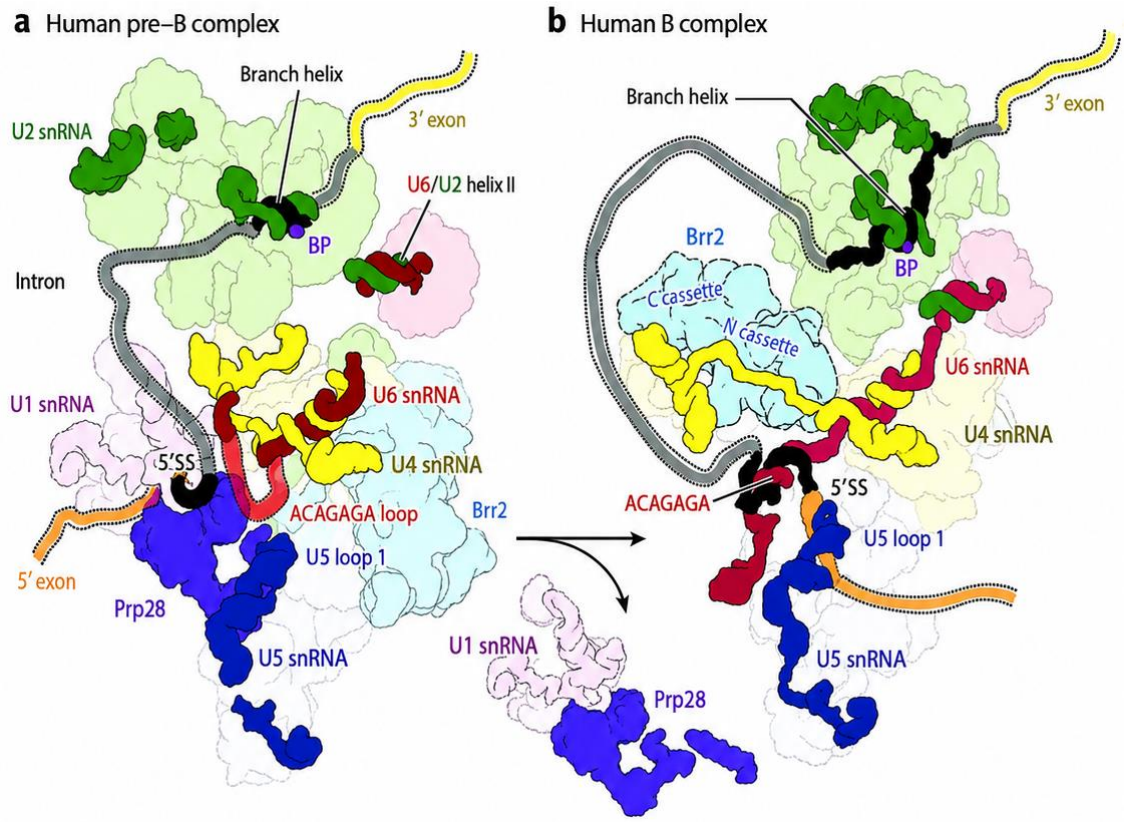


Figure 1. **Schematic representation of spliceosome assembly and function.** U1 recognizes the 5′ splice site, U2 binds the branch-point sequence, the U4/U6.U5 tri-snRNP is recruited to the pre-catalytic spliceosome, and subsequent rearrangements release U4, activate U6 and allow U5-mediated exon alignment. Adapted from Wilkinson, Charenton and Nagai (2020)[4].

1.3 Spliceosomal snRNAs and the emergence of RNUopathies

Spliceosomal snRNAs represent a particularly important class of non-coding RNAs because they form the RNA core of the spliceosome. They do not act as passive scaffolds; rather, they directly participate in splice-site recognition, spliceosome assembly, RNA-RNA rearrangements and catalysis. U1 recognizes the 5′ splice site, U2 recognizes the branch point, U4 regulates U6 and maintains the spliceosome in an inactive activation-competent state, U5 aligns exons for ligation, and U6 contributes to the catalytic center.

The involvement of several snRNA genes in Mendelian disorders indicates that the correct structure, expression and maturation of these RNAs is essential for normal human development[6].

As summarized in Table 1, genes encoding U2, U4, and U5 have previously been connected to pathogenic neurodevelopmental abnormalities, but no confirmed NDD genes have yet been established for U1 and U6[7,8].

Table 1. Functional comparison of major spliceosomal snRNAs and currently associated neurodevelopmental disease genes.

snRNA	Main role in splicing	Disease-associated gene
U1	Recognition of the 5′ splice site	No confirmed NDD-associated snRNA gene yet
U2	Recognition of the branch-point sequence	*RNU2-2*
U4	Regulation of U6 and spliceosome activation	*RNU4-2*
U5	Alignment of exons during ligation	*RNU5B-1*; *RNU5A-1* candidate
U6	Formation of the catalytic center with U2	No confirmed disease-causing snRNA gene yet

In 2025[9], recurrent germline variants in *RNU2-2*, encoding a U2 snRNA, were associated with a severe neurodevelopmental disorder characterized by intellectual disability, autism, microcephaly, hypotonia, epilepsy, and a propensity for seizures. Variants in *RNU5B-1*, encoding U5 snRNA, were independently implicated in NDDs by studies focusing on R-loop-forming regions and on major spliceosome snRNA genes[10,11]. These findings support the broader concept of "RNUopathies": neurodevelopmental disorders caused by pathogenic variants in non-coding spliceosomal RNA genes.

1.4 *RNU4-2*: characteristics, function and association with neurodevelopmental disorders

1.4.1 Genomic characteristics and U4 snRNA biogenesis

RNU4-2 is a compact non-protein-coding gene located on chromosome 12q24.23 and annotated on the reverse strand in the GRCh38/hg38 reference genome. It produces a 141-nucleotide U4 snRNA transcript, one of the U4 molecules incorporated into the major spliceosome. Its reverse-strand orientation means that transcript-level nucleotide numbering proceeds in the opposite direction to increasing genomic coordinates. It is relevant when comparing transcript-based HGVS nomenclature with genomic variant descriptions, as the

same molecular change may be represented differently depending on whether it is described at the RNA transcript or genomic DNA level. This information is summarized in Table 2[12].

Table 2. Basic annotation of *RNU4-2* relevant to variant interpretation.

Feature	Annotation
Official gene symbol	*RNU4-2*
Gene type	Small nuclear RNA gene / non-protein-coding RNA gene
RNA product	U4 small nuclear RNA / U4 snRNA
Chromosomal location	Chromosome 12q24.23
Reference genome	GRCh38 / hg38
Strand orientation	Reverse / minus strand
Transcript length	141 nucleotides
Disease-relevant region	Central 18-bp region, n.62-n.79
Associated disorder	*RNU4-2*-related neurodevelopmental disorder / ReNU syndrome

Like other RNA polymerase II-transcribed spliceosomal snRNAs, U4 snRNA undergoes a regulated maturation pathway. After transcription, precursor snRNAs are processed, exported to the cytoplasm, assembled with the Sm protein ring, modified, and re-imported into the nucleus. Further maturation occurs in nuclear Cajal bodies before U4 associates with U6 to form the U4/U6 di-snRNP, which then joins U5 to generate the U4/U6.U5 tri-snRNP[13]. This biogenesis pathway is important because pathogenicity may depend not only on the primary sequence of U4 but also on its folding, protein interactions and incorporation into spliceosomal particles.

1.4.2 Functional role of U4 snRNA in spliceosome activation

U4 snRNA plays a major inhibitory role by sequestering catalytically important regions of U6, thereby maintaining the spliceosome in an inactive but activation-competent state. When the tri-snRNP is drawn to the assembling spliceosome, significant structural alterations take place. The separation of U4 from U6, which permits U6 to take on the RNA conformations required

for catalytic activity, is one such alteration[4,14]. Thus, U4 is essential for both structurally maintaining the pre-catalytic spliceosome and ensuring correct spliceosome activation.

It is important to consider the secondary structure of the U4/U6 snRNA duplex in order to understand the pathogenic relevance of *RNU4-2* variants as U4 functions through direct base-pairing interactions with U6. The most notable features are presented in Table 3, all of which contribute to the conformational organization required for spliceosome activation[15]. In particular, most mutations cluster in the central region of U4, expressly around the Stems I and III and the intervening nucleotides[16]. This pattern suggests that disease-associated alterations predominantly affect structural elements critical for function, rather than being randomly distributed throughout the transcript. This is crucial for understanding how single-nucleotide changes may alter U4/U6 interactions, destabilize the inactive spliceosomal conformation, and ultimately impair splicing fidelity[14].

*Table 3.*Main U4 domains in the context of U4/U6 duplex.

U4 region	Approximate position	Function
Stem II	U4 5′ end, around n.3–16	Forms part of the U4/U6 base-pairing interaction; contributes to the stability of the U4/U6 duplex.
5′ stem-loop	Around n.20–52	Contributes to local U4 secondary structure; may help to preserve the correct folding of U4 snRNA; includes the k-turn region.
k-turn	Around n.27–n.35 and n.41–n.46	Functional element within the 5′ stem-loop; implicated in U4 folding and protein interaction, including SNU13/15.5K binding.
Stem I	Around n.56–62	U4/U6 paired region; helps stabilize the U4/U6 duplex and participates in the structural arrangement required before U6 is released during activation.
T-loop / quasi-pseudoknot	Around n.63–67	Critical structural region; helps position the U6 ACAGAGA box, which is essential for 5′ splice-site recognition.

RBM42 interaction region	Around n.68–70	Implicated in protein–RNA interaction, especially with RBM42.
Stem III	Around n.72–79	Critical paired structural region involved in stabilizing the U4/U6 conformation.
3′ stem-loop	Around n.85–117	Contributes to the overall secondary structure and stability of U4 snRNA.
Sm protein binding site	Around n.118–n.126	Required for Sm protein binding and snRNP maturation; implicated in the recessive *RNU4-2*-associated disorder.
Terminal stem-loop	Around n.127–n.141	Terminal structural element of U4 snRNA.

1.4.3 Pathogenic *RNU4-2* variants and dominant ReNU syndrome

RNU4-2-related neurodevelopmental disorder was independently described in large genome-sequenced cohorts in 2024. Chen et al.[17] identified *de novo* variants in an 18-bp region of *RNU4-2* in individuals with syndromic NDD, while Greene et al.[16] identified *RNU4-2* as the strongest non-coding gene associated with intellectual disability in a burden analysis of rare variants. The disorder is now commonly referred to as ReNU syndrome.

The most recurrent variant is n.64_65insT, a single-base insertion located within the T-loop/quasi-pseudoknot region[11]. This variant accounts for the majority of reported pathogenic *RNU4-2* cases and is absent or extremely depleted in population databases, consistent with strong negative selection. For better visualization, n.64_65insT is mapped on Figure 2.

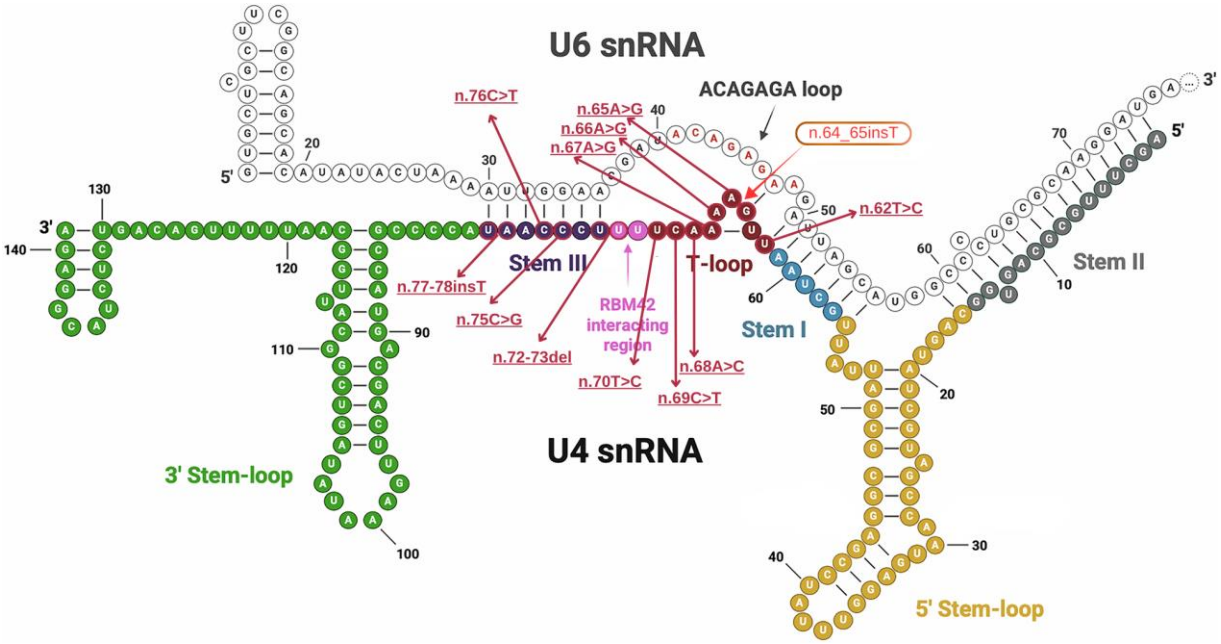


Figure 2. **Structural localization of reported pathogenic *RNU4-2* variants within the U4/U6 snRNA duplex.** The schematic shows the secondary structure of U4 snRNA in complex with U6 snRNA, highlighting the main functional domains of U4, including the 5′ stem-loop, Stem I, Stem II, Stem III, the T-loop/quasi-pseudoknot region, the RBM42 interaction region, and the 3′ stem-loop. Reported disease-associated *RNU4-2* variants are indicated by arrows and labelled according to U4 snRNA nucleotide numbering. Adapted from Chen, Dawes and Kim[17].

The central *RNU4-2* pathogenic hotspot maps to nucleotides n.62-n.79, a region involved in maintaining the architecture of the U4/U6 complex and positioning U6 for 5′ splice-site recognition. The hotspot includes Stem I, the T-loop/quasi-pseudoknot, the RBM42 interaction region, and Stem III[11]. Notably, variants located in the T-loop/quasi-pseudoknot and Stem III are believed to compromise the structural organization of the particle required to preserve splicing fidelity[4]. While mutations in Stem III may destabilize U4/U6 base pairing and lower the energy barrier that prevents activation of suboptimal splice sites, those near the quasi-pseudoknot may weaken the architecture that helps the correct positioning of the U6 ACAGAGA box for 5′ splice-site recognition[17].

Pathogenic *RNU4-2* variants are linked to altered spliceosome activation and abnormal 5′ splice-site selection[11]. RNA-sequencing studies showed increased use of unannotated cyptic 5′ splice sites in patients, validating the idea that the U4 central region is required for correct positioning of U6 during early catalytic activation. These findings connect the domain location

of variants directly to a measurable transcriptomic defect and reinforce the idea of ReNU syndrome as a spliceosome activation disorder.

Clinically, ReNU syndrome is characterized by developmental delay or intellectual disability, often moderate to severe, together with variable combinations of hypotonia, seizures, microcephaly, short stature, brain MRI abnormalities, autistic features, dysmorphic facial features and other neurological findings[18] (Fig. 3).

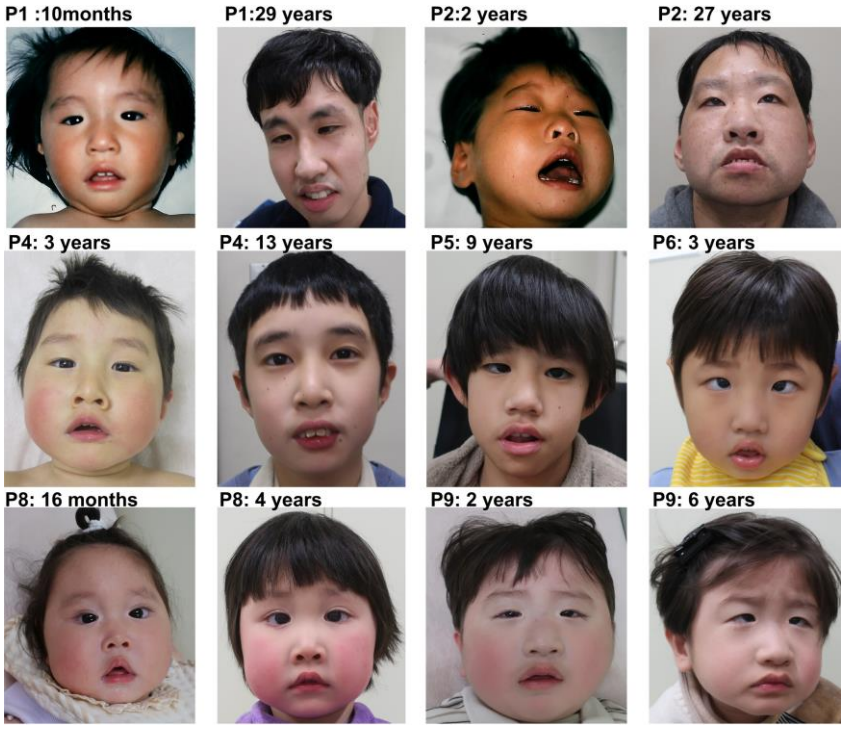


Figure 3. **Longitudinal facial phenotype of individuals affected by ReNU syndrome.**
Representative facial photographs of patients at different ages, showing the evolution of craniofacial features from infancy to childhood/adolescence/adulthood. Adapted from Okamoto et al. (2025)[19].

The recurrent n.64_65insT variant is generally associated with a severe phenotype, frequently including neonatal hypotonia, severe language impairment and abnormal brain imaging. Building on this, more recent genotype-phenotype analyses suggest that disease severity is influenced by variant location within the U4 structure. Variants in the T-loop/quasi-pseudoknot and RBM42 interaction region tend to cluster with more severe phenotypes, whereas some Stem III variants are associated with relatively milder neurodevelopmental presentations[11].

Accordingly, domain-based variant analysis is of particular interest when interpreting the "pathogenic landscape" of *RNU4-2*. Table 4 summarizes the most frequently reported pathogenic mutations, correlating each with its position in the structure, and its functional relevance[11].

Table 4. Main disease-relevant ReNU regions and examples of reported variants.

Region	Functional relevance	Examples of reported variants
Stem I	Contributes to U4/U6 interaction and the critical structural region	n.62T>C
T-loop / quasi-pseudoknot	Maintains central U4 architecture; associated with severe ReNU phenotype	n.64_65insT, n.65A>G, n.66A>G, n.67A>G
RBM42 interaction region	Potential RNA-protein interaction region involved in spliceosomal function	n.68A>C, n.69C>T, n.70T>C
Stem III	Supports U4/U6 structural stability; often associated with relatively milder phenotypes	n.72_73del, n.75C>G, n.76C>T, n.77_78insT

The estimated contribution of pathogenic *RNU4-2* variants to NDD is unusually high for a single non-coding gene. The initial studies suggested that variants in the critical region may explain approximately 0.4% of individuals with NDD, meaning that many patients worldwide could receive a diagnosis if *RNU4-2* is included in diagnostic pipelines or specifically screened in unresolved cases[12]. This prevalence is particularly important because the gene is not reliably assessed by standard exome sequencing. Consequently, patients with compatible phenotypes and negative prior genetic testing may benefit from WGS reanalysis or targeted sequencing of *RNU4-2*[20].

An additional notable feature of *RNU4-2*-related disease is the recurrent observation of *de novo* variants occurring preferentially on the maternal allele[17]. In the first large study, all phaseable *de novo* variants were maternal in origin, and later work confirmed a strong maternal predominance while also reporting rare paternal events for less severe variants. The biological explanation remains uncertain and may involve parent-of-origin effects, germline selection, or differential tolerance of pathogenic variants during gametogenesis. Although this

mechanism is still unresolved, it reinforces the need to interpret *RNU4-2* variants within both molecular and inheritance contexts.

1.4.4 Expansion of the *RNU4-2* disease spectrum: recessive *RNU4-2*-associated neurodevelopmental disorder

While *RNU4-2* was first associated with neurodevelopmental disease through heterozygous *de novo* variants causing dominant ReNU syndrome, evidence reported in 2026[15] broadened the genotypic and phenotypic spectrum of *RNU4-2*-related disorders. Biallelic variants in *RNU4-2* have been linked to a distinct autosomal recessive neurodevelopmental syndrome with variants lying outside the typical dominant hotspot. These affect other functionally important U4 elements, including Stem II, the k-turn within the 5′ stem-loop and the Sm protein binding site.

ReNU syndrome and the recessive *RNU4-2*-associated illness share some clinical similarities. Both include neurodevelopmental impairment, language delay, hypotonia, seizures, and other neurological manifestations. However, the recessive disorder has also been reported to show distinctive white matter abnormalities, like enlarged perivascular spaces, which may help distinguish it from dominant ReNU syndrome in certain patients. Consequently, at least two distinct clinical-genetic entities should be included in the current understanding of *RNU4-2*-related disease: a dominant disorder primarily caused by *de novo* variants in the central hotspot and a less common recessive disorder caused by biallelic variants in other functional regions of U4 snRNA.

Diagnostically, this expanded model is of relevance for patients with unresolved neurodevelopmental disorders. Variants located outside the dominant ReNU hotspot should not automatically be dismissed, especially when they fall within Stem II, the k-turn, or the Sm protein binding site. Nevertheless, mutations like these must be interpreted with care and caution, since heterozygous variants outside the hotspot are not sufficient on their own to establish a diagnosis of recessive *RNU4-2*-associated disease. Determining pathogenicity still requires functional proof, segregation analysis, population frequency assessment, and confirmation of biallelic status.

Taken together, these findings show that *RNU4-2* variant interpretation is evolving rapidly. The original disease model focused mainly on dominant *de novo* variants in the central ReNU

hotspot, whereas newer data indicate that additional U4 domains may also be clinically relevant when affected in a biallelic state.

2. AIM OF THE STUDY

The present thesis focuses on the molecular investigation of *RNU4-2* in a cohort of patients with neurodevelopmental disorders that remained genetically unresolved after previous diagnostic work-up. By combining Sanger sequencing with literature-based interpretation of *RNU4-2* structure and function, the study aims to evaluate whether variants in this non-coding spliceosomal RNA gene may contribute to the clinical phenotype of these patients.

Specifically, the work aims to identify and confirm *RNU4-2* variants, map them onto functionally relevant U4 snRNA domains, compare them with the published spectrum of dominant and recessive *RNU4-2*-related disorders, and assess their possible clinical relevance. This approach reflects the broader transition in medical genetics from a coding-gene-centerd model toward a more complete interpretation of the genome, including non-coding RNA genes with essential roles in neurodevelopment.

3. MATERIALS AND METHODS

3.1 Patient selection and clinical data review

This study included a subset of 300 unresolved patients with epilepsy and neurodevelopmental disorders referred from multiple Italian hospitals to the Genetics Laboratory between January 2019 and May 2026.

All patients had previously undergone molecular testing using a custom Next-Generation Sequencing (NGS) panel comprising 1,568 genes associated with neurodevelopmental disorders, including both established disease-causing genes and candidate genes involved in neuronal development, synaptic function, and RNA processing pathways.

Patients included in the present study had tested negative for pathogenic single nucleotide variants (SNVs) and small insertions/deletions (indels), or carried variants considered insufficient to fully explain their clinical phenotype. As *RNU4-2* was not included in this custom NGS panel, additional analysis of this gene was subsequently performed by Sanger sequencing.

Written informed consent for genetic testing was obtained from all patients or their legal guardians. Additional informed consent was acquired for segregation analyses performed in available family members.

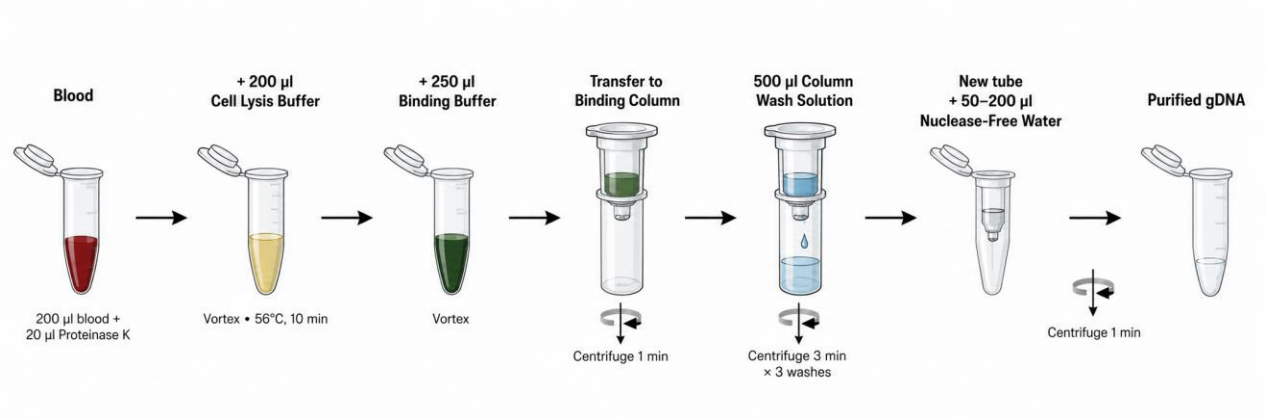
3.2 Genomic DNA extraction from peripheral blood

Genomic DNA was extracted from peripheral whole blood samples using a silica membrane-based spin column kit (Promega, Madison, WI, USA), according to the manufacturer's instructions (Fig. 4).

Briefly, 200 µL of whole blood was transferred into a 1.5 mL microcentrifuge tube, followed by the addition of 20 µL Proteinase K and 200 µL cell lysis buffer. Samples were vortexed thoroughly and briefly centrifuged to collect contents at the bottom of the tube, then incubated at 56 °C for 10 minutes. Following incubation, 250 µL of binding buffer was added, and the mixture was vortexed and centrifuged.

The lysate was transferred onto a silica-based spin column and centrifuged at 13,000 rpm for 1 minute. The flow-through was discarded, and the column was washed three times with 500 µL column wash buffer, each followed by centrifugation at 13,000 rpm for 3 minutes. A final 1 minute dry spin at 13,000 rpm was performed.

For elution, the spin column was placed into a clean 1.5 mL microcentrifuge tube. A volume of 50 µL nuclease-free water was applied directly onto the silica membrane and then centrifuged at 13,000 rpm for 1 minute, and the eluate containing purified genomic DNA was



collected. Extracted DNA was stored at −20 °C until downstream analysis.

Figure 4. **Workflow of genomic DNA extraction from peripheral blood using the ReliaPrep™ Blood gDNA Miniprep System.** Peripheral whole blood is mixed with Proteinase K and Cell Lysis Buffer, followed by incubation at 56 °C to promote cell lysis and protein digestion. Binding Buffer is then added to create conditions suitable for gDNA binding to the binding column. After centrifugation, DNA remains bound to the column membrane while contaminants are removed through sequential washing steps. Finally, genomic DNA is eluted with nuclease-free water and collected for downstream molecular analyses. Created with BioRender.com.

3.3 DNA quantification and quality assessment

DNA concentration and purity were measured using a NanoDrop One spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). The instrument was blanked using 1.5 µL of nuclease-free water. To verify baseline accuracy, a second aliquot of water was measured and confirmed to produce a flat absorbance spectrum. For each sample, 1.5 µL of DNA solution was loaded onto the pedestal, and absorbance was measured at 260 nm. DNA concentration (ng/µL) and purity ratios (A260/A280 and A260/A230) were recorded. Pedestals were cleaned with lint-free wipes between measurements (Fig. 5).

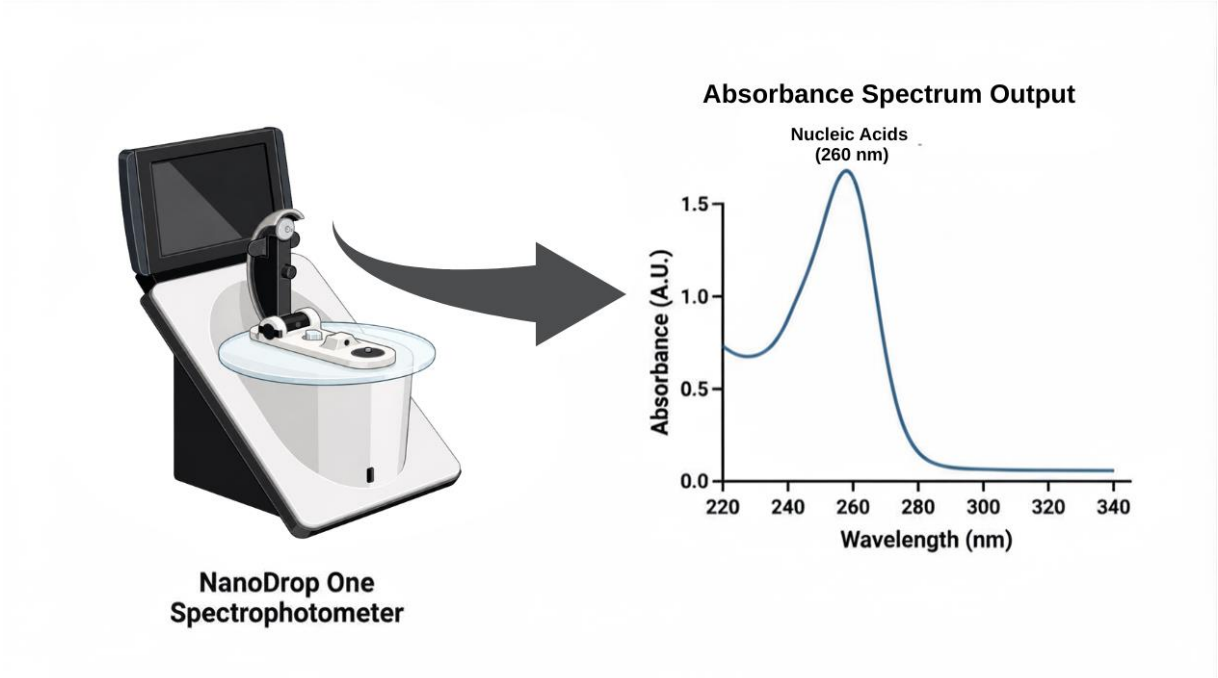


Figure 5. **DNA quantification and purity assessment using the NanoDrop One spectrophotometer.** A small-volume DNA sample is loaded directly onto the measurement pedestal, where UV–Vis absorbance is recorded across a defined wavelength range. The resulting absorbance spectrum is used to estimate nucleic acid concentration at 260 nm and to assess sample purity through absorbance ratios, A260/A280 and A260/A230. Created with BioRender.com.

For DNA samples, the A260/A280 ratio was considered an indicator of protein contamination, with an ideal value of approximately 1.8 for pure DNA. The A260/A230 ratio was used as an additional indicator of sample purity. Values between approximately 2.0 and 2.2 were considered acceptable for high-quality nucleic acid samples. Lower A260/A230 ratios may indicate the presence of residual contaminants, such as phenol, salts, carbohydrates, or other organic compounds carried over from the extraction procedure.

Samples with DNA concentration and quality deemed suitable for amplification and sequencing were selected for downstream applications.

3.4 PCR amplification of *RNU4-2*

The genomic region containing *RNU4-2* was amplified by polymerase chain reaction (PCR) using the following primers: forward 5′-CCTGAATGTTTGACACGGGA-3′ and reverse 5′-GAACCCCGGACATTCAATCA-3′. PCR reactions were prepared in a final volume of 15 µL per sample. The composition of the reaction mix for a single sample is reported in Table 5.

Table 5. PCR reaction mix used for amplification of the *RNU4-2* genomic region. Volumes refer to one sample.

Component	Volume per sample/reaction
Nuclease-free water	7.64 µL
5× Green GoTaq® Flexi Buffer	3.00 µL
PCR nucleotide mix	1.20 µL
$MgCl_2$ solution	0.90 µL
GoTaq® DNA Polymerase	0.06 µL
Primer mix	1.20 µL
Template DNA	1.00 µL
Final volume	**15.00 µL**

Amplification was carried out using a Mastercycler X50s (Eppendorf SE, Hamburg, Germany) under the following conditions: initial denaturation at 96 °C for 5 minutes, followed by 33 cycles of denaturation at 96 °C for 30 seconds, annealing at 58 °C for 30 seconds, and extension at 72 °C for 30 seconds, with a final extension at 72 °C for 7 minutes. Negative controls were included in all runs.

3.5 Agarose gel electrophoresis

PCR products were analyzed by agarose gel electrophoresis using a 1.5% agarose gel prepared in TAE buffer and stained with SYBR Green (Thermo Fisher Scientific, Waltham, MA, USA). Electrophoresis was performed, normally, at 130 V for 15 minutes, and amplified fragments were visualized under UV transillumination using the ChemiDoc imaging system (Bio-Rad Laboratories, Inc., Hercules, CA, USA).

3.6 PCR product purification

PCR products were purified using ExoSAP-IT™ (Thermo Fisher Scientific, Waltham, MA, USA). Prior to use, the reagent was diluted 1:20 in nuclease-free water. Subsequently, 1 µL of PCR product was added to 5 µL of the diluted ExoSAP-IT. The mixture was incubated at 37 °C for 15 minutes to allow enzymatic degradation of residual primers and dephosphorylation

of unincorporated nucleotides, followed by enzyme inactivation at 80 °C for 15 minutes in a thermal cycler.

3.7 Sanger sequencing

Purified PCR products were subjected to Sanger sequencing using BigDye™ Terminator v1.1 chemistry (Applied Biosystems, Foster City, CA, USA). Sequencing reactions were prepared in a final volume of 10 µL per sample/reaction using either the forward or reverse primer.

The composition of the sequencing reaction mix is reported in Table 6.

Table 6. Sanger sequencing reaction mix. Volumes refer to one sample.

Component	Volume per sample/reaction
BigDye™ Terminator v1.1	0.6 µL
Sequencing buffer	0.4 µL
Nuclease-free water	6.0 µL
Purified PCR product	1.0 µL
Primer (forward or reverse)	1.0 µL
Final volume	10.0 µL

Cycle sequencing was performed under the following conditions: initial denaturation at 96 °C for 3 minutes, followed by 25 cycles of denaturation at 96 °C for 20 seconds, annealing at 55 °C for 5 seconds, and extension at 60 °C for 4 minutes.

Post-reaction purification was performed using the SAM™/XTerminator™ system (Applied Biosystems, Foster City, CA, USA). A volume of 45 µL SAM solution was dispensed into the wells of an optical 96-well sequencing plate according to the number of amplicons processed, followed by the addition of 10 µL XTerminator reagent and 10 µL sequencing reaction product.

The plate was sealed, mixed at 1800 rpm for 30 minutes, and centrifuged at 1000 rpm for 2 minutes. The purified supernatant was directly analyzed by capillary electrophoresis using the SeqStudio Genetic Analyzer (Applied Biosystems, Foster City, CA, USA).

3.8 Variant analysis, classification, and genotype-phenotype correlation

Sequence chromatograms were analyzed using Chromas 2.6.6 trace viewer software and compared with the *RNU4-2* reference sequence based on the GRCh37/hg19 reference assembly.

Rare variants were prioritized by filtering against public population databases (1000 Genomes Project, ExAC, gnomAD), retaining only those with a minor allele frequency (MAF) <1%.

Variant pathogenicity was assessed according to the American College of Medical Genetics and Genomics (ACMG) guidelines (Richards et al., 2015), integrating cpopulation frequency data, computational evidence, segregation information, and variant-specific evidence when available. Additional interpretative support was obtained from Franklin by Genoox (https://franklin.genoox.com/clinical-db/home) and VarSome (https://sso.varsome.com/), which provide curated clinical annotations. Variants were classified as Pathogenic (P), Likely Pathogenic (LP), Variant of Uncertain Significance (VUS), Likely Benign (LB), or Benign (B).

For patients in whom an *RNU4-2* variant was identified, genotype–phenotype correlation was assessed by comparing the observed clinical features with the published spectrum of *RNU4-2*-related neurodevelopmental disorder/ReNU syndrome. The clinical features considered included developmental delay, intellectual disability, motor impairment, seizures, hypotonia, microcephaly, short stature, and characteristic dysmorphic features.

4. RESULTS

4.1 Patients cohort

The study cohort comprised 300 patients enrolled from the Clinic of Neuropsichiatry of the Maggiore Hospital in Novara based on a clinical suspicion of neurodevelopmental disorders, including intellectual disability, developmental delay, autism spectrum disorder, epilepsy, or other manifestations.

4.2 *RNU4-2* variants identification

In this cohort, no pathogenic or likely pathogenic variants were identified in *RNU4-2*. However, a total of 8 variants of uncertain significance (VUS) were detected in 7 of the 300 patients included in the study cohort, corresponding to 2.33% of the cohort. These individuals comprised 4 males (57.1%) and 3 females (42.8%). Only rare heterozygous variants with a minor allele frequency (MAF) below 1% in publicly available population databases were considered. The variants included n.12G>A, n.16C>G, n.29A>G, n.94A>C, n.96T>A, n.134T>C and n.138G>A, with n.29A>G observed in two unrelated patients and two variants, n.96T>A and n.12G>A, detected in the same patient. The associated clinical features were heterogeneous but mainly included developmental delay, speech or language delay, intellectual disability, autism spectrum disorder risk, hypotonia, epilepsy or seizures, dysmorphic features and behavioural abnormalities. A detailed overview of the identified variants, population frequencies, inheritance data and clinical signs is provided in Table 7.

Table 7. RNU4-2 variants identified in the study cohort.

Patient ID	Variant	Variant Classification (ACMG)*	Inheritance	Frequency (GnomAD)	Clinical sign
95-23	n.96T>A n.12G>A	VUS VUS	n.a.	- -	Developmental delay; speech delay
913-24	n.94A>C	VUS	Maternal	0.02%	Global psychomotor delay; hypotonia, facial dysmorphisms
380-22	n.134T>C	VUS	Maternal	0.1%	Mild global developmental delay

Patient ID	Variant	Variant Classification (ACMG)*	Inheritance	Frequency (GnomAD)	Clinical sign
755-23	n.29A>G	VUS	n.a.	0.02%	Mild intellectual disability; conduct disorder; macrocrania; hypertelorism
56-25	n.16C>G	VUS	Maternal	-	Language delay; moderate/severe ASD risk; possible hearing impairment
76-20	n.138G>A	VUS	n.a.	-	Generalized epilepsy; tonic-clonic seizures; personality disorder
927-24	n.29A>G	VUS	n.a.	0.02%	Developmental delay; speech delay; dysmorphism

* American College of Medical Genetics and Genomics (ACMG) guidelines (Richards et al., 2015)

In addition to the VUS identified, two recurrent variants were detected in the cohort: chr12:120729708G>A, observed in 4 patients and classified as benign, and chr12:120729707G>A, observed in 3 patients and classified as likely benign. Given their established benign or likely benign classification, these variants were not considered candidate disease-associated variants and were excluded from further analyses.

4.3 Segregation analysis

Segregation analysis was performed in cases for which DNA samples from family members were available. The variants n.16C>G (patient 56-25), n.94A>C (patient 913-24), and n.134T>C (patient 380-22) were found to be maternally inherited. Intriguingly, the majority of pathogenic *RNU4-2* variants reported to date in ReNU syndrome have arisen *de novo* in maternal allele. The familial transmission of these variants, particularly from apparently unaffected mothers, supports their current classification as variants of uncertain significance (VUS).

Nevertheless, inheritance alone does not exclude a potential contribution to the phenotype. Variable expressivity, incomplete penetrance, or the presence of subtle clinical manifestations in carrier mothers cannot be ruled out and would require further clinical characterization and functional studies to be assessed.

4.4 Structural localization of the identified variants

To evaluate the potential functional relevance of the VUS identified in this cohort, all variants were mapped onto the predicted secondary structure of U4 snRNA and its interaction with U6 snRNA (Fig. 6).

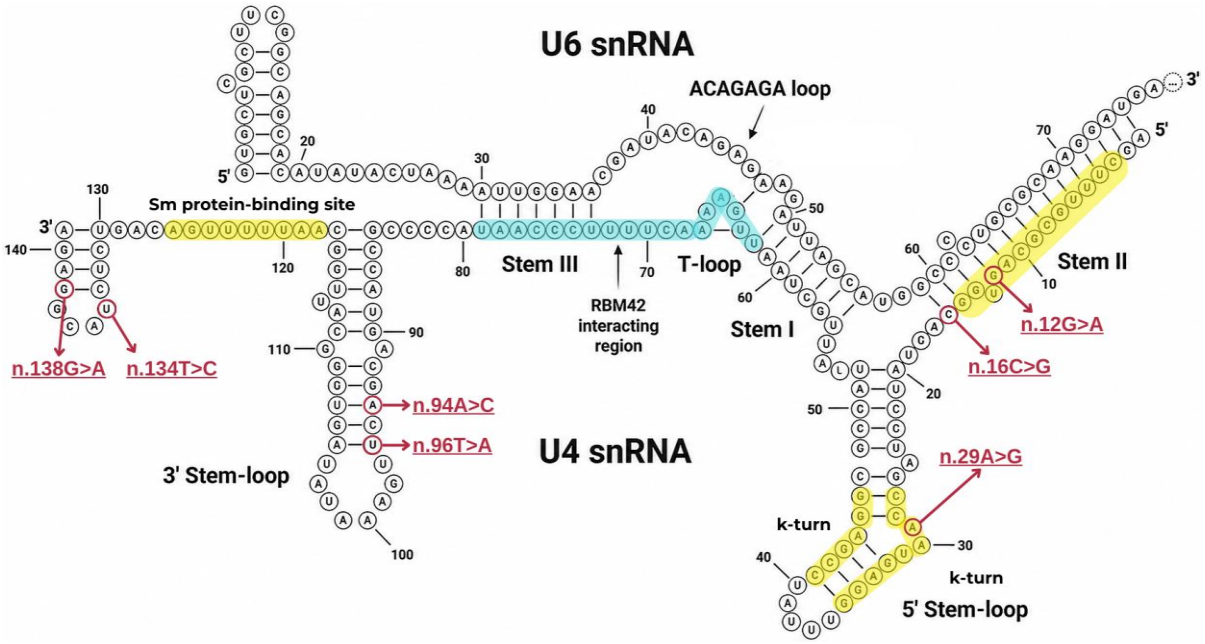


Figure 6. **Structural localization of *RNU4-2* variants identified in the cohort.** Identified variants are underscored in red. The central U4 region associated with dominant *RNU4-2*-related ReNU syndrome is highlighted in cyan, including the T-loop/quasi-pseudoknot, RBM42 interaction region and Stem III. Yellow highlights regions implicated in recessive *RNU4-2*-associated neurodevelopmental disorder, including the 5′ stem-loop/k-turn region, Stem II and the Sm protein-binding site. Adapted from Chen, Dawes and Kim[17] and interpreted in light of Rius et al[15].

The variants n.12G>A and n.16C>G localize to Stem II and to the boundary between Stem II and the 5′ stem-loop, respectively. Stem II contributes to the U4/U6 duplex architecture and to the maintenance of U6 snRNA in an inactive, activation-competent conformation before catalytic spliceosome activation.

The recurrent n.29A>G variant detected in two patients (Tab. 7) maps to the U4 5′ stem-loop, specifically within the k-turn region. This structural element contributes to the local architecture of U4 snRNA and may be relevant for RNA folding and protein-RNA interactions during snRNP maturation and spliceosome assembly. The detection of n.29A>G in two unrelated patients, corresponding to a frequency of 0.6% of the investigated patients, makes it a recurrent finding in this local dataset, much higher than that observed in the control

population (0,002%), representing an interesting finding that warrants further investigation (Tab. 7).

The variants n.94A>C and n.96T>A map to the 3′ stem-loop, a region that contributes to the structural stability of the mature U4 snRNA molecule.

The variants n.134T>C and n.138G>A are located closer to the terminal portion of the transcript and are therefore more appropriately assigned to the terminal stem-loop / 3′ region rather than to the canonical 3′ stem-loop. This region may contribute to the terminal folding and overall stability of U4 snRNA. Variant distribution is illustrated in Table 8.

Table 8. Structural localization of *RNU4-2* variants identified in the cohort.

Variant	**Structural localization**
n.12G>A	Stem II
n.16C>G	Stem II / boundary of 5′ stem-loop
n.29A>G	5′ stem-loop / k-turn
n.94A>C	3′ stem-loop
n.96T>A	3′ stem-loop
n.134T>C	Terminal stem-loop / 3′ region
n.138G>A	Terminal stem-loop / 3′ region

Overall, none of the variants identified in the present cohort mapped to the classical dominant ReNU syndrome hotspot, corresponding to nucleotides n.62–n.79. However, some variants were located within structural domains that have recently been implicated in recessive *RNU4-2*-associated neurodevelopmental disorder, highlighting their potential relevance despite their position outside the canonical hotspot. Specifically, n.12G>A and n.16C>G map to Stem II, a region involved in maintaining the secondary structure of U4 snRNA, whereas n.29A>G is located within the k-turn region of the U4 5′ stem-loop. Given the role of these structural elements in U4 snRNA folding, snRNP biogenesis, and spliceosome assembly, the presence of variants in these regions may be of functional interest and supports further investigation of their potential contribution to the neurodevelopmental phenotypes observed in this cohort.

5. DISCUSSION

This study aimed to investigate the presence of potentially disease-associated variants in the spliceosomal non-coding RNA gene *RNU4-2* in a cohort of 300 patients with neurodevelopmental disorders. No pathogenic or likely pathogenic variants were identified in the cohort. However, eight variants at the heterozygous state of uncertain significance were detected in seven patients, corresponding to 2.33% of the analysed cohort. Even though these findings do not provide a definitive molecular diagnosis, they provide valuable context to the rapidly expanding literature on spliceosomal snRNA-associated neurodevelopmental disorders.

5.1 Structural distribution of the identified *RNU4-2* variants

The patients were first grouped according to the localization of their mutations within the U4 architecture.

The first group included variants located in the 5′ region, namely n.12G>A and n.16C>G. This region contributes to the formation and stabilization of the U4/U6 duplex, which is necessary to maintain U6 in an inactive but activation-competent conformation before spliceosome activation. Variants affecting this region may therefore interfere with local RNA folding, U4/U6 pairing or the correct structural organization of the pre-catalytic spliceosome[4,15,21].

The second group included a recurring variant located in the k-turn region of the 5′ stem-loop, represented by n.29A>G. Overall the 5′ stem-loop contributes to the structural organization of U4 snRNA, while the k-turn, that is part of the 5′ stem-loop, represents an RNA architectural motif that may help maintain local conformation and support interactions required for snRNP assembly. Disruption of this region could affect U4 folding, RNA-protein interactions or incorporation of U4 into functional spliceosomal particles[15,21].

The third group included those with variants located in the 3′ stem-loop region, comprising n.94A>C, n.96T>A, n.134T>C and n.138G>A. This region contributes to the structural stability of U4 snRNA and may be relevant for correct snRNA maturation, stability and incorporation into spliceosomal ribonucleoprotein complexes. Variants in this domain could potentially impair U4 availability or function within the spliceosome[15,21].

5.2 Domain-based clinical interpretation

When the variants were considered according to U4 snRNA architecture, some domain-related clinical patterns could be observed.

Variants in the 5′ region, which contributes to U4/U6 duplex formation and stabilization[21], were mainly associated with speech and language delay, together with broader neurodevelopmental involvement. Since this domain participates in maintaining the pre-catalytic spliceosome in an inactive but activation-competent state, disruption of this region could affect correct spliceosome assembly or activation, potentially influencing the expression of genes important for neurodevelopment and language acquisition.

Variants in the k-turn region were associated with developmental and/or speech delay together with dysmorphic features. This may be of interest because this region contributes to local U4 folding[21] and may support RNA-protein interactions required for snRNP assembly. As such, alterations in this domain could interfere with U4 structural organization and lead to developmental consequences.

Variants in the 3′ stem-loop region showed a broader neurological pattern, including motor delay, language or communication impairment, global developmental delay and, in some cases, epilepsy or psychiatric manifestations. Since this region may contribute to U4 snRNA stability, maturation and incorporation into spliceosomal ribonucleoprotein complexes[21], its disruption could affect overall U4 availability or function, potentially explaining the wider neurological involvement observed in this group.

Dominant ReNU syndrome is primarily associated with heterozygous de novo variants clustered within the central n.62–n.79 region of *RNU4-2*[11]. In the present cohort, none of the identified variants mapped to this established dominant hotspot. Nevertheless, this finding does not exclude a possible contribution of *RNU4-2* to the patients' phenotypes. The interpretation of non-coding snRNA variants remains challenging, and pathogenicity may depend not only on their position within known disease-associated regions, but also on their potential effects on RNA folding, snRNP assembly, transcript stability, protein binding, or interactions with other spliceosomal components[22].

Several considerations support the biological relevance of variants located outside the classical dominant hotspot. Variants affecting structurally important regions of U4 snRNA may subtly

influence RNA secondary structure or U4/U6 duplex stability, without necessarily producing the same molecular signature described in classical ReNU syndrome. In addition, such variants may act as genetic modifiers rather than fully penetrant monogenic causes, contributing to neurodevelopmental susceptibility in combination with other coding or non-coding variants elsewhere in the genome[23]. Finally, variants in other snRNA genes or spliceosome-related components, not assessed in the present targeted analysis, may also cooperate in shaping the clinical presentation[10].

This interpretation is further supported by the recent description of a recessive *RNU4-2*-associated neurodevelopmental disorder, distinct from dominant ReNU syndrome. Rius et al.[15] reported that biallelic *RNU4-2* variants, either homozygous or compound heterozygous, can cause a recessive neurodevelopmental phenotype. Unlike dominant ReNU syndrome, in which pathogenic variants cluster within the central hotspot, the recessive disorder is associated with variants located in other functionally relevant U4 snRNA elements, including Stem II, the k-turn region, and the Sm protein-binding site. This distinction is particularly relevant to the present study, as some of the variants identified in our cohort map to regions implicated in the recessive *RNU4-2*-related disorder.

Specifically, n.12G>A and n.16C>G localize to Stem II or to the boundary between Stem II and the 5′ stem-loop, while n.29A>G maps to the k-turn region. The latter variant was detected in two unrelated patients, corresponding to a frequency of 0.6% in the investigated cohort, which is considerably higher than that observed in the control population, 0.002% (Tab.7). This represents an interesting recurrent finding in this local dataset and warrants further investigation in larger cohorts. The n.29A>G variant is also noteworthy because it lies close to n.27C>G, n.28C>G, and n.32G>A, which were reported in patients with the recessive *RNU4-2*-associated disorder (Fig. 6). Together, these variants cluster around the k-turn region of the U4 5′ stem-loop, a structural domain that may be important for U4 snRNA folding and snRNP assembly. From a clinical perspective, patients carrying n.29A>G in the present cohort showed developmental or speech delay and dysmorphic features, partially overlapping with manifestations reported in recessive *RNU4-2*-associated disease. However, the variants identified in our cohort were present in the heterozygous state, whereas the recessive disorder requires biallelic pathogenic variation. Therefore, n.29A>G cannot currently be interpreted as

causative on its own, although its recurrence and localization support continued attention to this region.

Finally, n.138G>A lies close to n.140G>A, a variant reported in the recessive disease cohort. Both variants are located near the terminal 3′ region of U4 snRNA, which may be relevant for snRNA maturation and structural stability. This proximity raises the possibility that the distal region of U4 snRNA may also contribute to disease-related mechanisms. Although a definitive interpretation is not possible in the absence of confirmed biallelic variation or functional validation, the localization of these variants near regions implicated in recessive *RNU4-2*-associated disease suggests that they should not be dismissed and may provide useful clues for future studies.

5.6 Limitations and future perspectives

The main consideration of this study is that the variants identified in the cohort are currently classified as variants of uncertain significance. This reflects the present challenges in interpreting non-coding snRNA variation, rather than excluding their potential biological relevance. Segregation data were available for only a subset of patients, which limited the possibility of establishing whether the variants occurred *de novo* or were inherited from an unaffected or mildly affected parent. Extending segregation analysis to additional family members will therefore be important to better define the clinical relevance of these findings.

Another aspect to consider is that functional validation was not performed in the present study. As a result, the potential effects of the identified variants on U4 snRNA structure, expression, molecular interactions, or spliceosome function could not be directly assessed. Nevertheless, the localization of some variants within structurally or functionally relevant regions of U4 snRNA provides a rationale for prioritizing them in future experimental studies. Approaches such as small RNA-seq, RNA structure probing, or assays evaluating snRNP assembly and splicing activity may help clarify their biological impact.

The analysis was focused specifically on *RNU4-2*, in line with the aim of the study. However, recent evidence linking other spliceosomal snRNA genes, including *RNU2-2*, *RNU5B-1*, and the candidate gene *RNU5A-1*, to neurodevelopmental disorders suggests that broader investigation of snRNA genes may be valuable in unsolved cases. Expanding the analysis to

additional spliceosomal snRNA loci could therefore provide a more comprehensive view of the contribution of RNUopathies to neurodevelopmental phenotypes.

The broad and heterogeneous clinical presentation of the cohort should also be taken into account. Neurodevelopmental disorders are genetically and phenotypically diverse, and features such as developmental delay, epilepsy, hypotonia, language impairment, behavioural abnormalities, and dysmorphic features are not specific to *RNU4-2*-related disease. For this reason, partial clinical overlap with dominant ReNU syndrome or with the recessive *RNU4-2*-associated disorder cannot alone establish pathogenicity. However, when combined with variant localization in relevant structural domains and recurrence in the cohort, these clinical observations may help identify candidate cases for further investigation.

Overall, the variants reported here should not currently be considered diagnostic, but they represent potentially informative candidate findings. Future studies including complete segregation analysis, clinical evaluation of variant-positive relatives, updated comparison with population databases, and functional characterization of selected variants will be essential to clarify their significance. Given the rapid evolution of the field of RNUopathies, periodic reinterpretation of these variants will also be important, as new genetic, clinical, and functional data become available.

6. CONCLUSION

In conclusion, this study contributes to the growing investigation of non-coding spliceosomal RNA genes in neurodevelopmental disorders by analysing RNU4-2 in a cohort of 300 genetically unresolved patients. The identification of eight heterozygous RNU4-2 variants of uncertain significance in seven patients highlights the complexity of interpreting variation in small non-coding RNA genes, where pathogenicity may depend not only on nucleotide position, but also on RNA structure, inheritance pattern, transcript stability, spliceosome assembly and downstream effects on splicing.

Although the identified variants do not correspond to the classical dominant ReNU syndrome hotspot, some are located in regions of potential structural or functional relevance, including areas close to variants recently implicated in recessive RNU4-2-associated neurodevelopmental disease. These findings suggest that regions outside the established dominant hotspot may deserve further attention, particularly as the RNU4-2 disease spectrum continues to evolve.

Overall, this work supports the importance of including RNU4-2 and other spliceosomal snRNA genes in the molecular evaluation of unresolved neurodevelopmental disorders. Future studies involving larger cohorts, broader analysis of snRNA genes, segregation and functional stusied will be essential to clarify the biological relevance of rare non-coding variants. A better understanding of how spliceosomal snRNA variation contributes to neurodevelopmental phenotypes may improve variant interpretation, expand diagnostic strategies and help resolve currently unexplained cases.

BIBLIOGRAPHY

1. Pallanti, S., & Salerno, L. (2023). Neurodevelopmental Disorders (NDDs): Beyond the Clinical Definition and Translational Approach. *Children (Basel, Switzerland)*, *10*(1), 99. https://doi.org/10.3390/children10010099.
2. Gao, S., Shan, C., Zhang, R., & Wang, T. (2024). Genetic advances in neurodevelopmental disorders. *Medical review (2021)*, *5*(2), 139–151. https://doi.org/10.1515/mr-2024-0040.
3. Chin, I.M., Gardell, Z.A., and Corces, M.R. (2024). Decoding polygenic diseases: advances in noncoding variant prioritization and validation. Trends Cell Biol. 34, 465-483. doi: 10.1016/j.tcb.2024.03.005.
4. Wilkinson, M.E., Charenton, C., and Nagai, K. (2020). RNA splicing by the spliceosome. Annu. Rev. Biochem. 89, 359-388. doi: 10.1146/annurev-biochem-091719-064225.
5. Younis, I., Dittmar, K., Wang, W., Foley, S.W., Berg, M.G., Hu, K.Y., Wei, Z., Wan, L., and Dreyfuss, G. (2013). Minor introns are embedded molecular switches regulated by highly unstable U6atac snRNA. eLife 2, e00780. doi: 10.7554/eLife.00780.
6. Antonarakis, S. E. (2025). *Small nuclear RNA genes in Mendelian disorders*. Nature Genetics 58. https://doi.org/10.1038/s41588-025-02440-7.
7. Farach, L.S., Little, M.E., Duker, A.L., Logan, C.V., Jackson, A., Hecht, J.T., et al. (2018). The expanding phenotype of RNU4ATAC pathogenic variants to Lowry Wood syndrome. Am. J. Med. Genet. A 176, 465-469. doi: 10.1002/ajmg.a.38581.
8. Elsaid, M.F., Chalhoub, N., Ben-Omran, T., Kumar, P., Kamel, H., Ibrahim, K., et al. (2017). Mutation in noncoding RNA RNU12 causes early onset cerebellar ataxia. Ann. Neurol. 81, 68-78. doi: 10.1002/ana.24826.
9. Greene, D., De Wispelaere, K., Lees, J., Codina-Sola, M., Jensson, B.O., Hales, E., et al. (2025). Mutations in the small nuclear RNA gene RNU2-2 cause a severe neurodevelopmental disorder with prominent epilepsy. Nat. Genet. 57, 1367-1373. doi: 10.1038/s41588-025-02159-5.
10. Jackson, A., Thaker, N., Blakes, A., Rice, G., Griffiths-Jones, S., Balasubramanian, M., et al. (2025). Analysis of R-loop forming regions identifies RNU2-2 and RNU5B-1 as neurodevelopmental disorder genes. Nat. Genet. 57, 1362-1366. doi: 10.1038/s41588-025-02209-y.
11. Nava, C., Cogne, B., Santini, A., et al. (2025). Dominant variants in major spliceosome U4 and U5 small nuclear RNA genes cause neurodevelopmental disorders through splicing disruption. Nat. Genet. 57, 1374-1388. doi: 10.1038/s41588-025-02184-4.
12. Burns, V. F., & Radford, E. J. (2024). ReNU syndrome - a newly discovered prevalent neurodevelopmental disorder. *Trends in genetics : TIG*, *40*(11), 914–916. doi:10.1016/j.tig.2024.09.005.

13. Fischer, U., Englbrecht, C., and Chari, A. (2011). Biogenesis of spliceosomal small nuclear ribonucleoproteins. Wiley Interdiscip. Rev. RNA 2, 718-731. doi: 10.1002/wrna.87.

14. Burke, J.E., Butcher, S.E., and Brow, D.A. (2015). Spliceosome assembly in the absence of stable U4/U6 RNA pairing. RNA 21, 923-934. doi: 10.1261/rna.048421.114.

15. Rius, R., Blakes, A.J.M., Chen, Y. *et al.* Biallelic variants in the noncoding RNA gene *RNU4-2* cause a recessive neurodevelopmental syndrome with distinct white matter changes. *Nat Genet* 58, 761–773 (2026). doi:10.1038/s41588-026-02554-6.

16. Greene, D., Thys, C., Berry, I.R., Jarvis, J., Ortibus, E., Mumford, A.D., Freson, K., and Turro, E. (2024). Mutations in the U4 snRNA gene RNU4-2 cause one of the most prevalent monogenic neurodevelopmental disorders. Nat. Med. 30, 2165-2169. doi: 10.1038/s41591-024-03085-5.

17. Chen, Y., Dawes, R., Kim, H.C., Stenton, S.L., Walker, S., et al. (2024). De novo variants in the RNU4-2 snRNA cause a frequent neurodevelopmental syndrome. Nature 632, 832-840. doi: 10.1038/s41586-024-07773-7.

18. Valenzuela, I., Codina-Sola, M., Vazquez, E., Cueto-Gonzalez, A., Leno-Colorado, J., Lasa-Aranzasti, A., et al. (2024). Deep phenotyping of 11 individuals with pathogenic variants in RNU4-2 reveals a clinically recognizable syndrome. Genet. Med. 26, 101288. doi: 10.1016/j.gim.2024.101288.

19. Okamoto, N., E. Nishi, Y. Hasegawa, et al. 2025. "A Clinical Study of Nine Patients With ReNU Syndrome." *American Journal of Medical Genetics Part A*197, no. 11: e64151. https://doi.org/10.1002/ajmg.a.64151.

20. Fan, S., Yang, S., Nie, X., Yu, Z., Jiang, Y., Sun, M., Gu, W., & Zhang, X. (2026). Reanalysis of Whole Genome Sequencing Resolves Genetically Undiagnosed Patients With "RNUopathies". *Clinical genetics*, *109*(3), 576–580. https://doi.org/10.1111/cge.70069.

21. Nguyen, T. H., Galej, W. P., Bai, X. C., Savva, C. G., Newman, A. J., Scheres, S. H., & Nagai, K. (2015). The architecture of the spliceosomal U4/U6.U5 tri-snRNP. *Nature*, *523*(7558), 47–52. https://doi.org/10.1038/nature14548.

22. Bruselles A, et al. *Expanding the mutational spectrum of ReNU syndrome.* American Journal of Human Genetics. 2025. doi: 10.1038/s41431-025-01820-1.

23. Kingdom R, et al. *Genetic modifiers of rare variants in monogenic developmental disorder loci.* Nature Genetics. 2024;56:861–868. doi: 10.1038/s41588-024-01710-0.