

DÉPARTEMENT DE BIOLOGIE

Mémoire présenté en vue de l'obtention du Diplôme de Master

Domaine : Sciences de la Nature et de la Vie

Filière : Science biologique

Spécialité : Génétique Fondamentale et Appliquée

Présenté par

MERZOUK Karima

NEFOUSSI Hafsa

**Analyse in silico de l'impact des variants de RNU4-2 sur la structure et la fonction de
l'ARN U4 dans le syndrome ReNU**

Soutenu le 22/06/2026

DEVANT LE JURY COMPOSÉ DE :

Président	GUEDOUAR Youcef	MCA	U. Mostaganem
Examineur	MANSOURI Toufik	MAA	U. Mostaganem
Encadrant	ADDA NEGGAZ Leila	MCB	USTO-MB
Co-Encadrant	DAHMANI Chahinez Amira	MCA	U. Mostaganem

Année universitaire 2025/2026

Remerciements

Au terme de ce travail, nous tenons à remercier en premier lieu Dieu Tout-Puissant de nous avoir accordé l'aptitude et la volonté nécessaires pour mener à bien ce projet.

Nous tenons à exprimer notre grand respect et nos sincères remerciements aux membres du jury qui ont aimablement accepté d'évaluer ce travail, en dépit de leurs nombreuses obligations.

À Monsieur **GUEDOUAR Youcef**, d'avoir accepté de présider ce jury. Votre présence et l'intérêt que vous portez à notre travail représentent pour nous un grand honneur. Nous vous exprimons notre profonde reconnaissance et notre plus sincère respect.

Monsieur **MANSOURI Toufik**, nous vous remercions sincèrement d'avoir accepté d'examiner ce travail. Nous vous exprimons notre gratitude pour l'intérêt accordé à notre mémoire et pour le temps consacré à son évaluation.

Nos plus vifs remerciements et notre profonde gratitude s'adressent à notre encadrante, Madame **ADDA NEGGAZ Leila**, qui nous a fait l'honneur d'accepter de diriger ce travail. Nous la remercions chaleureusement pour son encadrement rigoureux et méthodique, ainsi que pour les précieuses compétences dont elle nous a fait bénéficier tout au long de notre parcours universitaire.

Nos sincères remerciements s'adressent également à Madame **Dahmani Chahinez Amira** pour avoir encadré ce travail. Nous lui exprimons également notre reconnaissance pour ses conseils fructueux, son soutien continu et ses encouragements permanents.

Enfin, nous tenons à remercier chaleureusement tous ceux qui ont contribué, de près ou de loin, à l'accomplissement et à la concrétisation de ce travail.

Dédicace

Je dédie mon diplôme à celui qui m'a appris que la vie est un combat dont la science et le savoir sont les armes : mon cher père, mon amour et mon pilier.

Je dédie mon diplôme à ma chère mère, qui m'a poussée et encouragée à achever ce long voyage.

Sans oublier ma sœur bien-aimée, Sabah, ma seconde mère et mon unique amie.

À ceux qui ne m'ont jamais rien refusé, à ceux qui ont lutté et tout donné pour mon confort : mes frères, Abdou et Mohamed.

À moi-même : merci à cette âme bienveillante d'avoir persévéré malgré la difficulté du chemin.

En cet instant où je récolte les fruits de longues années de travail et de persévérance, je tiens à dédier cette réussite aux êtres les plus chers à mon cœur.

Je dédie cette réussite à mon père et à ma mère, pour leur amour, leurs sacrifices et leur soutien indéfectible tout au long de mon parcours.

À ma sœur Batoul, pour sa présence, son affection et ses encouragements.

À la famille Benhamou et à la famille Nefoussi, pour leurs prières et leur soutien précieux.

Enfin, à moi-même, pour ma patience, ma persévérance et ma détermination à atteindre cet objectif.

Cette réussite est la vôtre autant que la mienne.

En mon nom et au nom de ma collègue, nous dédions ce travail à notre chère encadrante : merci pour votre soutien, vos précieux conseils et votre temps précieux. Que Dieu vous protège et veille sur vous.

Puissiez-vous rester pour moi une source de bonheur, de santé et d'amour tout au long de votre vie. Puissiez-vous rester, pour toujours, mon pilier inébranlable et infatigable.

Résumé

Le syndrome neurodéveloppemental associé à *RNU4-2* (syndrome ReNU) est une cause monogénique fréquente de déficience intellectuelle sévère, causée par des variants de novo affectant le gène *RNU4-2*, qui code l'ARN nucléaire U4, composant clé du spliceosome majeur. L'interprétation fonctionnelle de ces variants reste difficile, les outils standards de prédiction étant peu adaptés aux ARN non codants.

Cette étude analyse in silico 61 variants de *RNU4-2* issus de ClinVar, répartis en 30 variants pathogènes ou probablement pathogènes (Patho/LP) et 31 variants de signification incertaine (VUS), à l'aide de trois métriques : $\Delta\Delta G$ RNAfold (stabilité de la structure secondaire), $\Delta\Delta G$ RNAcofold (interaction U4/U6) et RNAdistance (réorganisation topologique).

Le $\Delta\Delta G$ RNAfold est la seule métrique distinguant significativement les deux groupes ($p = 0,012$; delta de Cliff = $-0,37$), mais les variants Patho/LP présentent un $\Delta\Delta G$ médian nul, suggérant que la déstabilisation globale n'est pas le mécanisme pathogène principal. Le signal le plus discriminant reste la localisation dans la région critique (T-loop + Stem III) : 18 variants Patho/LP sur 30 s'y concentrent, contre un seul VUS (OR = 45,0 ; $p = 9,7 \times 10^{-7}$). Les variants n.7G>C, n.140G>A et n.129G>C/A sont identifiés comme candidats prioritaires pour une validation expérimentale.

Ces résultats soulignent l'hétérogénéité des mécanismes moléculaires du syndrome ReNU et l'intérêt d'une approche multi-critères combinant analyse structurale et localisation fonctionnelle pour l'interprétation des variants de gènes d'ARN non codants.

Mots-clés : RNU4-2 · Syndrome ReNU · ARN non codant · Analyse in silico · Spliceosome · VUS · Structure secondaire ARN · Bioinformatique.

Résumé en arabe

تشمل اضطرابات النمو العصبي حالاتٍ مرتبطةً باضطراباتٍ في نمو الجهاز العصبي، وغالبًا ما تعود إلى عوامل وراثية. تُعدُّ الأحماض النووية الريبية غير المشفرة (RNA) مكوناتٍ أساسيةً في تنظيم التعبير الجيني، لا سيما من خلال آلية التضفير تحت إشراف الجسيم التضفيري. يُنتج جين RNU4-2 الحمض النووي الريبوي U4، وهو مكونٌ أساسيٌّ لهذا المركب، ضروريٌّ لتنظيم عملية التضفير واستقرارها. وقد رُبطت الطفرات المرضية في هذا الجين مؤخرًا بمتلازمةٍ جديدةٍ في النمو العصبي تُسمى ReNU، والتي تتجلى في تأخر النمو، وضعف الإدراك، واضطراباتٍ عصبيةٍ متنوعة. يُعدُّ تحليل هذه الطفرات أمرًا صعبًا نظرًا للطبيعة البنوية للحمض النووية الريبية غير المشفرة، ومن هنا تبرز أهمية الطرق الحاسوبية لتقييم تأثيراتها على استقرار الحمض النووي الريبوي وبنيته. لذلك، يهدف هذا المشروع إلى استكشاف اختلافات جين RNU4-2 حسابيًا لفهم دورها الوظيفي ومشاركتها في إمراضية متلازمة ReNU بشكلٍ أفضل.

Abstract

RNU4-2-related neurodevelopmental syndrome (ReNU syndrome) is a frequent monogenic cause of severe intellectual disability, caused by de novo variants affecting the RNU4-2 gene, which encodes the U4 small nuclear RNA, a key component of the major spliceosome. Functional interpretation of these variants remains challenging, as standard prediction tools were designed for protein-coding sequences and perform poorly on non-coding RNAs.

This study presents an in silico analysis of 61 RNU4-2 variants retrieved from ClinVar, classified as either pathogenic or likely pathogenic (Patho/LP, $n = 30$) or of uncertain significance (VUS, $n = 31$), using three complementary metrics: $\Delta\Delta G$ RNAfold (secondary structure stability), $\Delta\Delta G$ RNAcofold (U4/U6 interaction energy), and RNAdistance (topological reorganization).

$\Delta\Delta G$ RNAfold was the only metric significantly differentiating the two groups ($p = 0.012$; Cliff's delta = -0.37), yet Patho/LP variants showed a median $\Delta\Delta G$ of zero, suggesting that global RNA destabilization is not the predominant pathogenic mechanism. The strongest discriminating signal was variant localization within the critical region (T-loop + Stem III): 18 of 30 Patho/LP variants mapped to this region versus only one VUS (OR = 45.0; $p = 9.7 \times 10^{-7}$). Variants n.7G>C, n.140G>A, and n.129G>C/A were identified as priority candidates for experimental functional validation.

These findings highlight the mechanistic heterogeneity underlying ReNU syndrome and support a multi-criteria approach combining structural analysis and functional localization for the interpretation of variants in non-coding RNA genes.

Keywords: RNU4-2 · ReNU syndrome · Non-coding RNA · In silico analysis · Spliceosome · Variant of uncertain significance · RNA secondary structure · Bioinformatics.

Liste des Abréviations

A

ADN : Acide désoxyribonucléique
ARN : Acide ribonucléique
ARNm : Acide ribonucléique messenger
ARNmi / miRNA : Micro-ARN
ARNnc : ARN non codant
ACMG/AMP : American College of Medical Genetics / Association for Molecular Pathology

B

BMJ : British Medical Journal

C

ClinVar : Clinical Variation (base de données NCBI)
CNV : Copy Number Variation (variation du nombre de copies)

D

DNA : Deoxyribonucleic acid

E

ENCODE : Encyclopedia of DNA Elements
Ensembl : Base de données génomique européenne

F

FASTA : Format standard de séquence biologique

G

ΔG : Énergie libre de Gibbs

H

HGVS : Human Genome Variation Society (nomenclature des variants)

I

IC95 % : Intervalle de confiance à 95 %

IQR : Intervalle interquartile

IRM : Imagerie par résonance magnétique

L

lncRNA : Long non-coding RNA (ARN long non codant)
LP : Likely pathogenic (probablement pathogène)

M

MECP2 : Methyl-CpG binding protein 2
MENA : Middle East and North Africa
MFE : Minimum Free Energy (énergie libre minimale de repliement)

N

NDD / NDDs : Neurodevelopmental disorders
NCBI : National Center for Biotechnology Information
NHPX : Protéine 15.5K de liaison au k-turn de l'ARN U4
NLGN3 : Neuroligine 3
NNDB : Nearest Neighbor Database
NRXN / NRXN1 : Neurexine 1
nt : Nucléotide(s)

P

P : Pathogenic (pathogène)
P/LP : Pathogenic / Likely pathogenic
Patho/LP : Regroupement des variants P, P/LP et LP dans l'étude
pb : Paire de bases

R

ReNU : RNU4-2-related Neurodevelopmental syndrome
RNAcofold : Logiciel de prédiction de structure hybride ARN-ARN (ViennaRNA)

RNAdistance : Logiciel de calcul de distance structurale entre structures ARN (ViennaRNA)

RNAfold : Logiciel de prédiction de structure secondaire d'ARN (ViennaRNA)

RNP : Ribonucléoprotéique

RNU4-2 : RNA, U4 small nuclear 2 (gène codant l'ARN U4)

RNU6-1 : RNA, U6 small nuclear 1 (gène codant l'ARN U6)

S

SGE : Saturation Genome Editing

SHANK / SHANK3 : SH3 and multiple ankyrin repeat domains protein 3

snARN / snRNA : Small nuclear RNA (petit ARN nucléaire)

snoARN / snoRNA : Small nucleolar RNA (petit ARN nucléolaire)

snRNP : Small nuclear ribonucleoprotein

SNV : Single Nucleotide Variant

T

TDH : Trouble déficitaire de l'attention avec ou sans hyperactivité

TDN : Troubles du développement neurologique

TED : Tree Edit Distance (distance d'édition d'arbres)

TND : Troubles neurodéveloppementaux

TSA : Trouble du spectre de l'autisme

U

U1, U2, U4, U5, U6 : Petits ARN nucléaires du spliceosome majeur

U4/U6 : Duplex intermoléculaire ARN U4 / ARN U6

UNICEF : United Nations Children's Fund

V

VUS : Variant of Uncertain Significance (variant de signification incertaine)

W

WT : Wild-type (type sauvage)

Table des figures

Figure 1 : Manifestations cliniques, dysmorphie cranio-faciale et anomalies cérébrales.....	4
Figure 2 : Représentation schématique du mécanisme général de l'épissage constitutif des .	10
Figure 3 : Cycle d'assemblage	12
Figure 4 : Structure tridimensionnelle du duplex d'ARN snU4/U6 humain à l'état libre.....	14
Figure 5 : Structure secondaire du duplex U4/U6	14
Figure 6 : Page d'accueil du navigateur génomique Ensembl (Release 115).	19
Figure 7 : Interface de recherche de la base de données ClinVar (NCBI).....	20
Figure 8 : Interface de RNAfold.....	23
Figure 9 : Interface de RNAcifold.	25
Figure 10 : Carte des 61 variants RNU4-2 par région fonctionnelle	31
Figure 11 : $\Delta\Delta G$ RNAfold de chaque variant RNU4-2	32
Figure 12 : Analyse comparative des structures secondaires de l'ARN U4.....	33
Figure 13 : Distribution des trois métriques	37
Figure 14 : Comparaison des trois métriques	38

Table des tableaux

Tableau 1 : Principales caractéristiques cliniques des individus RNU4-2.....	3
Tableau 2 : Régions mutationnelles de RNU4-2	16
Tableau 3 : Répartition des 61 variants selon la classification ClinVar	30
Tableau 4 : $\Delta\Delta G$ RNAfold par groupe	32
Tableau 5 : $\Delta\Delta G$ RNAcofold par groupe	34
Tableau 6 : RNAdistance par groupe.....	35
Tableau 7 : Médiane [intervalle interquartile] des trois métriques par région fonctionnelle .	36
Tableau 8 : Comparaison Patho/LP vs VUS tests de Mann-Whitney U et tailles d'effet.	38
Tableau 9 : Répartition des variants selon la région critique (T-loop + Stem III)	39
Tableau 10 : Répartition des variants par niveau d'impact structural et par classification.	40

Table des matières

Liste des Abréviations.....	i
Liste des figures	ii
Liste des tableaux.....	iii
Introduction Générale.....	1
CHAPITRE I : Revue de la littérature	3
1.1 Troubles neurodéveloppementaux.....	4
1.2 L'épissage des pré-ARNm et le spliceosome	8
1.3 Le gène RNU4-2 et l'ARN U4.....	12
1.4 Le syndrome ReNU.....	14
CHAPITRE II : Matériel et méthodes.....	19
2.1 Données utilisées	19
2.2 Générations des séquences mutantes.....	21
2.3 Prédiction de la structure secondaire par RNAfold.....	22
2.4 Analyse de l'interaction U4/U6 par RNAcifold.....	24
2.5 Variations d'énergie libre de repliement ($\Delta\Delta G$ structure)	25
2.6 Calcul de la distance structurale (RNAdistance).....	26
2.7 Analyses statistiques.....	27
2.8 Stratégie d'analyse comparative VUS / variants pathogènes.....	28
CHAPITRE III : Résultats	30
3.1 Description des variants étudiés.....	30
3.2 Impact des variants sur la structure secondaire de l'ARN U4 (RNAfold).....	31
3.3 Impact des variants sur l'interaction U4/U6 (RNAcifold).....	33
3.4 Analyse de la distance structurale (RNAdistance).....	35
3.5 Variations par région fonctionnelle.....	35
3.6 Comparaison globale Patho/LP – VUS.....	37
3.7 Enrichissement des variants pathogènes dans la région critique	39
3.8 Corrélations entre les trois métriques	39
3.9 Classification de l'impact structural (notation dot-bracket)	40
CHAPITRE IV : Discussion.....	41
Conclusion	45
Annexes.....	46
Références bibliographiques.....	54

Introduction Générale

Les troubles du développement neurologique (TDN) forment une catégorie variée d'affections marquées par des modifications dans le développement du système nerveux central, entraînant des déficits cognitifs, moteurs, comportementaux ou adaptatifs selon des niveaux divers. Ces troubles représentent une des principales sources de handicap chez les enfants et sont souvent liés à des origines génétiques complexe, impliquant des variations qui affectent des gènes critiques pour le développement cérébral. Les avancées récentes dans les technologies de séquençage à haute capacité ont permis d'identifier un grand nombre de gènes impliqués dans ces affections, soulignant l'importance non seulement des régions codantes du génome, mais également des zones non codantes qui jouent un rôle clé dans la régulation de l'expression génique et des mécanismes cellulaires.

Parmi les éléments non codants ayant un impact sur les maladies humaines se trouvent les petits ARN nucléaires (snARN), qui participent à l'épissage des ARN pré-messagers à l'intérieur du spliceosome. Ce complexe ribonucléoprotéique garantit l'élimination précise des introns ainsi que la jonction appropriée des exons, étape essentielle pour produire des ARN messagers fonctionnels. Tout dysfonctionnement de ce processus peut induire des anomalies d'expression génique à grande échelle et favoriser l'émergence de diverses pathologies génétiques.

Le gène *RNU4-2* code pour l'ARN nucléaire U4, l'un des composants clés du spliceosome majeur. L'ARN U4 établit un duplex avec l'ARN U6 et joue un rôle dans la régulation de l'activation catalytique du spliceosome. Cette interaction est vitale pour assurer l'exactitude de l'épissage et éviter une activation prématurée du complexe catalytique. À la différence des gènes codant des protéines, la fonction biologique de *RNU4-2* est principalement déterminée par la structure secondaire et tridimensionnelle de son transcrit. Par conséquent, une simple substitution d'un nucléotide peut altérer le repliement de l'ARN, perturber ses interactions moléculaires et compromettre son rôle dans le processus d'épissage.

En 2024, des recherches génétiques ont mis en évidence des variants pathogènes récurrents de *RNU4-2* comme étant à l'origine d'un nouveau syndrome neurodéveloppemental, désormais connu sous le nom de syndrome neurodéveloppemental associé à *RNU4-2* (syndrome ReNU).

Cette condition se caractérise principalement par un retard général du développement, une déficience intellectuelle de gravité variable, des troubles du langage, une hypotonie et diverses anomalies neurologiques. La découverte de ce syndrome a mis en lumière l'importance essentielle des ARN non codants dans le développement neurologique humain et a souligné la nécessité de l'intégrité structurale de l'ARN U4 pour le bon fonctionnement du spliceosome.

L'interprétation des variants identifiés dans *RNU4-2* représente cependant un défi majeur. En effet, les approches classiques de prédiction de l'impact fonctionnel, développées principalement pour les protéines, sont peu adaptées aux ARN non codants. Dans ce contexte, les analyses *in silico* constituent des outils particulièrement précieux pour évaluer les conséquences potentielles des variants sur la stabilité thermodynamique, la structure secondaire et les interactions moléculaires de l'ARN. Des logiciels spécialisés tels que RNAfold permettent notamment de prédire les modifications de repliement induites par les substitutions nucléotidiques et d'estimer leurs effets sur la stabilité de la molécule à travers les variations d'énergie libre minimale.

L'analyse structurale computationnelle offre ainsi une approche pertinente pour explorer les mécanismes moléculaires sous-jacents à la pathogénicité des variants de *RNU4-2*. En identifiant les altérations susceptibles d'affecter les domaines fonctionnels de l'ARN U4, notamment les régions impliquées dans la liaison aux protéines du spliceosome ou dans la formation du duplex U4/U6, il devient possible de mieux comprendre les relations entre génotype et phénotype dans le syndrome ReNU.

Dans cette perspective, le présent travail vise à réaliser une analyse *in silico* des variants du gène *RNU4-2* associés au syndrome neurodéveloppemental ReNU afin d'évaluer leur impact potentiel sur la structure et la stabilité de l'ARN U4. Cette étude contribue à une meilleure compréhension des mécanismes moléculaires impliqués dans cette pathologie récemment décrite et met en évidence l'intérêt des approches bioinformatiques dans l'interprétation fonctionnelle des variants affectant les ARN non codants.

Chapitre I : Revue de la littérature

1.1 Troubles neurodéveloppementaux

1.1.1 Définition

Les troubles neurodéveloppementaux (TND) regroupent un ensemble de conditions qui apparaissent pendant le développement, généralement dès la petite enfance. Ils se traduisent par des difficultés durables dans le fonctionnement personnel, social, scolaire ou professionnel (Cernigliaro et al., 2024 ; Guerrera et al., 2024). Ces troubles comprennent notamment la déficience intellectuelle, les troubles du spectre de l'autisme (TSA), le trouble déficitaire de l'attention avec ou sans hyperactivité (TDAH), les troubles spécifiques des apprentissages, les troubles de la communication ainsi que les troubles moteurs. Bien que variés, les TND partagent souvent des bases génétiques, neurobiologiques et développementales communes.

1.1.2 Caractéristiques générales

Les TND se manifestent par des difficultés précoces touchant plusieurs domaines du développement, notamment la communication, les interactions sociales, l'apprentissage, les fonctions exécutives, le comportement adaptatif ainsi que les habiletés motrices (Valenzuela et al., 2024a) Leur évolution est généralement chronique, avec des symptômes qui peuvent persister durant l'adolescence et à l'âge adulte. Cependant, l'expression clinique des TND peut varier au cours du temps selon la sévérité du trouble, l'environnement, les prises en charge éducatives et thérapeutiques, ainsi que les capacités d'adaptation de l'individu (Figure 1) (Barbour et al., 2024).

Tableau 1. Principales caractéristiques cliniques des individus présentant un variant de *RNU4-2*

Catégorie Clinique	Caractéristiques
Anomalie neurodéveloppementale	Retard global du développement, déficience intellectuelle, hypotonie, crises épileptiques, troubles autistiques

Anomalies cérébrales (IRM)	Anomalie du corps calleux, ventriculomégalie
Croissance	Retard de croissance intra-utérin, petite taille , faible poids
Crâne	Micro-céphalie primaire ou progressive
Dysmorphie cranio-facial	Épicanthus, paupières en capuchon, joues pleines, philtrum en tente, lèvre inférieure éversée, palais étroit
Digestif	Difficultés alimentaires, reflux, constipation
Ophthalmologique	Strabisme, nystagmus
Musculo-squelettique	Scoliose, fractures, ostéoporose

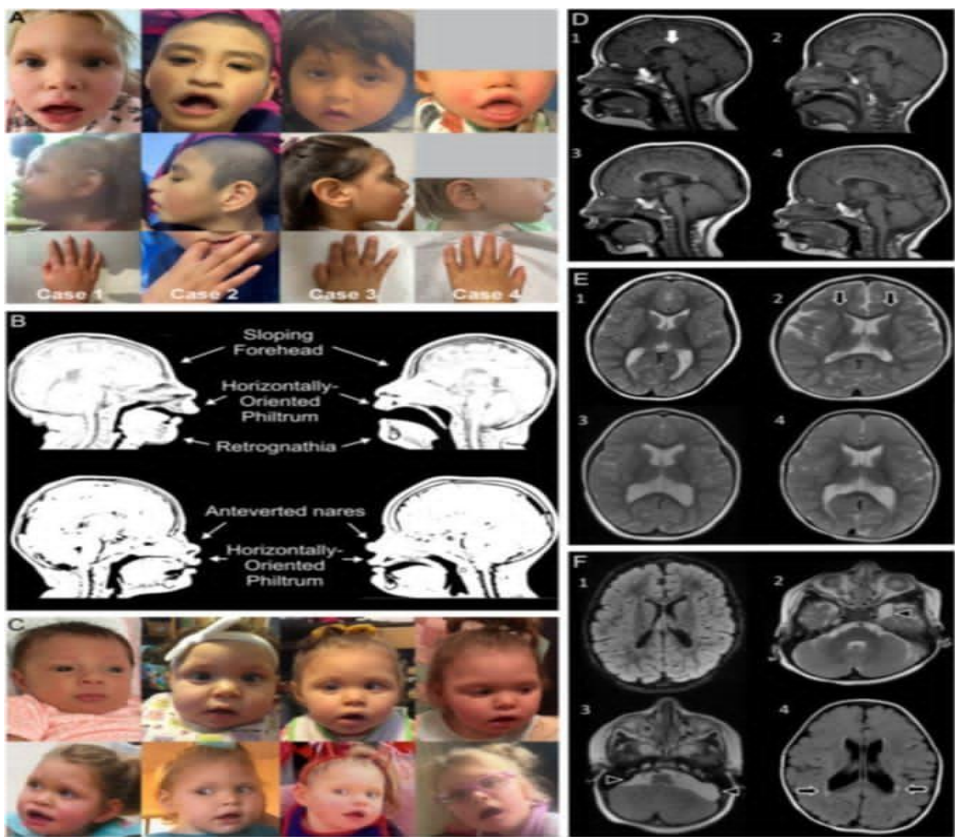


Figure 1 : Manifestations cliniques, dysmorphie cranio-faciale et anomalies cérébrales observées chez des patients atteints d’un syndrome neurodéveloppemental (TND) d’origine génétique (Barbour et al., 2024).

1.1.3 Épidémiologie des troubles neurodéveloppementaux

Les TND comptent parmi les troubles chroniques les plus fréquents de l'enfance à l'échelle mondiale. Une méta-analyse portant sur les pays à revenu faible et intermédiaire estime leur prévalence globale à environ 10 %, avec des variations importantes selon le contexte géographique et socio-économique. Parmi les sous-types les mieux documentés figurent la déficience intellectuelle (12 %), le trouble déficitaire de l'attention avec ou sans hyperactivité (TDAH, 11 %) et le trouble du spectre de l'autisme (TSA, 10 %). Ces chiffres globaux masquent toutefois des disparités considérables : les taux les plus élevés sont observés en Afrique subsaharienne (14 %) et dans les pays à faible revenu (13 %), contre 7 % en Asie.

Dans la région Moyen-Orient et Afrique du Nord (MENA), les données épidémiologiques disponibles restent hétérogènes, en partie du fait de l'insuffisance des infrastructures diagnostiques et du manque d'études comparables entre pays. Les estimations existantes révèlent néanmoins une charge substantielle. Pour le TSA, une méta-analyse récente rapporte une prévalence estimée à 0,14 % dans la région, avec une variabilité frappante entre pays de 0,01 % à Oman à 6,50 % en Irak (**Akomolafe et al., 2025**). Pour le TDAH, une méta-analyse publiée dans le BMJ Open estime la prévalence à 10,3 % dans la région MENA, avec des taux allant de 1,3 % au Yémen à 22,2 % en Iran chez les enfants et adolescents (**Al-Wardat et al., 2024**). De tels écarts ne reflètent pas nécessairement des différences biologiques réelles entre populations : ils traduisent avant tout des inégalités d'accès au diagnostic, des niveaux variables de formation des professionnels de santé et, dans certains contextes, une stigmatisation persistante qui freine le recours aux soins. Le rapport UNICEF (2023) rappelle à cet égard que la région MENA compte près de 21 millions d'enfants vivant avec un handicap, soit environ un enfant sur sept, un chiffre qui illustre l'ampleur des besoins non couverts.

En Algérie, les données épidémiologiques sur les TND demeurent limitées. La première cohorte clinique consacrée au TSA, rapportée par (**Loumi & Andres, 2025**) a porté sur 186 cas diagnostiqués entre 2013 et 2021. Les profils cliniques observés sont globalement concordants avec les données internationales, avec une prédominance masculine marquée, un âge au diagnostic tardif et une fréquence élevée de comorbidités associées à des facteurs périnataux défavorables. En l'absence de données populationnelles nationales, les estimations du Global Burden of Disease 2019 indiquent, pour la région Afrique du Nord et Moyen-Orient, un taux de prévalence standardisé pour le TSA de 304,4 pour 100000 habitants (IC 95% : 251,2-366,1) avec un ratio hommes/femmes de 2,9 (**Ebrahimi Meimand et al., 2023**).

Ces données épidémiologiques ont des implications directes sur le plan de la santé publique. Les TND génèrent des besoins en soins complexes et prolongés, mobilisant simultanément la neurologie pédiatrique, la psychiatrie, l'orthophonie, la psychomotricité et l'éducation spécialisée. Leur retentissement dépasse la sphère individuelle : les familles supportent une charge psychologique, professionnelle et financière considérable, tandis que les systèmes de santé doivent absorber des coûts importants sur l'ensemble de la trajectoire de vie des patients. Plusieurs facteurs de risque sont bien établis : sexe masculin, faible poids de naissance, prématurité, âge maternel avancé, complications obstétricales, bas niveau d'éducation maternelle et conditions socio-économiques défavorables ce qui laisse entrevoir des marges d'action préventive réelles, à condition que les interventions ciblent les déterminants modifiables dès la période périnatale (Almahmoud & Abushaikha, 2023; Mehari & Zeleke, 2026).

1.1.4 Étiologie génétique et épigénétique

L'étiologie des TND est multifactorielle. Les études des jumeaux et les analyses de ségrégation familiale ont établi une composante génétique significative pour la majorité de ces troubles, avec des estimations d'héritabilité variant de 50 à 90% selon le trouble considéré (Tick et al., 2016). Cette composante recouvre des mécanismes variés : Mutations ponctuelles, variations du nombre de copies (CNV), mutations de novo et variants rares à effet intermédiaire.

Parmi les variants les mieux caractérisés figurent les mutations de novo absentes chez les deux parents et apparues spontanément chez l'enfant atteint. Ces mutations touchent fréquemment des gènes impliqués dans le développement cérébral, comme *NRXN1*, *NLGN3* ou *SHANK3*, associés à des formes syndromiques de TSA. Leur identification a été rendue possible par l'essor du séquençage de l'exome et, plus récemment, du génome entier. Parallèlement, les CNV gains ou pertes de segments chromosomiques perturbent la dosimétrie de gènes entiers et constituent une autre source majeure de vulnérabilité génétique. Les variants rares, définis par une fréquence inférieure à 1 % dans la population générale, représentent quant à eux une catégorie hétérogène dont l'effet pathogène est souvent difficile à établir en dehors d'études fonctionnelles dédiées (Redon et al., 2006).

Les mécanismes épigénétiques jouent un rôle reconnu dans la physiopathologie des TND. La méthylation de l'ADN, les modifications post-traductionnelles des histones et la régulation par les ARN non codants modulent l'expression des gènes de façon réversible, sans altérer la

séquence sous-jacente. Des anomalies de méthylation au niveau des régions CpG, en particulier dans les gènes synaptiques, ont été documentées chez des individus présentant un TSA ou une déficience intellectuelle. Le syndrome de Rett, causé par des mutations du gène *MECP2* qui code un lecteur de méthylation, illustre de manière paradigmatique le rôle des perturbations épigénétiques dans les TND (**Barros & Offenbacher, 2009**).

Enfin, les facteurs environnementaux agissent en interaction avec ce substrat génétique et épigénétique. L'exposition prénatale à des toxiques (alcool, tabac, métaux lourds, bisphénol A), les infections maternelles, le stress gestatif ou la malnutrition peuvent modifier durablement les profils épigénétiques et influencer l'expression de gènes critiques pour le neurodéveloppement, particulièrement durant la neurogenèse et la synaptogenèse (**Kundakovic & Champagne, 2015**). Cette interaction gène–environnement est indispensable pour comprendre pourquoi deux individus portant une même mutation peuvent présenter des tableaux cliniques sensiblement différents (**Burgunder, 2023**).

1.1.5 Mécanismes moléculaires impliqués

Les TND sont associés à des altérations génétiques qui affectent des mécanismes biologiques clés du développement et du fonctionnement du cerveau. Trois grands niveaux mécanistiques sont aujourd'hui bien documentés : la fonction synaptique, la neurogenèse et le contrôle de l'expression génique.

Les dysfonctionnements synaptiques occupent une place centrale dans la physiopathologie de nombreux TND, en particulier le TSA. Des mutations affectant les protéines synaptiques, comme les neurexines (NRXN) et les protéines de la famille SHANK, altèrent la structure, la plasticité et la transmission des synapses excitatrices, compromettant la communication entre neurones (**Michetti et al., 2022**). Ces anomalies ne se limitent pas à des défauts locaux : elles perturbent l'équilibre excitation/inhibition à l'échelle des circuits, avec des conséquences sur le traitement de l'information sensorielle, sociale et cognitive. D'autre part, des perturbations dans la formation des neurones, y compris des problèmes de prolifération, de migration et de maturation neuronale, ont aussi été notées. Ces mécanismes sont cruciaux pour une construction adéquate des circuits neuronaux durant le développement du cerveau, et leurs dérèglements peuvent engendrer des déficits structurels et fonctionnels persistants (**Washbourne, 2015**).

Enfin, la régulation de l'expression génique constitue le troisième mécanisme notamment via le contrôle de la transcription et l'épissage de pré-ARNm. Des mutations touchant des facteurs de transcription neuronaux ou des composants de la machinerie d'épissage comme les snRNA du spliceosome altèrent simultanément l'expression de nombreux gènes cibles, générant des phénotypes complexes et pléiotropiques (Plowright et al., 2024).

1.2 L'épissage des pré-ARNm et le spliceosome

1.2.1 Les ARN non codants

longtemps qualifiés d'« ADN poubelle » et considérés comme dépourvus de fonction, les séquences génomiques non codantes ont fait l'objet d'une réévaluation majeure grâce aux progrès du séquençage à haut débit et des projets d'annotation à grande échelle, notamment ENCODE Project Consortium une grande partie du génome humain est transcrite, et une fraction importante de ces transcrits exerce des fonctions biologiques sans être traduite en protéines (ENCODE Project Consortium, 2012).

Les ARN non codants (ARNnc) forment une classe fonctionnelle subdivisés ARNnc courts (<200nt), des ARNnc longs (lncRNA, >200nt). Parmi les ARNnc courts fonctionnellement caractérisés figurent les microARN (miRNA), impliqués dans la régulation post-transcriptionnelle, les ARN de petit nucléole (snoRNA), qui guident les modifications chimiques des ARN ribosomiques, et les petits ARN nucléaires (snRNA), composants essentiels de la machinerie d'épissage. Ces molécules n'agissent pas isolément : elles s'associent à des protéines spécifiques pour former des complexes ribonucléoprotéiques (RNP) dont la structure tridimensionnelle conditionne la fonction (Matera et al., 2007; Mattick et al., 2023; Zampetaki et al., 2018).

Contrairement aux ARNm, l'activité biologique des ARNnc repose sur leur architecture tridimensionnelle plutôt que sur une information codante. Une substitution nucléotidique qui modifie le repliement d'un ARNnc peut suffire à perturber sa fonction ce qui rend leur analyse structurale particulièrement informative dans le contexte de l'interprétation des variants génétiques (Miao & Westhof, 2017; Zampetaki et al., 2018)

.1.2.2 Les snRNA et leur rôle dans l'épissage

Le spliceosome est un complexe ribonucléoprotéique conservé, présent dans le noyau des cellules eucaryotes, dont la fonction est d'exciser les introns des pré-ARNm et de lier les exons flanquants pour produire un ARNm mature. Il existe sous deux formes : le spliceosome majeur, responsable de l'épissage de plus de 99% des introns, et le spliceosome mineur, dédié à une classe restreinte d'introns à séquences consensus atypiques. Les snRNA U1, U2, U4, U5 et U6 constituent le cœur fonctionnel du spliceosome majeur, dont l'activité catalytique repose principalement sur les réarrangements impliquant U2 et U6. Leur taille varie de 100 à 200 nucléotides, et chacun adopte une structure secondaire conservée, maintenue par des appariements intramoléculaires précis. Contrairement aux ARNm, les snRNA ne subissent pas de traduction : ils fonctionnent directement à l'état d'ARN, en s'associant à des protéines spécifiques les protéines Sm pour U1, U2, U4 et U5, et les protéines Lsm pour U6 pour former des particules snRNP (small nuclear ribonucleoprotein particles).

Chaque snRNA occupe une fonction distincte et non redondante dans le cycle d'épissage. U1 reconnaît le site donneur 5' de l'intron par complémentarité de séquence. U2 s'associe au point de branchement en amont du site accepteur 3'. U4 et U6 forment un duplex stable dans lequel U4 maintient U6 dans une conformation inactive. U5, enfin, contribue à positionner les exons adjacents pour la réaction de ligation (**Figure 2**) (Cooper, 2000)

Plusieurs régions fonctionnelles de U2, U4 et U6 sont orthologues entre *Saccharomyces cerevisiae* et l'humain témoignant du rôle important que jouent ces ARN dans le spliceosome. Toute substitution dans ces régions risque de compromettre l'assemblage l'activation du spliceosome ce qui confère aux variants du gène *RNU4-2* un potentiel pathogène direct lorsqu'ils touchent des positions structurellement conservées (Will & Lührmann, 2011 ;Chen et al., 2024a).

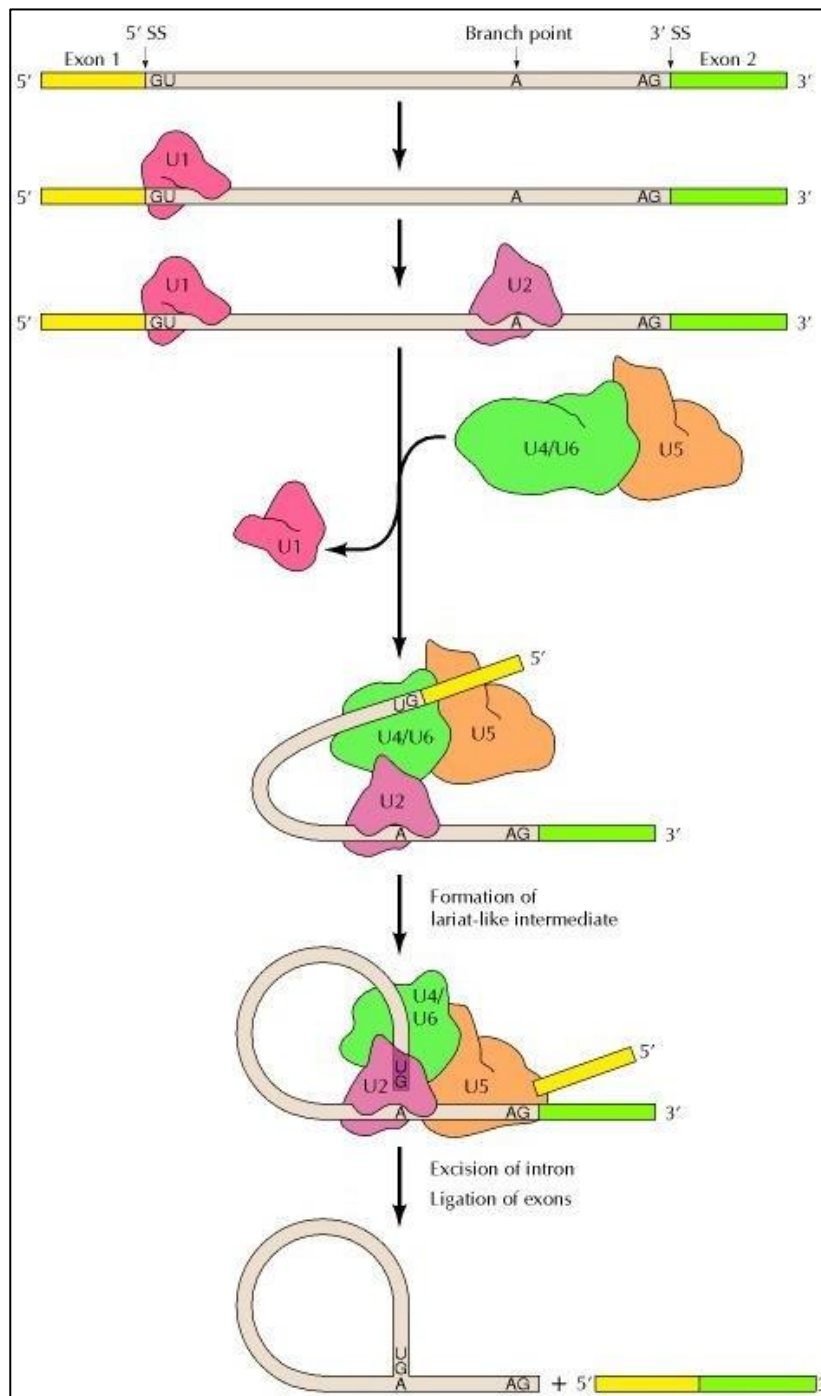


Figure 2 : Représentation schématique du mécanisme général de l'épissage constitutif des pré-ARNm (retrait de l'intron sous forme de lariat et jonction des exons (**Cooper, 2000**)).

1.2.3 Organisation et cycle du spliceosome majeur :

Le spliceosome majeur est l'un des complexes macromoléculaires les plus grands et les plus dynamiques de la cellule eucaryote. Il comprend les cinq particules snRNP ainsi qu'une centaine de protéines auxiliaires impliquées dans son assemblage, son activation et son recyclage (**Will & Lührmann, 2011**). Son assemblage sur le pré-ARNm est séquentiel et ordonné, procédant par étapes distinctes que l'on désigne classiquement par les complexes E, A, B, B* et C (**Figure 3**) (**Nguyen et al., 2016; Sander et al., 2006**).

la reconnaissance du site donneur 5' par U1 snRNP initie l'assemblage (complexe E), suivie du recrutement de U2 snRNP au niveau du point de branchement intronique (complexe A). Le tri-snRNP U4/U6·U5 rejoint ensuite l'ensemble pour former le complexe B, encore catalytiquement inactif. Transition vers le complexe B* implique un remaniement structural majeur : le duplex U4/U6 est dissocié, libéré de U4 et U6 adopte une conformation active et établit de nouvelles interactions avec U2 et le pré-ARNm. U4, devenu dispensable à ce stade, se dissocie du complexe (**Wahl et al., 2009**).

Le complexe C ainsi formé catalyse les deux réactions de transestérification successives qui constituent l'épissage : la première génère l'intermédiaire en attaquant le site donneur 5', la seconde assure la ligation des deux exons et libère l'intron. Après catalyse, le spliceosome se dissocie et ses composants dont U4 sont recyclés pour un nouveau cycle d'épissage.

U4 y joue un rôle de régulateur négatif de l'activation catalytique, il séquestre U6 dans une conformation inactive et il prévient toute activation prématurée du centre catalytique. Ce mécanisme de contrôle est essentiel à la fidélité de l'épissage, et sa perturbation notamment par des variants pathogènes de *RNU4-2* peut conduire à des erreurs de reconnaissance des sites d'épissage et à des anomalies transcriptomiques globales (**Greene et al., 2024; Nielsen & Staley, 2012**).

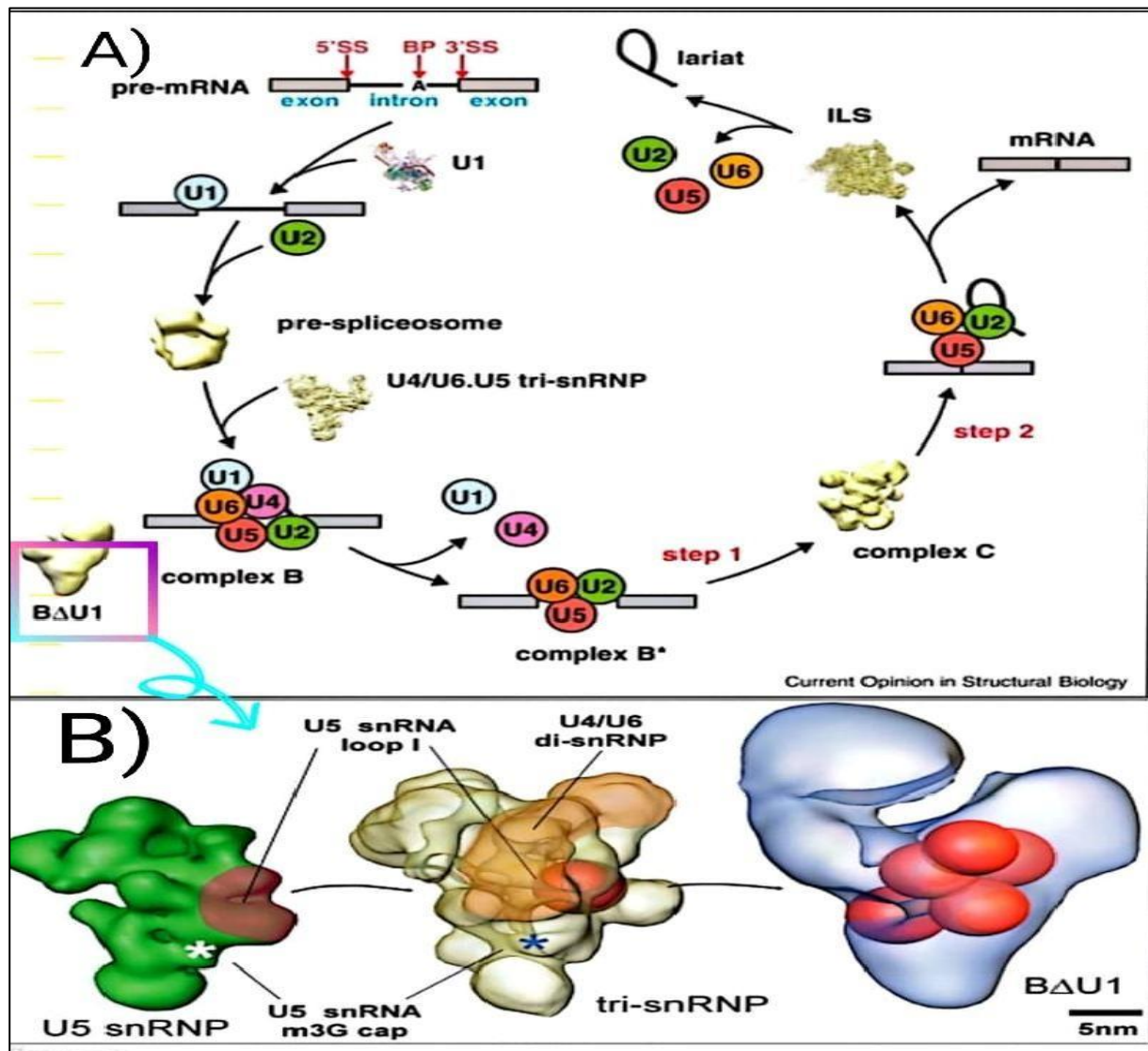


Figure 3 : Cycle d'assemblage, d'activation catalytique et de désassemblage du spliceosome majeur au cours du processus d'épissage (Nguyen et al., 2016; Sander et al., 2006).

1.3 Le gène RNU4-2 et l'ARN U4

1.3.1 Caractéristiques génomiques de RNU4-2

Le gène *RNU4-2* code le snRNA U4 du spliceosome majeur. Localisé sur le chromosome 12(12q24.31), il occupe les positions chr12 : 120 291 759 – 120 291 903 sur le brin antisens, pour une longueur génomique 145 nucléotides. mono exonique dépourvu de cadre de lecture ouvert, ce gène ne fait l'objet d'aucun épissage post-transcriptionnel. La longueur de son transcrit l'ARN U4 snRNA est donc quasiment identique à sa longueur génomique, soit environ 145 nucléotides. Cette compacité est typique des snRNA, dont la séquence est intégralement

fonctionnelle dès la transcription. Le gène est également référencé sous plusieurs synonymes : *RNU4-1B*, *RNU4C*, *U4A*, *U4B* et *U4C*.

1.3.2 Structure secondaire de l'ARN U4

L'ARN U4 présente une structure secondaire organisée en plusieurs motifs tige boucle (*stem-loops*) conservés, comprenant la tige-boucle 5', le k-turn (*kink-turn*), la T-loop et la tige III (**Figure 4**) (Nava et al., 2025; Will & Lührmann, 2011). Le k-turn est reconnu par la protéine 15.5K/NHPX, dont la liaison constitue l'une des premières étapes de l'assemblage de la particule U4 snRNP. La T-loop et la tige III sont quant à elles impliquées dans la formation du duplex U4/U6 et représentent le principal foyer mutationnel du syndrome ReNU ce qui souligne l'importance fonctionnelle de leur intégrité structurale. Toute substitution nucléotidique dans ces régions est susceptible de modifier le repliement de la molécule, de déstabiliser ses interactions avec U6 ou avec les protéines spliceosomales, et d'altérer en conséquence le bon déroulement de l'épissage (**Figure 5**) (Nottrott et al., 1999; Vidovic et al., 2000).

La T-loop et la tige III, impliquées dans la formation du duplex U4/U6, concentrent la majorité des variants pathogènes du syndrome ReNU. Au sein du spliceosome, U4 agit comme régulateur négatif de l'activation catalytique : en maintenant U6 dans une conformation inactive via ce duplex, il prévient toute activation prématurée du centre catalytique. Les variants pathogènes de *RNU4-2* perturbent précisément cette régulation, induisant des erreurs de sélection des sites d'épissage et des anomalies transcriptomiques délétères dans les neurones en développement (Chen et al., 2024a).

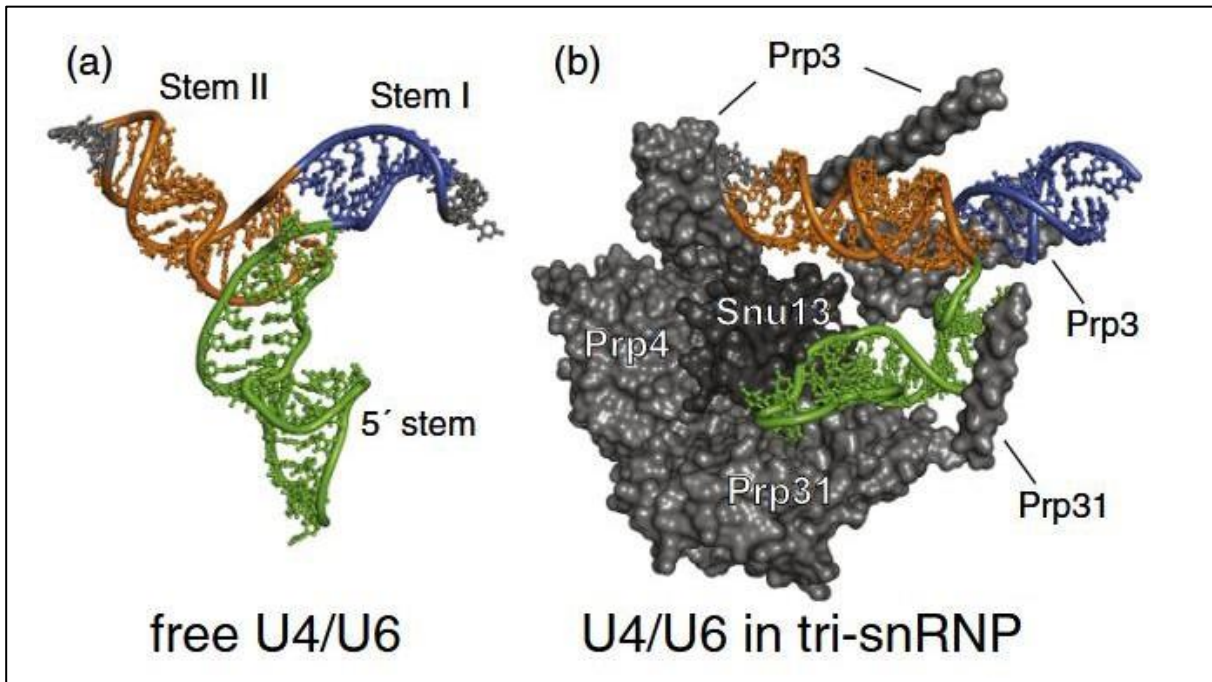


Figure 4 : Structure tridimensionnelle du duplex d'ARN snU4/U6 humain (a) à l'état libre.

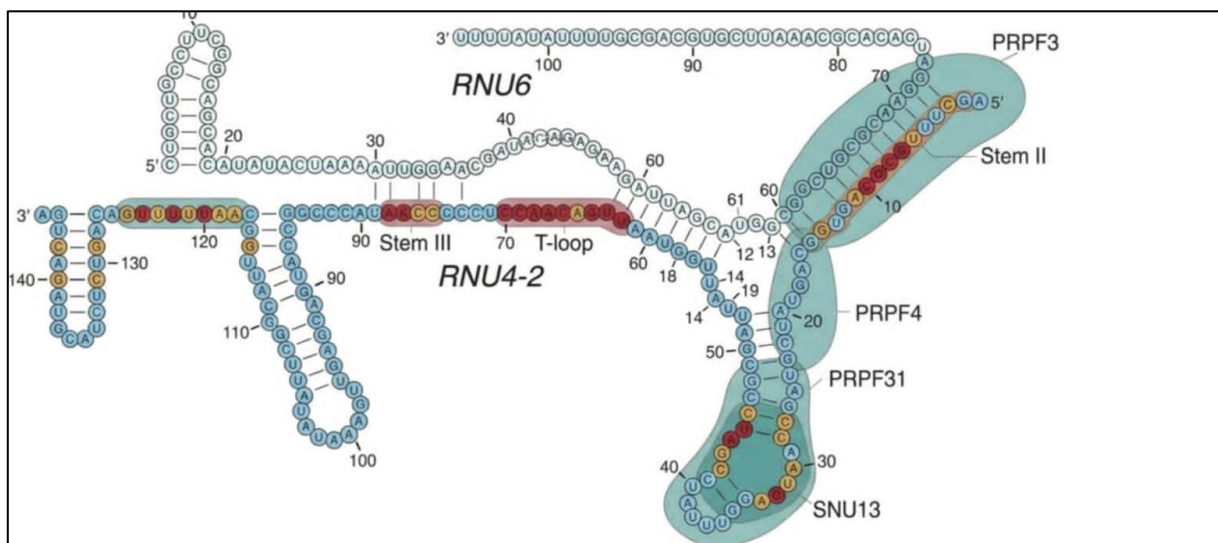


Figure 5 : Structure secondaire du duplex U4/U6 avec numérotation des ribonucléotides de *RNU4-2*. Structure secondaire du duplex U4/U6 au sein du tri-snRNP U4/U6·U5. L'ARN U4 forme trois motifs tige-boucle intramoléculaires (nucléotides 20–52, 85–117 et 127–144) et interagit avec U6 via des appariements intermoléculaires dans la tige I (nt 56–61), la tige II (nt 1–16), la tige III (nt 73–79) et le quasi-pseudo-nœud (nt 62 et 64) Adapté de (Greene et al., 2024).

1.4 Le syndrome ReNU

1.4.1 Définition

Le syndrome ReNU (*RNU4-2-related Neurodevelopmental syndrome*) a été caractérisé pour la première fois en 2024 comme une cause monogénique distincte de TND sévères, liée à des variants pathogènes du gène *RNU4-2* (**Chen et al., 2024a; Greene et al., 2024**). La convergence de plusieurs cohortes indépendantes vers la même entité clinique a conduit à reconnaître RNU4-2 comme contributeur monogénique fréquent des TND (**Burns & Radford, 2024; Chen et al., 2024a**).

Ce syndrome se distingue des autres causes de NDD par deux caractéristiques. D'une part, il est causé par des variants dans un gène d'ARN non codant, invisible aux pipelines d'analyse standards centrés sur les séquences codantes. D'autre part, la quasi-totalité des variants pathogènes décrits sont apparus de novo, sans transmission parentale identifiable.

1.4.2 Spectre mutationnel de RNU4-2

La forme classique du syndrome ReNU est causée par des variants hétérozygotes de novo localisés dans une région critique de 18 nucléotides englobant la T-loop et la tige III de l'ARN U4 deux éléments structuraux essentiels à l'interaction U4/U6 et à l'assemblage du spliceosome (**Chen et al., 2024a**). Cette région constitue le principal foyer mutationnel (*hotspot*) du gène, et le variant le plus fréquemment rapporté est l'insertion n.64_65insT, retrouvée chez une majorité des patients décrits dans les premières cohortes (**Chen et al., 2024a; Jonghe et al., 2025; Okamoto et al., 2025**).

Un second hotspot a été identifié dans la tige-boucle 5' du gène, associé à un phénotype ReNU comparable mais transmis selon un mode souvent hérité plutôt que de novo. Par ailleurs, des variants bialléliques touchant des régions telles que la tige II, le k-turn ou le site Sm ont été décrits comme responsables d'une forme récessive distincte, caractérisée par des anomalies spécifiques de la substance blanche à l'IRM cérébrale. Ces données suggèrent que l'effet pathogène des variants de *RNU4-2* dépend à la fois de leur localisation dans la structure secondaire de l'ARN U4 et du mode d'hérédité.

Les variants pathogènes altèrent la sélection du site donneur 5' d'épissage, entraînant des anomalies transcriptomiques globales particulièrement délétères dans les cellules neuronales en

développement (Chen et al., 2024a; Hayashi et al., 2026). Une cartographie fonctionnelle par saturation genome editing suggère que les positions les plus intolérantes aux substitutions correspondent précisément aux régions les plus conservées structuralement de l'ARN U4, résultats qui nécessitent une confirmation par des cohortes indépendantes (Hayashi et al., 2026).

Tableau 2 : Régions mutationnelles de RNU4-2, mode d'hérédité et sévérité clinique associée

Région de <i>RNU4-2</i>	Mode d'hérédité	Sévérité clinique typique
T-loop	Dominant de novo	Retard sévère, langage absent
Tige III	Dominant de novo	Formes parfois plus modérées
Tige 5'	Souvent hérité	Phénotype ReNU homogène
Tige II / k-turn / site Sm	Récessif biallélique	NDD distinct, anomalies de substance blanche

1.4.3 Présentation clinique

Le tableau clinique du syndrome ReNU est dominé par un retard neuro-développemental global sévère. La déficience intellectuelle est constante, généralement modérée à sévère, et s'accompagne d'une absence ou d'une réduction majeure du langage expressif. Un retard moteur, une hypotonie axiale et des crises épileptiques sont fréquemment rapportés, ainsi qu'une petite taille et une microcéphalie postnatale progressive (Jonghe et al., 2025; Okamoto et al., 2025). L'imagerie cérébrale révèle des anomalies structurales variées, incluant une ventriculomégalie, une hypoplasie du corps calleux et une réduction de la substance blanche.

Sur le plan dysmorphologique, les patients présentent un faciès reconnaissable qui facilite l'orientation diagnostique clinique : paupières supérieures en capuchon, épicanthus, joues pleines, philtrum en tente, bouche entrouverte avec lèvre inférieure éversée, hypersialorrhée, ainsi que des anomalies nasales et auriculaires (Rosenblum et al., 2025; Valenzuela et al., 2024b). Ces traits, bien que non spécifiques pris isolément, forment un gestalt clinique suffisamment cohérent pour orienter vers un séquençage ciblé de RNU4-2.

Des complications musculosquelettiques ont été documentées, notamment des fractures récurrentes et une ostéoporose sévère chez certains patients adultes, suggérant un dysfonctionnement systémique lié aux anomalies d'épissage généralisée (Peñafiel-Sam et al., 2025). Enfin, des anomalies prénatales du système nerveux central dont une microcéphalie fœtale peuvent constituer le premier signe d'appel, ouvrant la voie à un diagnostic anténatal dans les familles à risque (**Chen et al., 2024a**).

Objectif :

Ce travail vise à analyser *in silico* l'impact des variants du gène RNU4-2 sur la structure et la stabilité de l'ARN U4, afin de mieux comprendre les mécanismes moléculaires associés au syndrome ReNU.

Pour ce faire, trois métriques complémentaires ont été calculées pour chacun des 61 variants issus de ClinVar : la variation d'énergie libre de repliement ($\Delta\Delta G$ RNAfold), la variation d'énergie libre d'interaction U4/U6 ($\Delta\Delta G$ RNACofold) et la distance structurale (RNAdistance). Ces paramètres ont ensuite été comparés entre les variants pathogènes et les variants de signification incertaine (VUS), en tenant compte de leur localisation dans les domaines fonctionnels de l'ARN U4, afin d'identifier des critères discriminants et des candidats VUS prioritaires pour une validation expérimentale.

Chapitre II : Matériel et méthodes

2.1 Données utilisées

2.1.1 Sources des données

Les données biologiques utilisées dans cette étude ont été obtenues à partir de deux bases de données publiques de référence. Les séquences nucléotidiques des gènes *RNU4-2* et *RNU6-1* ont été extraites de la base de données Ensembl accessible via le navigateur correspondant (**Figure 6**). Cette plateforme fournit des annotations génomiques intégrées, régulièrement mises à jour, et constitue une source de référence pour les études de génomique fonctionnelle. Les variants génétiques associés au gène *RNU4-2* ont quant à eux été récupérés depuis la base de données ClinVar (NCBI). Cette plateforme centralise les interprétations cliniques des variants génétiques humains soumis par des laboratoires accrédités à travers le monde. La consultation de Clinvar a été effectuée en mai 2025, et l'ensemble des variants annotés pour *RNU4-2* a été extrait via l'interface de recherche par gène (terme de requête : *RNU4-2*) (**Figure 7**).

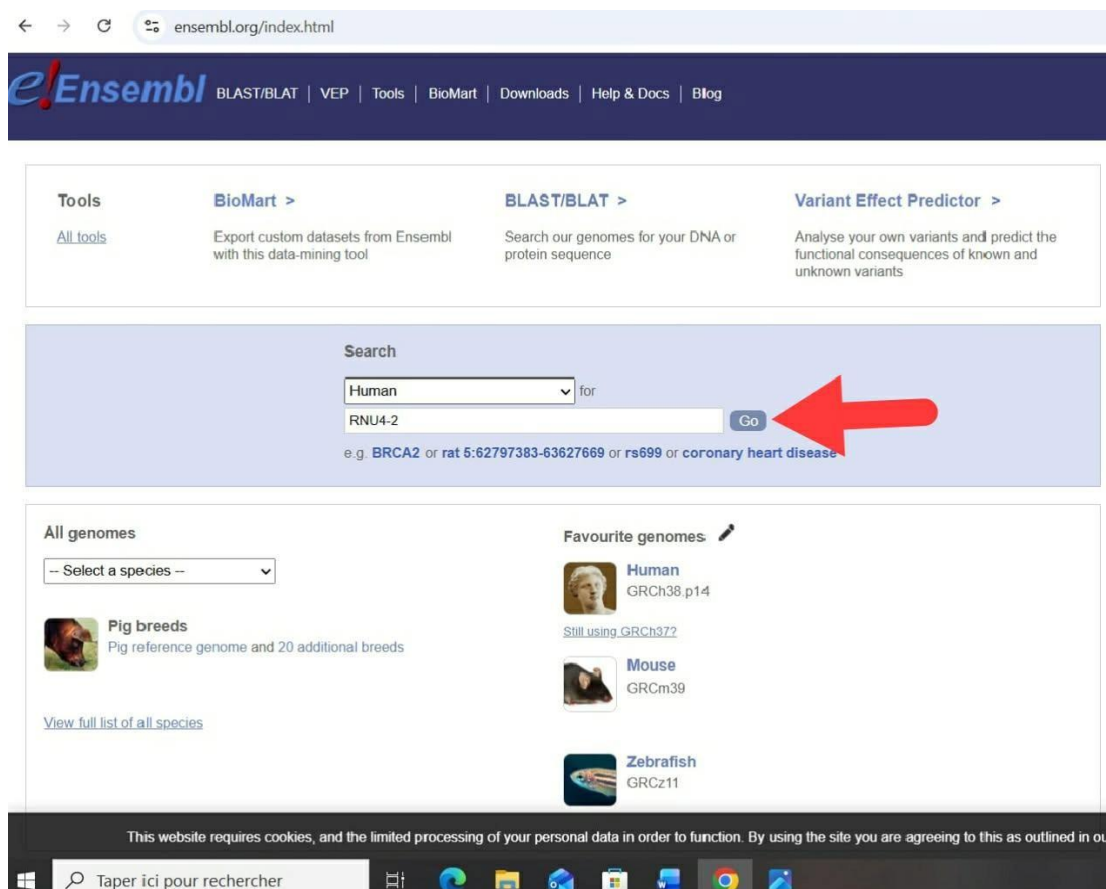


Figure 6 : Page d'accueil du navigateur génomique Ensembl (Release 115).

Search (all fields optional) [Clear all](#)

Variant [?](#)

Gene [?](#)

RNU4-2

Genomic Location [?](#)

GRCh38

More search options [v](#)

[Advanced Search](#) **Search**

Figure 7 : Interface de recherche de la base de données ClinVar (NCBI).

2.1.2 séquences nucléotidiques de référence

Deux séquences de référence ont servi de base à l'ensemble des analyses structurales menées dans cette étude. La première correspond à l'ARN snU4, produit du gène *RNU4-2* : il s'agit du transcrit RNU4-2-201 (identifiants Ensembl ENST00000365668.2 et RefSeq NR_003137.3), long de 145 nucléotides, localisé sur le brin antisens du chromosome 12 aux coordonnées GRCh38 chr12:120 291 759–120 291 903. Cette séquence constitue la référence sauvage (*wild-type*, WT) pour l'ensemble des prédictions et comparaisons structurales. La seconde séquence correspond à l'ARN snU6, produit du gène *RNU6-1* : transcrit RNU6-1-201 (ENST00000383898.1), long de 107 nucléotides, localisé sur le brin antisens du chromosome

15 aux coordonnées GRCh38 chr15:67 839 939–67 840 045. Cette séquence a été utilisée comme partenaire moléculaire dans les analyses d'hybridation intermoléculaire par RNAcofold.

La séquence WT de RNU4-2 utilisée pour l'ensemble des prédictions structurales est la suivante :

**5'AGCUUUGCGCAGUGCAGUAUCGUAGCCAAUGAGGUUUAUCCGAGGCGCGAUUAUUGCUAA
UUGAAAACUUUUCCTAAUACCCCGCCAUGACGACUUGAAAUAUAGUCGGCAUUGGCAAUUU
UGACAGUCUCUACGGAGACUGA-3'**

La séquence WT de RNU6-1 est la suivante :

**5'GUGCUCGCUUCGGCAGCACAUAUACUAAAAUUGGAACGAUACAGAGAAGAUUAGCAUGGC
CCCTGCGCAAGGAUGACACGCAAAUUCGUGAAGCGUCCAUAUUUUU-3'**

2.1.3 Sélection et classification des variants

Les variants inclus dans cette étude ont été sélectionnés sur la base de leur pertinence biologique et de leur validation dans des bases de données fiables et des publications scientifiques

L'inclusion des variants dans l'étude a été soumise aux critères suivants :

Inclusion : variants disposant d'une annotation HGVS en nomenclature n. (position dans la séquence transcrite) ; variants situés dans l'exon unique de RNU4-2 ; variants classés dans Clinvar comme étant Pathogenic (p), Likely pathogenic(LP), pathogenic/ Likely pathogenic(P/LP) OU Uncertain significance (VUS)

Exclusion : variants classés Benign ou likely benign ; grands remaniements chromosomiques (CNV) sans annotation. ; variants soumis en doublon avec annotation différente (une seule entrée retenue par variant).

2.2. Générations des séquences mutantes

Pour chacun des 61 variants inclus, une séquence mutante a été générée à partir de la séquence WT de *RNU4-2* en appliquant la modification nucléotidique décrite par la notion HGVS correspondante

Quatre types de variants ont été traités : **Substitutions (SNV)** : Changement d'un seul nucléotide (ex : n.140G>A). **Insertions** : Ajout de nucléotides (ex : n.45_46insT). **Délétions** : Perte de nucléotides (ex : n.127_128del). **Duplications** : Copie en double d'une séquence de nucléotides (ex : n.68dup). Seuls les variants exoniques ont été inclus, l'exon unique du gène couvrant l'intégralité de la séquence transcrite.

2.3 Prédiction de la structure secondaire par RNAfold

La prédiction des structures secondaires intramoléculaires et le calcul des paramètres thermodynamiques associés ont été réalisés à l'aide du programme web RNAfold (**Figure 8**), intégré à la suite ViennaRNA Package 2.0 (**Lorenz et al., 2011**). Ce programme est fondé sur l'algorithme de programmation dynamique de Zuker (**Zuker & Stiegler, 1981**), dont l'objectif est d'identifier la structure secondaire qui minimise l'énergie libre de Gibbs (ΔG , en kcal/mol) l'algorithme décompose la séquence en sous séquences imbriquées et évalue, pour chaque perte de positions(i,j) l'énergie de l'ensemble des conformations secondaires compatibles Tiges(stems) Boucles terminales (hairpin loops) Boucles internes (internal loops), bulges et jonctions multibranches (multiloops). Les paramètres thermodynamiques utilisés Sont issus des tables d'énergie De Turner, établies Expérimentalement Par calorimétrie (**Mathews et al., 1999; Nearest Neighbor Database, 2024**).

La conformation retenue comme prédiction principale correspond à la structure à énergie libre minimale (Minimum Free Energy, MFE), qui correspond à l'état d'équilibre thermodynamique global de la molécule. Cette approche suppose que la molécule d'ARN se replie vers son état d'équilibre thermodynamique global, ce qui constitue une approximation valable dans de nombreux contextes biologiques, bien que des effets cinétiques ou co-transcriptionnels puissent s'en écarter.

RNAfold WebServer 1 Enter Input Parameters 2 View Results

[Home] [New job] [Help]

The **RNAfold web server** will predict secondary structures of single stranded RNA or DNA sequences. Current limits are 7,500 nt for partition function calculations and 10,000 nt for minimum free energy only predictions.

Simply paste or upload your sequence below and click *Proceed*. To get more information on the meaning of the options click the symbols. You can test the server using [this sample sequence](#).

Sequence Input

Paste or type your **sequence** here: [clear]

AGCTTTGCGCAGTGGCAGTATCGTAGCCAATGAGGTTTATCCGAGGCGCGATTATTGCTA
 ATTGAAACTTTTCCCAATACCCGCCATGACGACTTGAAATATAGTCGGCATTGGCAAT
 TTTTGGACAGTCTCTACGGAGACTGA

Or upload a file in FASTA format: Choisir un fichier Aucun fichier choisi

Folding Constraints

Fold algorithms and basic options

- ☒ minimum free energy (MFE) and partition function
- ☐ minimum free energy (MFE) only
- ☐ no GU pairs at the end of helices
- ☒ avoid isolated base pairs
- ☐ assume RNA molecule to be circular
- ☐ Incorporate G-Quadruplex formation into the structure prediction algorithm

Dangling end options

Energy Parameters

Modified Bases

SHAPE reactivity data

Output options

- ☒ interactive RNA secondary structure plot
- ☒ RNA secondary structure plots with reliability annotation (Partition function folding only)
- ☒ Mountain plot

Notification via e-mail upon completion of the job (optional):

Proceed »

Institute for Theoretical Chemistry | University of Vienna | rna@tbi.univie.ac.at

Figure 8 : Interface de RNAfold.

2.3.1 Sorties obtenues (dot-bracket, MFE)

RNAfold produit deux sorties principales :

La représentation dot-bracket qui est une représentation linéaire de la structure secondaire utilisant trois types de caractères :

. (point) : nucléotide non apparié (simple brin) ; (et) : nucléotides formant une paire de bases ; La valeur de MFE (en kcal/mol) est indiquée entre parenthèses à la suite de la structure. Une valeur plus négative reflète une structure plus stable. Cette valeur peut être utilisée pour comparer des structures entre elles, ou pour évaluer l'impact de mutations sur la stabilité de repliement.

2.4 Analyse de l'interaction U4/U6 par RNAcofold

L'impact des variants sur la formation du duplex intermoléculaire U4/U6 a été évalué à l'aide du programme web RNAcofold (ViennaRNA Package 2.0) (**Lorenz et al.,2011**). RNAcofold est conçu pour la prédiction de structures hybrides entre deux molécules d'ARN. Il concatène les deux séquences en les séparant par un opérateur de coupure (&), puis applique l'algorithme de programmation dynamique de Zuker sur la séquence combinée, en traitant la jonction comme une boucle externe. Cette approche permet d'estimer simultanément les interactions intramoléculaires au sein de chaque brin et les appariements intermoléculaires entre les deux partenaires, à partir des paramètres thermodynamiques du modèle de Turner.

Pour chaque variant étudié, la séquence mutante de l'ARN U4 a été concaténée à la séquence WT de l'ARN U6 selon le format FASTA attendu par RNAcofold (U4_variant&U6_WT). Seules les séquences matures (sans modifications post-transcriptionnelles) ont été utilisées, en accord avec les recommandations d'utilisation du logiciel. Les analyses ont été conduites à 37 °C avec les paramètres par défaut, afin d'assurer la cohérence avec les conditions employées pour RNAfold. La sortie principale utilisée est l'énergie libre de liaison ($\Delta G_{binding}$, en kcal/mol), qui reflète la stabilité thermodynamique du duplex U4/U6 hybride.

RNAcofold WebServer 1 Enter Input Parameters 2 View Results

[Home] [New job] [Help]

The **RNAcofold web server** will predict secondary structures of single stranded RNA or DNA sequences upon dimer formation.

Simply paste or upload your sequences below and click **Proceed**. To get more information on the meaning of the options click the symbols. You can test the server using these [sample sequences](#).

Paste or type your **first sequence** here: [clear]
 ATTGAAACTTTTCCCAATACCCGCCATGACGACTTGAAATATAGTCGGCATTGGCAATTTTGTACAGTCTCTACGGAGACTGA **RNU4-2**

Or upload a file containing the **first sequence** in FASTA format: Aucun fichier choisi

Paste or type your **second sequence** here: [clear]
 GTGCTCGCTTCGGCAGCACATATACTAAAATTGGAACGATACAGAGAAGATTAGCATGGCCCTGCGCAAGGATGACACGCAAATTCGTGAAGCGTCCATATTTT **RNU6-1**

Or upload a file containing the **second sequence** in FASTA format: Aucun fichier choisi

Fold algorithms and basic options

- ☒ minimum free energy (MFE) and partition function (for all species)
- ☐ minimum free energy (MFE) and partition function (hetero-dimer only)
- ☐ minimum free energy (MFE) only
- ☐ no GU pairs at the end of helices
- ☒ avoid isolated base pairs

Show advanced options

Concentrations

- ☒ Use no concentration parameters
- ☐ Vary concentration (generates concentration dependency plots)
- ☐ Use user specified concentrations

Notification via e-mail upon completion of the job (optional):

Figure 9 : Interface de RNAcofold

2.5 variations d'énergie libre de repliement ($\Delta\Delta G$ structure)

L'effet de chaque variant sur la stabilité de la structure secondaire de l'ARN U4 a été quantifié par le calcul de la variation d'Energie libre de repliement ($\Delta\Delta G$ structure), définie par l'équation :

$$\Delta\Delta G_{structure} = \Delta G_{variant} - \Delta G_{WT}$$

Où $\Delta G_{variant}$ - ΔG_{WT} désignent respectivement les valeurs MFE calculées par RNAfold pour la séquence mutante et pour la séquence sauvage de référence. Une valeur de $\Delta\Delta G_{structure} > 0$ indique que la mutation déstabilise la structure secondaire de l'ARN U4 (repliement

moins favorable), tandis qu'une valeur < 0 traduit un effet stabilisant. Le modèle des plus proches voisins (*Nearest-Neighbor Model*) sous-jacent à ces calculs stipule que la stabilité globale d'une structure résulte de la somme des contributions énergétiques de chaque paire de bases consécutive et des éléments de boucle associées (Mathews et al., 1999).

2.5.1 Variation d'énergie libre de liaison ($\Delta\Delta G$ binding)

L'effet des variants sur la force de l'interaction intermoléculaire U4 /U6 a été quantifié par le calcul du $\Delta\Delta G$ binding, défini par l'équation :

$$\Delta\Delta G_{\text{binding}} = \Delta G_{\text{binding,variant}} - \Delta G_{\text{binding,WT}}$$

Où $\Delta G_{\text{binding,variant}}$ et $\Delta G_{\text{binding,WT}}$ correspondent aux énergies libres de liaison du complexe U4/U6 calculées par RNAcofold pour la séquence mutante et pour la séquence sauvage, respectivement. Une valeur positive indique un affaiblissement de l'interaction U4/U6 induit par le variant, tandis qu'une valeur négative suggère un renforcement de cette interaction.

2.5.2 Interprétation des résultats

L'interprétation des valeurs de $\Delta\Delta G$ permet de classer les variants selon leur effet sur la conformation de l'ARN.

- $\Delta\Delta G \approx 0 \rightarrow$ neutre
- $\Delta\Delta G > +2 \rightarrow$ déstabilisant
- $\Delta\Delta G < -2 \rightarrow$ stabilisant

2.6 Calcul de la distance structurale (RNAdistance)

En complément des métriques, les réarrangements topologiques induits par chaque variant ont été quantifiés par le calcul de la distance structurale entre la structure WT et chaque structure mutante. A l'aide du programme RNAdistance (Lorenz et al., 2011). RNAdistance implémente la distance d'édition d'arbres (Tree Edit Distance, TED), qui représente chaque structure dot-bracket comme un arbre hiérarchique et calcule le nombre minimal d'opérations (insertion, suppression, substitution de nœuds) nécessaires pour transformer une structure en l'autre.

Cette métrique est particulièrement adaptée à la détection de variants qui induisent des réarrangements conformationnels majeurs sans nécessairement modifier l'énergie libre globale phénomène fréquemment observé pour les variants du hotspot Stem II de RNU4-2 (positions 62–79).

Une distance structurale élevée suggère que la mutation ne change pas seulement une base, mais déstabilise une tige (stem) ou crée une nouvelle boucle (loop), ce qui peut :

- **Altérer les interactions protéine-ARN** : Si le site de liaison est masqué ou modifié.
- **Modifier la stabilité de l'ARN** : Rend la molécule plus sensible à la dégradation.
- **Perturber la fonction catalytique** : Crucial pour les ARN non-codants (comme les snRNA du complexe spliceosome).

Pour les structures complexes incluant des pseudonoeuds, la notation est étendue avec des crochets [], des accolades {} ou des lettres $\langle \rangle$, permettant de différencier les hélices qui s'entrecroisent.

2.7 Analyses statistiques

L'ensemble des analyses statistiques a été réalisé en Python à l'aide des bibliothèques SciPy et statsmodels (interface web google colab). Les distributions n'étant pas normales et les effectifs étant limités, des tests non paramétriques ont été employés. Les trois métriques ($\Delta\Delta G$ RNAfold, $\Delta\Delta G$ RNAcofold, RNAdistance) ont été comparées entre les groupes Patho/LP ($n = 30$) et VUS ($n = 31$) par le test de Mann-Whitney U bilatéral. Une correction de Bonferroni a été appliquée pour les trois comparaisons simultanées, fixant le seuil de significativité à $\alpha = 0,0167$. La taille d'effet a été estimée par le delta de Cliff, dont la magnitude a été interprétée selon les seuils de Romano et al. (2006).

L'effet de la région fonctionnelle sur chaque métrique a été évalué par le test de Kruskal-Wallis, restreint aux régions comptant au moins cinq variants, suivi d'une analyse post-hoc de Dunn avec correction de Bonferroni. L'enrichissement des variants pathogènes dans la région critique (T-loop, positions 62 à 70, et Stem III, positions 75 à 78) a été quantifié par le test exact de Fisher, avec estimation de l'intervalle de confiance à 95 % de l'odds ratio. Enfin, les relations entre les trois métriques ont été analysées par la corrélation de rang de Spearman.

2.8 Stratégie d'analyse comparative VUS / variants pathogènes

Les distributions de chaque métrique ont été comparées entre les deux groupes (Patho/LP, n = 30 ; VUS, n = 31). La localisation de chaque variant dans les domaines fonctionnels de l'ARN U4 a également été prise en compte comme variable qualitative complémentaire, afin d'examiner si les variants pathogènes se concentrent préférentiellement dans les régions structurellement les plus contraintes. L'ensemble de cette démarche vise à contribuer, à terme, au développement de critères de classification fonctionnelle applicables aux variants des gènes de petits ARN nucléaires, pour lesquels les outils standards d'interprétation clinique (conçus pour les séquences codantes) sont peu performants.

Chapitre III : Résultats

3.1 Description des variants étudiés

Soixante-et-un variants RNU4-2 ont été retenus pour l'analyse comparative, les deux variants à classification conflictuelle ayant été écartés. Suivant les recommandations ACMG/AMP, ils se répartissent en deux catégories : trente variants pathogènes ou probablement pathogènes (Patho/LP, 49,2 %) et trente-et-un variants de signification incertaine (VUS, 50,8 %). Le **Tableau 3** détaille cette répartition.

Le long de l'ARN U4, les variants ne se distribuent pas uniformément. Le duplex stem II U4/U6 en rassemble à lui seul un tiers, loin devant les autres domaines (**Figure 10**). Le contexte structural est en revanche presque équilibré : 33 variants occupent une boucle (loop) et 28 une tige (stem).

Tableau 3 : Répartition des 61 variants selon la classification ClinVar.

Classification ClinVar	n	%
Pathogenic (P)	9	14,8
Pathogenic/Likely pathogenic (P/LP)	3	4,9
Likely pathogenic (LP)	18	29,5
Uncertain significance (VUS)	31	50,8
Total	61	100

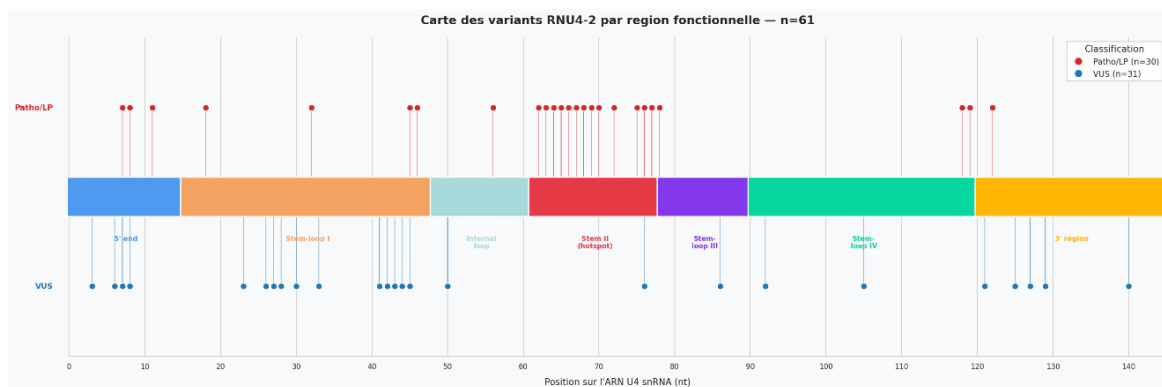


Figure 10: Carte des 61 variants RNU4-2 par région fonctionnelle (Patho/LP en rouge, VUS en bleu).

3.2 Impact des variants sur la structure secondaire de l'ARN U4 (RNAfold)

Pour chaque variant, RNAfold a fourni l'énergie libre minimale de la séquence mutante, rapportée à celle du type sauvage (ΔG WT = $-44,0$ kcal/mol). Le **Tableau 4** réunit les valeurs de $\Delta\Delta G$ par groupe. Le $\Delta\Delta G$ mesure l'effet de chaque variant sur la stabilité de la structure secondaire de l'ARN U4, une valeur négative traduit une stabilisation alors qu'une valeur positive une déstabilisation. La majorité des variants présentent un $\Delta\Delta G$ proche de zéro, sans effet marqué sur la stabilité globale.

Chez les Patho/LP, n.75C>G (duplex stem II U4/U6) présente un $\Delta\Delta G$ de $-3,0$ kcal/mol, correspondant à une stabilisation de la structure. À l'inverse, n.56T>C, situé dans la même région, atteint $+4,0$ kcal/mol, traduisant une déstabilisation (**Figure 11**).

Chez les VUS, n.140G>A présente le $\Delta\Delta G$ le plus élevé ($+6,4$ kcal/mol), correspondant à la déstabilisation la plus forte parmi tous les variants étudiés. Les autres VUS restent proches de zéro.

Plusieurs variants ont un $\Delta\Delta G$ nul : n.69C>T et n.68A>C (Patho/LP, duplex stem II U4/U6), n.125G>A (VUS, région 3') et n.7G>C (VUS, extrémité 5') conservent la valeur du type sauvage (**Annexe 1**).

Tableau 4 : $\Delta\Delta G$ RNAfold par groupe.

Groupe	n	$\Delta\Delta G$ moyen (kcal/mol)	Médiane (kcal/mol)
Pathogenic (P)	9	-0,02	0,00
P/LP	3	-0,73	0,00
Likely pathogenic (LP)	18	-0,24	0,00
VUS	31	+1,54	+1,80

Les médianes des trois sous-groupes pathogènes sont nulles. Les VUS, eux, se déplacent vers les valeurs positives (médiane +1,80 kcal/mol) et couvrent un intervalle plus large. Certains VUS atteignent même les $\Delta\Delta G$ RNAfold les plus élevés observés, tels que n.140G>A (+6,4 kcal/mol), n.129G>C et n.129G>A (+6,2 kcal/mol), des valeurs comparables aux plus extrêmes mesurées parmi les variants pathogènes.

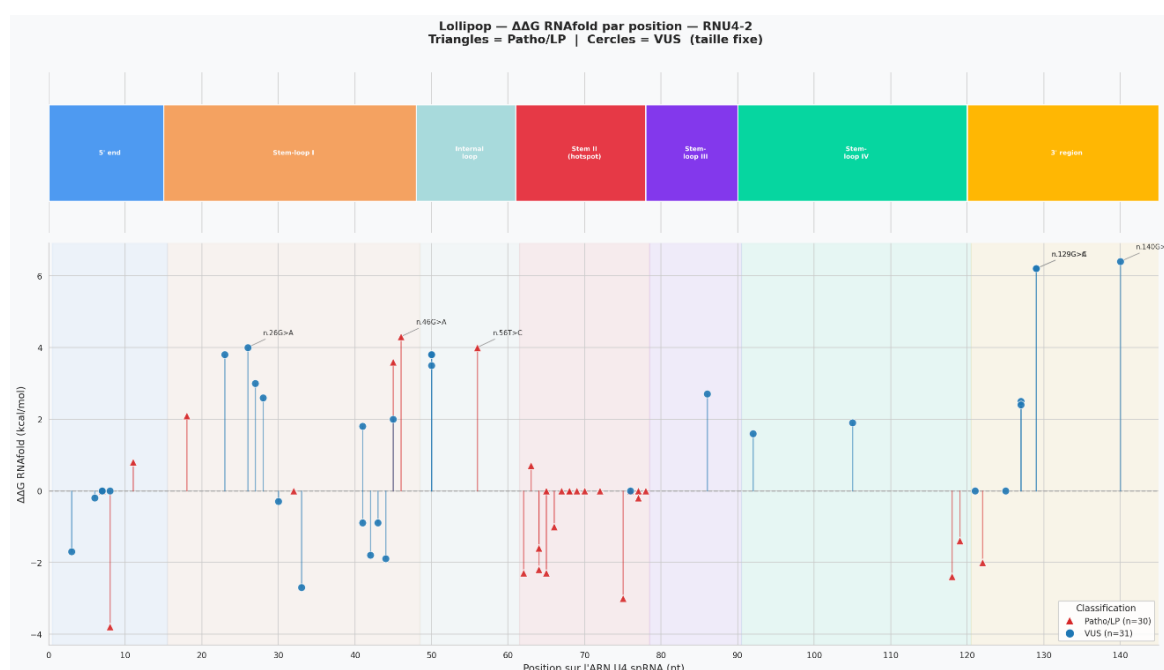


Figure 11 : $\Delta\Delta G$ RNAfold de chaque variant RNU4-2 selon sa position sur l'ARN U4 snRNA. Triangles = variants Patho/LP (n=30), cercles = variants VUS (n=31). Les variants présentant les $\Delta\Delta G$ les plus extrêmes sont annotés.

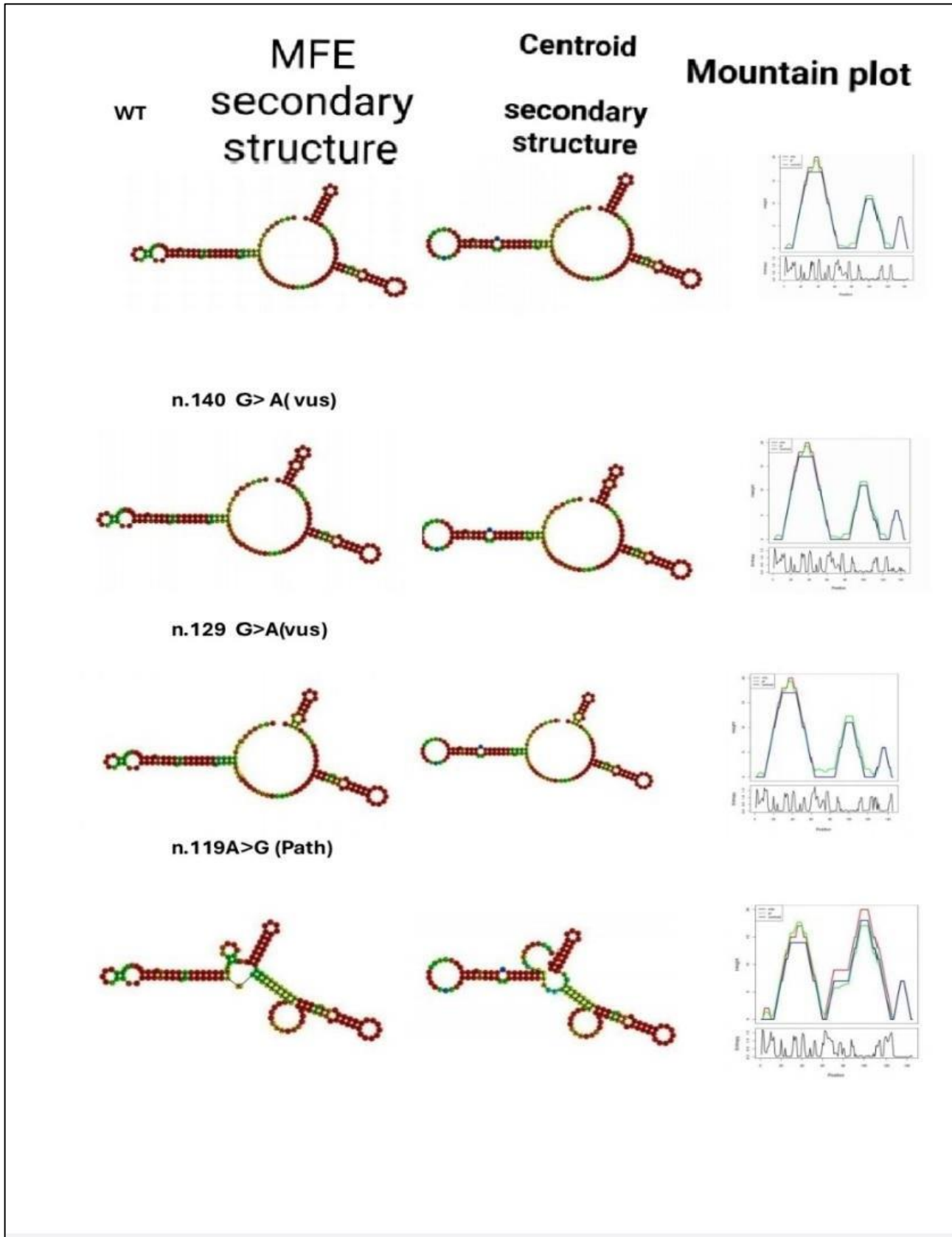


Figure 12 : Analyse comparative des structures secondaires de l'ARN U4 sauvage et des variants RNU4-2 par RNAfold.

3.3 Impact des variants sur l'interaction U4/U6 (RNAcofold)

RNAcofold modélise le duplex U4/U6, dont l'énergie de liaison atteint $-82,6$ kcal/mol pour le type sauvage. Le $\Delta\Delta G$ mesure l'effet de chaque variant sur cette interaction : une valeur négative traduit un renforcement de l'appariement, une valeur positive un affaiblissement. La médiane du $\Delta\Delta G$ est de $+0,10$ kcal/mol chez les Patho/LP et de $+2,30$ kcal/mol chez les VUS, ces derniers couvrant un éventail plus dispersé (**Tableau 5**).

Chez les VUS, n.33A>G (stem-loop I) présente le $\Delta\Delta G$ le plus négatif ($-2,7$ kcal/mol), correspondant au renforcement le plus marqué de l'appariement par rapport au type sauvage. À l'inverse, n.50G>C (internal loop) atteint $+2,1$ kcal/mol, traduisant un affaiblissement.

n.7G>C, au même nucléotide que le variant n.7G>A, présente le $\Delta\Delta G$ le plus élevé parmi tous les variants étudiés ($+6,7$ kcal/mol), correspondant à l'affaiblissement le plus fort de l'appariement avec l'ARN U6.

Plusieurs variants ont un $\Delta\Delta G$ nul : n.119A>G (stem-loop IV) et n.69C>T (duplex stem II U4/U6), classés Patho/LP, ainsi que n.125G>A (VUS, région 3'). Ces substitutions ne modifient pas la stabilité du duplex selon RNAcofold. L'effet des variants sur l'interaction U4/U6 dépend ainsi de leur localisation et de leur nature (**Annexe 2**).

Tableau 5 : $\Delta\Delta G$ RNAcofold par groupe.

Groupe	n	$\Delta\Delta G$ RNAcofold moyen	Médiane
Pathogenic (P)	9	+0,50	+0,20
P/LP	3	-0,67	0,00
Likely pathogenic (LP)	18	+1,36	+1,25
VUS	31	+2,06	+2,30

3.4 Analyse de la distance structurale (RNAdistance)

La distance entre paires de bases mesure l'ampleur du réarrangement topologique par rapport au type sauvage. Les médianes des deux groupes sont identiques (4 paires de bases), mais le groupe Patho/LP se distingue par une moyenne plus élevée (11,07 contre 6,81) et par des valeurs extrêmes atteignant 44 paires de bases, contre 24 au maximum chez les VUS.

Tableau 6 : RNAdistance par groupe.

Groupe	n	RNAdistance moyenne	Médiane
Pathogenic (P)	9	16,22	10,00
P/LP	3	9,00	0,00
Likely pathogenic (LP)	18	8,83	3,00
VUS	31	6,81	4,00

La Figure 12 rapproche les structures secondaires (MFE et centroïde) de quelques variants représentatifs de celle du type sauvage. Le pathogène n.67A>G (stem II, RNAdistance = 0) conserve la conformation native, là où n.66A>G (RNAdistance = 36) et n.119A>G (RNAdistance = 44) la remanient profondément. Le VUS n.140G>A entraîne lui aussi une réorganisation visible.

3.5 Variations par région fonctionnelle

Indépendamment de la comparaison entre groupes cliniques (section 3.6), un test de Kruskal-Wallis a évalué l'effet de la région fonctionnelle, restreint aux régions comptant au moins cinq variants (**Figure 13**).

$\Delta\Delta G$ RNAfold

Le test atteint $H = 11,22$ ($p = 0,0243$). L'analyse post-hoc de Dunn-Bonferroni ne désigne toutefois aucune paire de régions, signe d'une hétérogénéité diffuse plutôt que localisée. Les $\Delta\Delta G$ les plus positifs se rencontrent dans la région 3' et la boucle interne ; les plus négatifs à l'extrémité 5', dans le stem II et le stem-loop IV.

$\Delta\Delta G$ RNAfold

Ici, l'effet régional est plus marqué ($H = 15,95$; $p = 0,0031$). Le post-hoc isole une seule paire significative, l'extrémité 5' contre le stem II ($p = 0,003$). C'est aussi à l'extrémité 5' que se concentrent les valeurs positives les plus fortes, tandis que la région 3' affiche la plus grande dispersion.

RNAdistance

Aucune différence entre régions n'est mise en évidence ($H = 5,54$; $p = 0,2359$).

Tableau 7 : Médiane [intervalle interquartile] des trois métriques par région fonctionnelle (ensemble des variants, $n = 61$).

Région	n	$\Delta\Delta G$ RNAfold (méd. [IQR])	$\Delta\Delta G$ RNAcofold (méd. [IQR])	RNAdistance (méd. [IQR])
5' end	7	0,0 [-0,9 ; 0,0]	+3,8 [+3,0 ; +5,2]	7 [0,5 ; 18]
Stem-loop I	9	+2,1 [0,0 ; +3,0]	+2,6 [0,0 ; +3,6]	7 [4 ; 12]
Internal loop	11	+2,0 [-0,9 ; +3,7]	+2,0 [-0,9 ; +3,3]	8 [6,5 ; 14]
Stem II (hotspot)	21	0,0 [-1,0 ; 0,0]	0,0 [-0,2 ; +1,4]	1 [0 ; 16]
Stem-loop III	2	+2,2 [+1,9 ; +2,4]	+2,2 [+1,9 ; +2,4]	2 [1 ; 3]
Stem-loop IV	3	-1,4 [-1,9 ; +0,2]	0,0 [-1,2 ; +0,9]	4 [4 ; 24]
3' region	8	+2,5 [0,0 ; +6,2]	+2,5 [0,0 ; +6,2]	4 [0 ; 4]

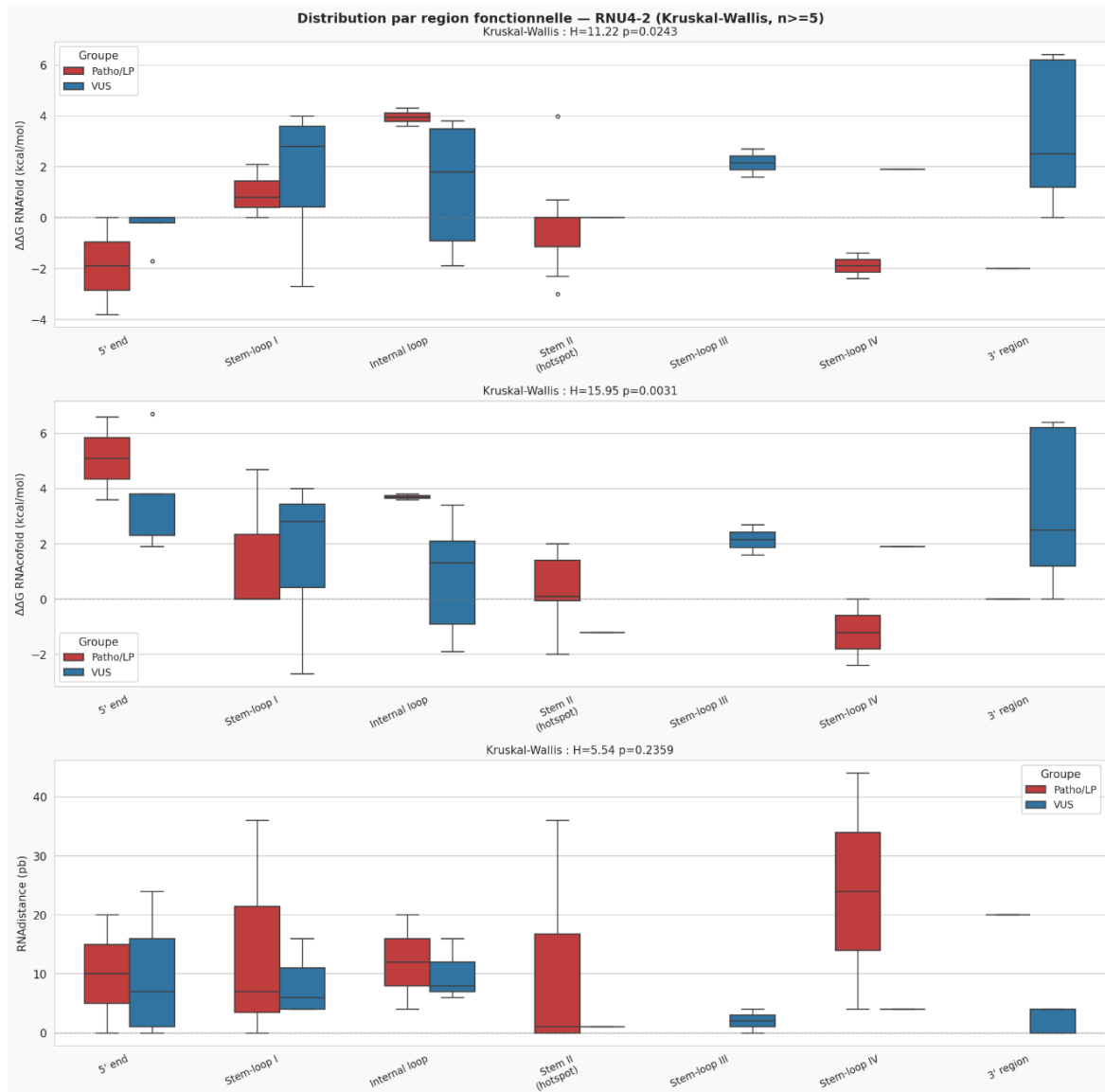


Figure 13 : Distribution des trois métriques ($\Delta\Delta G$ RNAfold, $\Delta\Delta G$ RNAcifold, RNAdistance) par région fonctionnelle, comparaison Patho/LP et VUS (Kruskal-Wallis, post-hoc Dunn-Bonferroni).

3.6 Comparaison globale Patho/LP vsVUS

Les trois métriques ont fait l'objet d'un test de Mann-Whitney U bilatéral entre Patho/LP (n = 30) et VUS (n = 31), assorti d'une correction de Bonferroni pour trois comparaisons ($\alpha = 0,0167$) et d'une taille d'effet (delta de Cliff). Une fois la correction appliquée, seul le $\Delta\Delta G$ RNAfold sépare les deux groupes (p = 0,012) ; le $\Delta\Delta G$ RNAcifold (p = 0,049) et la RNAdistance (p =

0,861) restent en deçà du seuil. Un delta de Cliff négatif correspond à des valeurs plus basses dans le groupe Patho/LP. Seuils de magnitude selon Romano et al. (2006).

Tableau 8 : Comparaison Patho/LP vs VUS tests de Mann-Whitney U et tailles d'effet.

Métrique	p brut	Bonferroni ($\alpha=0,0167$)	Delta de Cliff	Magnitude
$\Delta\Delta G$ RNAfold	0,012	significatif	-0,37	moyen
$\Delta\Delta G$ RNAcofold	0,049	non significatif	-0,29	faible
RNAdistance	0,861	non significatif	+0,03	négligeable

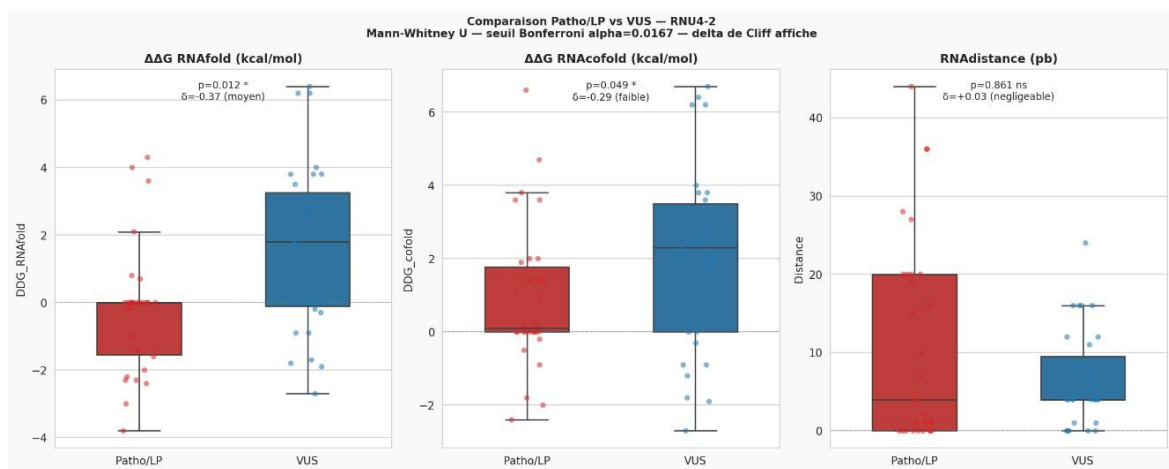


Figure 14 : Comparaison des trois métriques ($\Delta\Delta G$ RNAfold, $\Delta\Delta G$ RNAcofold, RNAdistance) entre les groupes Patho/LP et VUS. Test de Mann-Whitney U, seuil de Bonferroni $\alpha = 0,0167$, delta de Cliff indiqué.

3.7 Enrichissement des variants pathogènes dans la région critique

La région critique de RNU4-2 regroupe deux sous-structures conservées, le T-loop (positions 62 à 70) et le Stem III (positions 75 à 78), telles que définies par Chen et al. (2024) puis par

l'étude de saturation genome editing (2026). Un test exact de Fisher a quantifié la répartition des variants de part et d'autre de cette région (9).

Tableau 9: Répartition des variants selon la région critique (T-loop + Stem III), test exact de Fisher.

Localisation	Patho/LP	VUS	Total
Région critique	18	1	19
Hors région critique	12	30	42
Total	30	31	61
Fisher exact	OR = 45,0	$p = 9,7 \times 10^{-7}$	
IC95 % (OR)	5,4 – 375,7		

Dans la région critique se trouvent 18 des 30 variants Patho/LP, contre un seul des 31 VUS, soit un rapport de cotes de 45,0 ($p = 9,7 \times 10^{-7}$; IC95 % : 5,4 à 375,7). L'intervalle de confiance, large, tient au faible nombre de variants concernés ; la valeur du test exact confirme néanmoins la robustesse de l'association.

3.8 Corrélations entre les trois métriques

Les corrélations de rang de Spearman, calculées sur les 61 variants, lient étroitement le $\Delta\Delta G$ RNAfold et le $\Delta\Delta G$ RNAcofold ($\rho = 0,62$; $p < 0,001$). La RNAdistance, en revanche, reste indépendante des deux mesures énergétiques ($\rho = -0,19$ et $\rho = -0,10$; $p > 0,15$).

3.9 Classification de l'impact structural (notation dot-bracket)

Chaque structure dot-bracket a été comparée à celle du type sauvage, puis classée par niveau d'impact. Le Tableau 10 croise cette classification avec le statut ClinVar.

Tableau 10 : Répartition des variants par niveau d'impact structural et par classification.

Classification	Neutre	Faible	Modéré	Majeur	n
Pathogenic (P)	0	2	6	1	9
P/LP	2	0	1	0	3
Likely pathogenic (LP)	3	2	9	4	18
Patho/LP (total)	5	4	16	5	30
VUS	2	2	9	18	31

Un contraste ressort du tableau : l'impact majeur touche 18 VUS mais seulement 5 variants Patho/LP, tandis que l'impact modéré domine chez ces derniers (16 sur 30).

Chapitre IV : Discussion

Discussion

Le gène RNU4-2 code l'ARN nucléaire U4, composant du spliceosome majeur, dont les variants de novo sont responsables du syndrome ReNU, décrit pour la première fois en 2024 et caractérisé par une déficience intellectuelle, une hypotonie, un retard moteur et des crises convulsives (Greene et al., 2024 ; Chen et al., 2024). Ce syndrome représente l'une des causes monogéniques de trouble neurodéveloppemental les plus fréquentes, avec une prévalence estimée à 0,4 % des cas (Chen et al., 2024). Au sein du spliceosome, l'ARN U4 s'apparie à l'ARN U6 pour former le duplex U4/U6, intégré au tri-snRNP U4/U6*U5, précurseur de la forme catalytiquement active (Nguyen et al., 2016). Les variants pathogènes se concentrent dans une région critique de dix-huit nucléotides et perturbent l'épissage en altérant la sélection du site donneur en 5' (Chen et al., 2024). Ce travail avait pour objectif d'analyser, par des approches *in silico*, l'impact des variants de RNU4-2 sur la structure et la fonction de l'ARN U4, afin de mieux comprendre les mécanismes moléculaires impliqués dans le syndrome ReNU.

Le $\Delta\Delta G$ RNAfold est la seule métrique distinguant significativement les groupes Patho/LP et VUS ($p = 0,012$; delta de Cliff = $-0,37$). Cependant, les variants Patho/LP présentent un $\Delta\Delta G$ médian nul, ce qui indique que la déstabilisation globale de la structure secondaire n'est pas la règle parmi les variants pathogènes. RNAfold modélise l'ARN U4 hors de son contexte fonctionnel, sans tenir compte des interactions protéiques ni de la structure du duplex U4/U6 ; la pathogénicité des variants de RNU4-2 s'exerce précisément à ce niveau. Les $\Delta\Delta G$ les plus élevés sont observés chez des variants classés VUS (n.140G>A, +6,4 kcal/mol ; n.129G>C et n.129G>A, +6,2 kcal/mol), non chez les variants pathogènes connus. La stabilité globale de l'ARN U4, prise isolément, ne suffit donc pas à rendre compte de la pathogénicité dans ce corpus, et les effets délétères pourraient dépendre davantage de perturbations localisées dans des régions fonctionnellement contraintes que d'une perte d'énergie libre globale.

L'appariement U4/U6 conditionne l'assemblage du tri-snRNP (Nguyen et al., 2016). La différence entre groupes pour le $\Delta\Delta G$ RNAfold n'atteint pas le seuil de Bonferroni ($p = 0,049$), mais les valeurs tendent à être plus élevées chez les VUS. La position 7 fournit un exemple instructif : n.7G>A, classé probablement pathogène, et le VUS n.7G>C affaiblissent tous deux fortement l'appariement ($\Delta\Delta G$ RNAfold de +6,6 et +6,7 kcal/mol respectivement), alors que leur $\Delta\Delta G$ RNAfold est nul.

Ces variants illustrent un mécanisme dans lequel l'effet s'exerce exclusivement sur l'interface d'interaction U4/U6, sans altération de la structure isolée de l'ARN U4, mécanisme que la seule analyse RNAfold ne permettrait pas de détecter. À l'inverse, quelques variants présentent un $\Delta\Delta G$ RNAfold négatif, ce qui pourrait théoriquement gêner la dissociation du duplex U4/U6 nécessaire à l'activation du spliceosome, bien que cette hypothèse reste spéculative en l'absence de validation fonctionnelle. Ces observations suggèrent que certaines altérations associées au syndrome ReNU résultent d'une perturbation de l'assemblage ou du remodelage du complexe U4/U6 plutôt que d'une déstabilisation globale de l'ARN U4.

La RNAdistance quantifie la réorganisation topologique indépendamment de l'énergie. Plusieurs variants pathogènes combinent un $\Delta\Delta G$ faible et un réarrangement topologique majeur : n.119A>G (stem-loop IV) atteint la distance la plus élevée du jeu de données (44) avec un $\Delta\Delta G$ RNAfold de $-1,4$ kcal/mol et une interaction U4/U6 inchangée ; n.66A>G (duplex stem II) présente une distance de 36 pour un $\Delta\Delta G$ faiblement stabilisant ($-1,0$ kcal/mol). Ces cas suggèrent un mécanisme impliquant une distorsion de la topologie moléculaire sans signature énergétique nette, susceptible d'influer sur certaines interactions protéine-ARN sans modifier l'énergie libre globale.

La saturation genome editing de RNU4-2 (De Jonghe et al., 2026) a précisé, au nucléotide près, les régions fonctionnellement contraintes du gène : Le duplex d'interaction avec l'ARN U6 en stem II, les segments flanquant le k-turn de liaison à la protéine 15.5K, le site Sm (nt 118–126), et la région critique ReNU subdivisée en T-loop (nt 62–70) et Stem III (nt 75–78). Ces interfaces protéine-ARN ne sont pas reproduites par la modélisation thermodynamique de la séquence isolée, ce qui explique que certains variants pathogènes situés dans le k-turn ou le site Sm présentent des métriques in silico proches de zéro malgré un impact fonctionnel établi.

Le signal le plus fort observé dans cette étude est la localisation des variants. Dix-huit des 30 variants Patho/LP se situent dans la région critique, contre un seul des 31 VUS (OR = 45,0 ; IC95 % : 5,4–375,7 ; $p = 9,7 \times 10^{-7}$). Cette association dépasse en amplitude celle des métriques structurales et est cohérente avec les données publiées sur le syndrome ReNU (**Chen et al., 2024 ; De Jonghe et al., 2026**). L'intervalle de confiance large tient au faible effectif ; la valeur du test exact confirme la robustesse de l'association. Cela ne signifie pas que les variants situés hors de la région critique soient bénins, ni que les VUS en soient exclus par définition, mais indique que la localisation constitue le critère le plus discriminant dans ce corpus.

Un résultat méritant attention est la fréquence élevée d'altérations structurales majeures parmi les VUS : dix-huit présentent un impact structural majeur, contre cinq des variants Patho/LP. Cette observation peut refléter la présence de variants potentiellement pathogènes non encore reconnus parmi les VUS, ou indiquer que certaines modifications structurales importantes ne se traduisent pas nécessairement par une altération fonctionnelle du spliceosome.

Le groupe VUS est hétérogène par construction. Parmi les candidats prioritaires pour une validation fonctionnelle, n.7G>C se distingue : situé au même nucléotide que le variant probablement pathogène n.7G>A, il présente un profil quasi identique ($\Delta\Delta G$ RNAcofold +6,7 contre +6,6 kcal/mol) et un score de fonction déplété en SGE (-0,97), par ailleurs associé à une forme récessive du trouble (**De Jonghe et al., 2026**). Les variants n.140G>A et n.129G>A, qui présentent les déstabilisations énergétiques les plus fortes, sont également déplétés en SGE (scores -0,40 et -0,47). À l'inverse, n.56T>C, rapporté dans un contexte de rétinite pigmentaire sans lien établi avec le syndrome ReNU, obtient un score de fonction proche de zéro (-0,26), ce qui est cohérent avec l'absence de mécanisme commun. Ces données illustrent la complémentarité des approches *in silico* et expérimentales : les métriques structurales permettent de hiérarchiser les candidats, mais la reclassification requiert une validation fonctionnelle.

Les résultats indiquent que les mécanismes associés au syndrome ReNU ne se réduisent pas à un unique mode d'altération structurale. Certains variants n'affectent que l'interface U4/U6 sans modifier la structure isolée (n.7G>A, n.7G>C) ; d'autres réorganisent la topologie à énergie constante (n.119A>G, n.66A>G). La corrélation entre $\Delta\Delta G$ RNAfold et $\Delta\Delta G$ RNAcofold ($\rho = 0,62$; $p < 0,001$) indique qu'une partie des perturbations de la structure intrinsèque se répercute sur l'interaction U4/U6, mais plusieurs variants présentent un découplage entre ces deux paramètres, montrant que certains effets sont spécifiques de l'interface intermoléculaire. Cette diversité de mécanismes souligne l'intérêt d'une approche multi-critères, en complément des méthodes fonctionnelles expérimentales.

Cette étude présente néanmoins plusieurs limites. Les prédictions *in silico* ne reproduisent pas le contexte cellulaire, les modèles n'intègrent pas la structure tridimensionnelle, les protéines du spliceosome, les pseudo-nœuds, ni les modifications post-transcriptionnelles. Les

seuils d'interprétation des $\Delta\Delta G$ sont indicatifs, aucun seuil validé n'existant pour RNU4-2. Les effectifs par région sont parfois faibles, ce qui limite la puissance des tests post-hoc. Les classifications ClinVar utilisées comme référence reposent, pour 47 des 61 variants, sur un seul soumetteur, ce qui correspond à un niveau de preuve limité, et elles sont susceptibles d'évoluer avec l'accumulation des données pour un gène aussi récemment décrit.

Conclusion

Cette recherche a permis d'examiner, par le biais de méthodes *in silico*, les effets structurels éventuels des variantes du gène RNU4-2 associées au syndrome neurodéveloppemental ReNU. L'examen des variantes a révélé leur localisation préférentielle dans des zones fonctionnelles cruciales de l'ARN U4, en particulier dans les domaines concernés par l'établissement du duplex U4/U6 et par les interactions nécessaires au bon fonctionnement du spliceosome.

Les prévisions de la structure secondaire ont montré que plusieurs variations sont liées à des changements dans la stabilité thermodynamique et la conformation de l'ARN. Certaines substitutions provoquent d'importantes perturbations au sein de l'architecture moléculaire, susceptibles de modifier les interactions de l'ARN U4 avec ses partenaires moléculaires et de nuire à l'efficacité du mécanisme d'épissage. À l'opposé, d'autres variantes montrent des effets structurels réduits, laissant penser à un impact fonctionnel possiblement moins marqué.

L'étude comparative entre les variants pathogènes ou probablement pathogènes et ceux de signification indéterminée a révélé que les premiers se trouvent principalement dans des zones hautement préservées et fonctionnellement importantes, soulignant l'importance de la stabilité structurelle de l'ARN U4 pour maintenir le fonctionnement normal du spliceosome. Ces constatations renforcent l'idée que les changements dans la structure de l'ARN représentent l'un des principaux mécanismes moléculaires à l'origine du syndrome ReNU.

Bien que les méthodes bioinformatiques fournissent des données précieuses sur les impacts potentiels des variants, leurs résultats doivent être considérés avec précaution et complétés par des investigations expérimentales pour vérifier les effets fonctionnels constatés. Toutefois, cette étude illustre l'utilité des outils de prédiction structurelle dans l'analyse des ARN non codants et dans la compréhension des variants génétiques liés aux maladies humaines.

En conclusion, ce travail participe à une meilleure appréhension des mécanismes moléculaires en jeu dans le syndrome neurodéveloppemental ReNU et met en lumière l'importance cruciale de la structure de l'ARN U4 dans la pathogénie de cette maladie. Les résultats obtenus offrent une base pratique pour des recherches futures visant à approfondir les relations entre structure, fonction et pathogénicité des variantes du gène RNU4-2.

ANNEXE 1

Tableau de RNAfold :

Variant	Classification	Position	Contexte	Région	Long. WT	Long. mut	ΔG WT	ΔG mut	$\Delta \Delta G$ RNAfold	Interprétation
n.119A>G	Pathogenic	119	loop	stem_loop_IV	145	145	-44	-45.4	-1.4	Modérément stabilisant
n.77_78insT	Pathogenic	77	loop	stem_II_U4-U6_duplex	145	146	-44	-44	+0	Neutre
n.76C>T	Pathogenic	76	loop	stem_II_U4-U6_duplex	145	145	-44	-44	+0	Neutre
n.68dup	Pathogenic	68	loop	stem_II_U4-U6_duplex	145	146	-44	-44	+0	Neutre
n.66A>G	Pathogenic	66	loop	stem_II_U4-U6_duplex	145	145	-44	-45	-1	Neutre
n.65A>G	Pathogenic	65	loop	stem_II_U4-U6_duplex	145	145	-44	-46.3	-2.3	Fortement stabilisant
n.64_65insT	Pathogenic	64	stem	stem_II_U4-U6_duplex	145	146	-44	-45.6	-1.6	Modérément stabilisant
n.56T>C	Pathogenic	56	stem	stem_loop_I	145	145	-44	-40	+4	Fortement déstabilisant
n.18_19insA	Pathogenic	18	stem	stem_loop_I	145	146	-44	-41.9	+2.1	Fortement déstabilisant
n.69C>T	Pathogenic/LP	69	loop	stem_II_U4-U6_duplex	145	145	-44	-44	+0	Neutre
n.67A>G	Pathogenic/LP	67	loop	stem_II_U4-U6_duplex	145	145	-44	-44	+0	Neutre
n.64dup	Pathogenic/LP	64	stem	stem_II_U4-U6_duplex	145	146	-44	-46.2	-2.2	Fortement stabilisant
n.122T>G	Likely pathogenic	122	loop	3' region	145	145	-44	-46	-2	Modérément stabilisant
n.118A>G	Likely pathogenic	118	loop	stem_loop_IV	145	145	-44	-46.4	-2.4	Fortement stabilisant
n.78A>C	Likely pathogenic	78	loop	stem_II_U4-U6_duplex	145	145	-44	-44	+0	Neutre
n.77_78insG	Likely pathogenic	77	loop	stem_II_U4-U6_duplex	145	146	-44	-44.2	-0.2	Neutre

ANNEXES

n.76_77insT	Likely pathogenic	76	loop	stem_II_U4-U6_duplex	145	146	-44	-44	+0	Neutre
n.75C>G	Likely pathogenic	75	stem	stem_II_U4-U6_duplex	145	145	-44	-47	-3	Fortement stabilisant
n.72_73del	Likely pathogenic	72	loop	stem_II_U4-U6_duplex	145	143	-44	-44	+0	Neutre
n.70T>C	Likely pathogenic	70	loop	stem_II_U4-U6_duplex	145	145	-44	-44	+0	Neutre
n.68A>C	Likely pathogenic	68	loop	stem_II_U4-U6_duplex	145	145	-44	-44	+0	Neutre
n.65_66insT	Likely pathogenic	65	loop	stem_II_U4-U6_duplex	145	146	-44	-44	+0	Neutre
n.63T>C	Likely pathogenic	63	stem	stem_II_U4-U6_duplex	145	145	-44	-43.3	+0.7	Neutre
n.62T>C	Likely pathogenic	62	stem	stem_II_U4-U6_duplex	145	145	-44	-46.3	-2.3	Fortement stabilisant
n.46G>A	Likely pathogenic	46	stem	internal_loop	145	145	-44	-39.7	+4.3	Fortement déstabilisant
n.45G>C	Likely pathogenic	45	stem	internal_loop	145	145	-44	-40.4	+3.6	Fortement déstabilisant
n.32G>A	Likely pathogenic	32	loop	stem_loop_I	145	145	-44	-44	+0	Neutre
n.11A>C	Likely pathogenic	11	stem	stem_loop_I	145	145	-44	-43.2	+0.8	Neutre
n.8C>G	Likely pathogenic	8	loop	5'_end	145	145	-44	-47.8	-3.8	Fortement stabilisant
n.7G>A	Likely pathogenic	7	loop	5'_end	145	145	-44	-44	+0	Neutre
n.140G>A	VUS	140	stem	3'_region	145	145	-44	-37.6	+6.4	Fortement déstabilisant
n.129G>C	VUS	129	stem	3'_region	145	145	-44	-37.8	+6.2	Fortement déstabilisant
n.129G>A	VUS	129	stem	3'_region	145	145	-44	-37.8	+6.2	Fortement déstabilisant
n.127_128del	VUS	127	stem	3'_region	145	143	-44	-41.5	+2.5	Fortement déstabilisant
n.127C>T	VUS	127	stem	3'_region	145	145	-44	-41.6	+2.4	Fortement déstabilisant
n.125G>A	VUS	125	loop	3'_region	145	145	-44	-44	+0	Neutre
n.121T>A	VUS	121	loop	3'_region	145	145	-44	-44	+0	Neutre
n.105A>T	VUS	105	stem	stem_loop_IV	145	145	-44	-42.1	+1.9	Modérément déstabilisant
n.92C>G	VUS	92	stem	stem_loop_III	145	145	-44	-42.4	+1.6	Modérément déstabilisant
n.86C>T	VUS	86	stem	stem_loop_III	145	145	-44	-41.3	+2.7	Fortement déstabilisant

ANNEXES

n.76del	VUS	76	loop	stem_II_U4-U6_duplex	145	144	-44	-44	+0	Neutre
n.50G>T	VUS	50	stem	internal_loop	145	145	-44	-40.5	+3.5	Fortement déstabilisant
n.50G>C	VUS	50	stem	internal_loop	145	145	-44	-40.2	+3.8	Fortement déstabilisant
n.50G>A	VUS	50	stem	internal_loop	145	145	-44	-40.2	+3.8	Fortement déstabilisant
n.45_46insT	VUS	45	stem	internal_loop	145	146	-44	-42	+2	Modérément déstabilisant
n.44A>T	VUS	44	loop	internal_loop	145	145	-44	-45.9	-1.9	Modérément stabilisant
n.43_44insT	VUS	43	loop	internal_loop	145	146	-44	-44.9	-0.9	Neutre
n.42C>G	VUS	42	stem	internal_loop	145	145	-44	-45.8	-1.8	Modérément stabilisant
n.41C>T	VUS	41	stem	internal_loop	145	145	-44	-44.9	-0.9	Neutre
n.41C>G	VUS	41	stem	internal_loop	145	145	-44	-42.2	+1.8	Modérément déstabilisant
n.33A>G	VUS	33	loop	stem_loop_I	145	145	-44	-46.7	-2.7	Fortement stabilisant
n.30A>T	VUS	30	loop	stem_loop_I	145	145	-44	-44.3	-0.3	Neutre
n.28C>G	VUS	28	stem	stem_loop_I	145	145	-44	-41.4	+2.6	Fortement déstabilisant
n.27C>G	VUS	27	stem	stem_loop_I	145	145	-44	-41	+3	Fortement déstabilisant
n.26G>A	VUS	26	stem	stem_loop_I	145	145	-44	-40	+4	Fortement déstabilisant
n.23G>A	VUS	23	stem	stem_loop_I	145	145	-44	-40.2	+3.8	Fortement déstabilisant
n.8C>T	VUS	8	loop	5'_end	145	145	-44	-44	+0	Neutre
n.7_8insA	VUS	7	loop	5'_end	145	146	-44	-44	+0	Neutre
n.7G>C	VUS	7	loop	5'_end	145	145	-44	-44	+0	Neutre
n.6_7insA	VUS	6	loop	5'_end	145	146	-44	-44.2	-0.2	Neutre
n.3C>T	VUS	3	loop	5'_end	145	145	-44	-45.7	-1.7	Modérément stabilisant
n.111C>T	Conflicting	111	stem	stem_loop_IV	145	145	-44	-43.3	+0.7	Neutre
n.22C>T	Conflicting	22	stem	stem_loop_I	145	145	-44	-41.9	+2.1	Fortement déstabilisant

ANNEXE 2

Tableau de RNAcifold :

Variant	Classification	Position	Contexte	Région	ΔG WT (kcal/mol)	ΔG mut (kcal/mol)	$\Delta\Delta G$ RNAcifold	Interprétation
n.119A>G	Pathogenic	119	loop	stem_loop_IV	-82.6	-82.6	+0.0	Interaction U4/U6 non perturbée
n.77_78insT	Pathogenic	77	loop	stem_II_U4-U6_duplex	-82.6	-81.2	+1.4	Interaction U4/U6 modérément affaiblie
n.76C>T	Pathogenic	76	loop	stem_II_U4-U6_duplex	-82.6	-80.6	+2.0	Interaction U4/U6 modérément affaiblie
n.68dup	Pathogenic	68	loop	stem_II_U4-U6_duplex	-82.6	-82.4	+0.2	Interaction U4/U6 non perturbée
n.66A>G	Pathogenic	66	loop	stem_II_U4-U6_duplex	-82.6	-82.6	+0.0	Interaction U4/U6 non perturbée
n.65A>G	Pathogenic	65	loop	stem_II_U4-U6_duplex	-82.6	-82.8	-0.2	Interaction U4/U6 non perturbée
n.64_65insT	Pathogenic	64	stem	stem_II_U4-U6_duplex	-82.6	-82.4	+0.2	Interaction U4/U6 non perturbée
n.56T>C	Pathogenic	56	stem	stem_loop_I	-82.6	-81.7	+0.9	Interaction U4/U6 non perturbée
n.18_19insA	Pathogenic	18	stem	stem_loop_I	-82.6	-82.6	+0.0	Interaction U4/U6 non perturbée
n.69C>T	Pathogenic/LP	69	loop	stem_II_U4-U6_duplex	-82.6	-82.6	+0.0	Interaction U4/U6 non perturbée
n.67A>G	Pathogenic/LP	67	loop	stem_II_U4-U6_duplex	-82.6	-82.6	+0.0	Interaction U4/U6 non perturbée
n.64dup	Pathogenic/LP	64	stem	stem_II_U4-U6_duplex	-82.6	-84.6	-2.0	Interaction U4/U6 modérément renforcée
n.122T>G	Likely pathogenic	122	loop	3'_region	-82.6	-82.6	+0.0	Interaction U4/U6 non perturbée
n.118A>G	Likely pathogenic	118	loop	stem_loop_IV	-82.6	-85.0	-2.4	Interaction U4/U6 fortement renforcée
n.78A>C	Likely pathogenic	78	loop	stem_II_U4-U6_duplex	-82.6	-80.6	+2.0	Interaction U4/U6 modérément affaiblie
n.77_78insG	Likely pathogenic	77	loop	stem_II_U4-U6_duplex	-82.6	-81.2	+1.4	Interaction U4/U6 modérément affaiblie
n.76_77insT	Likely pathogenic	76	loop	stem_II_U4-U6_duplex	-82.6	-81.2	+1.4	Interaction U4/U6 modérément affaiblie
n.75C>G	Likely pathogenic	75	stem	stem_II_U4-U6_duplex	-82.6	-83.5	-0.9	Interaction U4/U6 non perturbée

ANNEXES

n.72_73del	Likely pathogenic	72	loop	stem_II_U4-U6_duplex	-82.6	-81.5	+1.1	Interaction U4/U6 modérément affaiblie
n.70T>C	Likely pathogenic	70	loop	stem_II_U4-U6_duplex	-82.6	-82.6	+0.0	Interaction U4/U6 non perturbée
n.68A>C	Likely pathogenic	68	loop	stem_II_U4-U6_duplex	-82.6	-82.6	+0.0	Interaction U4/U6 non perturbée
n.65_66insT	Likely pathogenic	65	loop	stem_II_U4-U6_duplex	-82.6	-83.1	-0.5	Interaction U4/U6 non perturbée
n.63T>C	Likely pathogenic	63	stem	stem_II_U4-U6_duplex	-82.6	-84.4	-1.8	Interaction U4/U6 modérément renforcée
n.62T>C	Likely pathogenic	62	stem	stem_II_U4-U6_duplex	-82.6	-80.7	+1.9	Interaction U4/U6 modérément affaiblie
n.46G>A	Likely pathogenic	46	stem	internal_loop	-82.6	-78.8	+3.8	Interaction U4/U6 fortement affaiblie
n.45G>C	Likely pathogenic	45	stem	internal_loop	-82.6	-79.0	+3.6	Interaction U4/U6 fortement affaiblie
n.32G>A	Likely pathogenic	32	loop	stem_loop_I	-82.6	-82.6	+0.0	Interaction U4/U6 non perturbée
n.11A>C	Likely pathogenic	11	stem	stem_loop_I	-82.6	-77.9	+4.7	Interaction U4/U6 fortement affaiblie
n.8C>G	Likely pathogenic	8	loop	5'_end	-82.6	-79.0	+3.6	Interaction U4/U6 fortement affaiblie
n.7G>A	Likely pathogenic	7	loop	5'_end	-82.6	-76.0	+6.6	Interaction U4/U6 fortement affaiblie
n.140G>A	VUS	140	stem	3'_region	-82.6	-76.2	+6.4	Interaction U4/U6 fortement affaiblie
n.129G>C	VUS	129	stem	3'_region	-82.6	-76.4	+6.2	Interaction U4/U6 fortement affaiblie
n.129G>A	VUS	129	stem	3'_region	-82.6	-76.4	+6.2	Interaction U4/U6 fortement affaiblie
n.127_128del	VUS	127	stem	3'_region	-82.6	-80.1	+2.5	Interaction U4/U6 fortement affaiblie
n.127C>T	VUS	127	stem	3'_region	-82.6	-80.2	+2.4	Interaction U4/U6 fortement affaiblie
n.125G>A	VUS	125	loop	3'_region	-82.6	-82.6	+0.0	Interaction U4/U6 non perturbée
n.121T>A	VUS	121	loop	3'_region	-82.6	-82.6	+0.0	Interaction U4/U6 non perturbée
n.105A>T	VUS	105	stem	stem_loop_IV	-82.6	-80.7	+1.9	Interaction U4/U6 modérément affaiblie
n.92C>G	VUS	92	stem	stem_loop_III	-82.6	-81.0	+1.6	Interaction U4/U6 modérément affaiblie
n.86C>T	VUS	86	stem	stem_loop_III	-82.6	-79.9	+2.7	Interaction U4/U6 fortement affaiblie
n.76del	VUS	76	loop	stem_II_U4-U6_duplex	-82.6	-83.8	-1.2	Interaction U4/U6 modérément renforcée
n.50G>T	VUS	50	stem	internal_loop	-82.6	-79.2	+3.4	Interaction U4/U6 fortement affaiblie
n.50G>C	VUS	50	stem	internal_loop	-82.6	-80.5	+2.1	Interaction U4/U6 fortement affaiblie

ANNEXES

n.50G>A	VUS	50	stem	internal_loop	-82.6	-79.3	+3.3	Interaction U4/U6 fortement affaiblie
n.45_46insT	VUS	45	stem	internal_loop	-82.6	-80.6	+2.0	Interaction U4/U6 modérément affaiblie
n.44A>T	VUS	44	loop	internal_loop	-82.6	-84.5	-1.9	Interaction U4/U6 modérément renforcée
n.43_44insT	VUS	43	loop	internal_loop	-82.6	-83.5	-0.9	Interaction U4/U6 non perturbée
n.42C>G	VUS	42	stem	internal_loop	-82.6	-84.4	-1.8	Interaction U4/U6 modérément renforcée
n.41C>T	VUS	41	stem	internal_loop	-82.6	-83.5	-0.9	Interaction U4/U6 non perturbée
n.41C>G	VUS	41	stem	internal_loop	-82.6	-81.3	+1.3	Interaction U4/U6 modérément affaiblie
n.33A>G	VUS	33	loop	stem_loop_I	-82.6	-85.3	-2.7	Interaction U4/U6 fortement renforcée
n.30A>T	VUS	30	loop	stem_loop_I	-82.6	-82.9	-0.3	Interaction U4/U6 non perturbée
n.28C>G	VUS	28	stem	stem_loop_I	-82.6	-80.0	+2.6	Interaction U4/U6 fortement affaiblie
n.27C>G	VUS	27	stem	stem_loop_I	-82.6	-79.6	+3.0	Interaction U4/U6 fortement affaiblie
n.26G>A	VUS	26	stem	stem_loop_I	-82.6	-79.0	+3.6	Interaction U4/U6 fortement affaiblie
n.23G>A	VUS	23	stem	stem_loop_I	-82.6	-78.6	+4.0	Interaction U4/U6 fortement affaiblie
n.8C>T	VUS	8	loop	5'_end	-82.6	-80.7	+1.9	Interaction U4/U6 modérément affaiblie
n.7_8insA	VUS	7	loop	5'_end	-82.6	-78.8	+3.8	Interaction U4/U6 fortement affaiblie
n.7G>C	VUS	7	loop	5'_end	-82.6	-75.9	+6.7	Interaction U4/U6 fortement affaiblie
n.6_7insA	VUS	6	loop	5'_end	-82.6	-78.8	+3.8	Interaction U4/U6 fortement affaiblie
n.3C>T	VUS	3	loop	5'_end	-82.6	-80.3	+2.3	Interaction U4/U6 fortement affaiblie
n.111C>T	Conflicting	111	stem	stem_loop_IV	-82.6	-81.9	+0.7	Interaction U4/U6 non perturbée
n.22C>T	Conflicting	22	stem	stem_loop_I	-82.6	-80.5	+2.1	Interaction U4/U6 fortement affaiblie

ANNEXE 3

Tableau de RNA Distance :

Variant	Classification	Position	Contexte	Région	RNAdistance (TED)	Interprétation
n.119A>G	Pathogenic	119	loop	stem_loop_IV	44	Réarrangement structural majeur
n.77_78insT	Pathogenic	77	loop	stem_II_U4-U6_duplex	1	Changement local faible
n.76C>T	Pathogenic	76	loop	stem_II_U4-U6_duplex	0	Aucun changement structural
n.68dup	Pathogenic	68	loop	stem_II_U4-U6_duplex	1	Changement local faible
n.66A>G	Pathogenic	66	loop	stem_II_U4-U6_duplex	36	Réarrangement structural majeur
n.65A>G	Pathogenic	65	loop	stem_II_U4-U6_duplex	28	Réarrangement structural majeur
n.64_65insT	Pathogenic	64	stem	stem_II_U4-U6_duplex	19	Réarrangement structural majeur
n.56T>C	Pathogenic	56	stem	stem_loop_I	10	Réarrangement structural majeur
n.18_19insA	Pathogenic	18	stem	stem_loop_I	7	Réarrangement structural modéré
n.69C>T	Pathogenic/LP	69	loop	stem_II_U4-U6_duplex	0	Aucun changement structural
n.67A>G	Pathogenic/LP	67	loop	stem_II_U4-U6_duplex	0	Aucun changement structural
n.64dup	Pathogenic/LP	64	stem	stem_II_U4-U6_duplex	27	Réarrangement structural majeur
n.122T>G	Likely pathogenic	122	loop	3'_region	20	Réarrangement structural majeur
n.118A>G	Likely pathogenic	118	loop	stem_loop_IV	4	Réarrangement structural modéré
n.78A>C	Likely pathogenic	78	loop	stem_II_U4-U6_duplex	0	Aucun changement structural
n.77_78insG	Likely pathogenic	77	loop	stem_II_U4-U6_duplex	15	Réarrangement structural majeur
n.76_77insT	Likely pathogenic	76	loop	stem_II_U4-U6_duplex	1	Changement local faible
n.75C>G	Likely pathogenic	75	stem	stem_II_U4-U6_duplex	20	Réarrangement structural majeur
n.72_73del	Likely pathogenic	72	loop	stem_II_U4-U6_duplex	2	Changement local faible
n.70T>C	Likely pathogenic	70	loop	stem_II_U4-U6_duplex	0	Aucun changement structural
n.68A>C	Likely pathogenic	68	loop	stem_II_U4-U6_duplex	0	Aucun changement structural
n.65_66insT	Likely pathogenic	65	loop	stem_II_U4-U6_duplex	1	Changement local faible
n.63T>C	Likely pathogenic	63	stem	stem_II_U4-U6_duplex	16	Réarrangement structural majeur
n.62T>C	Likely pathogenic	62	stem	stem_II_U4-U6_duplex	0	Aucun changement structural
n.46G>A	Likely pathogenic	46	stem	internal_loop	20	Réarrangement structural majeur

ANNEXES

n.45G>C	Likely pathogenic	45	stem	internal_loop	4	Réarrangement structural modéré
n.32G>A	Likely pathogenic	32	loop	stem_loop_I	0	Aucun changement structural
n.11A>C	Likely pathogenic	11	stem	stem_loop_I	36	Réarrangement structural majeur
n.8C>G	Likely pathogenic	8	loop	5'_end	20	Réarrangement structural majeur
n.7G>A	Likely pathogenic	7	loop	5'_end	0	Aucun changement structural
n.140G>A	VUS	140	stem	3'_region	4	Réarrangement structural modéré
n.129G>C	VUS	129	stem	3'_region	4	Réarrangement structural modéré
n.129G>A	VUS	129	stem	3'_region	4	Réarrangement structural modéré
n.127_128del	VUS	127	stem	3'_region	4	Réarrangement structural modéré
n.127C>T	VUS	127	stem	3'_region	0	Aucun changement structural
n.125G>A	VUS	125	loop	3'_region	0	Aucun changement structural
n.121T>A	VUS	121	loop	3'_region	0	Aucun changement structural
n.105A>T	VUS	105	stem	stem_loop_IV	4	Réarrangement structural modéré
n.92C>G	VUS	92	stem	stem_loop_III	4	Réarrangement structural modéré
n.86C>T	VUS	86	stem	stem_loop_III	0	Aucun changement structural
n.76del	VUS	76	loop	stem_II_U4-U6_duplex	1	Changement local faible
n.50G>T	VUS	50	stem	internal_loop	12	Réarrangement structural majeur
n.50G>C	VUS	50	stem	internal_loop	16	Réarrangement structural majeur
n.50G>A	VUS	50	stem	internal_loop	16	Réarrangement structural majeur
n.45_46insT	VUS	45	stem	internal_loop	7	Réarrangement structural modéré
n.44A>T	VUS	44	loop	internal_loop	6	Réarrangement structural modéré
n.43_44insT	VUS	43	loop	internal_loop	11	Réarrangement structural majeur
n.42C>G	VUS	42	stem	internal_loop	6	Réarrangement structural modéré
n.41C>T	VUS	41	stem	internal_loop	8	Réarrangement structural modéré
n.41C>G	VUS	41	stem	internal_loop	8	Réarrangement structural modéré
n.33A>G	VUS	33	loop	stem_loop_I	8	Réarrangement structural modéré
n.30A>T	VUS	30	loop	stem_loop_I	12	Réarrangement structural majeur
n.28C>G	VUS	28	stem	stem_loop_I	4	Réarrangement structural modéré
n.27C>G	VUS	27	stem	stem_loop_I	4	Réarrangement structural modéré

ANNEXES

n.26G>A	VUS	26	stem	stem_loop_I	4	Réarrangement structural modéré
n.23G>A	VUS	23	stem	stem_loop_I	16	Réarrangement structural majeur
n.8C>T	VUS	8	loop	5'_end	16	Réarrangement structural majeur
n.7_8insA	VUS	7	loop	5'_end	1	Changement local faible
n.7G>C	VUS	7	loop	5'_end	0	Aucun changement structural
n.6_7insA	VUS	6	loop	5'_end	7	Réarrangement structural modéré
n.3C>T	VUS	3	loop	5'_end	24	Réarrangement structural majeur
n.111C>T	Conflicting	111	stem	stem_loop_IV	8	Réarrangement structural modéré
n.22C>T	Conflicting	22	stem	stem_loop_I	0	Aucun changement structural

Références bibliographiques

A

- Almahmoud, O. H., & Abushaikha, L. (2023). Prevalence and risk factors of developmental disabilities among preschool children in the Arab world : A narrative literature review. *Child Health Nursing Research*, 29(2), 101–110. <https://doi.org/10.4094/chnr.2023.29.2.101>
- Antczak, M., Popenda, M., Zok, T., Zurkowski, M., Adamiak, R. W., & Szachniuk, M. (2018). New algorithms to represent complex pseudoknotted RNA structures in dot-bracket notation. *Bioinformatics*, 34(8), 1304–1312. <https://doi.org/10.1093/bioinformatics/btx783>

B

- Barbour, K., Bainbridge, M. N., Wigby, K., Besterman, A. D., Chuang, N. A., Tobin, L. E., Del Campo, M., Lenberg, J., Bird, L. M., & Friedman, J. (2024). The face and features of RNU4-2 : A new, common, recognizable, yet hidden neurodevelopmental disorder. *Pediatric Neurology*, 161, 188–193. <https://doi.org/10.1016/j.pediatrneurol.2024.09.015>
- Barros, S. P., & Offenbacher, S. (2009). Epigenetics : Connecting environment and genotype to phenotype and disease. *Journal of Dental Research*, 88(5), 400–408. <https://doi.org/10.1177/0022034509335868>
- Burgunder, J.-M. (2023). Mechanisms underlying phenotypic variation in neurogenetic disorders. *Nature Reviews. Neurology*, 19(6), 363–370. <https://doi.org/10.1038/s41582-023-00811-4>
- Burns, V. F., & Radford, E. J. (2024). ReNU syndrome – a newly discovered prevalent neurodevelopmental disorder. *Trends in Genetics*, 40(11), 914–916. <https://doi.org/10.1016/j.tig.2024.09.005>

C

- Chen, Y., Dawes, R., Kim, H. C., Ljungdahl, A., Stenton, S. L., Walker, S., Lord, J., Lemire, G., Martin-Geary, A. C., Ganesh, V. S., Ma, J., Ellingford, J. M., Delage, E., D'Souza, E. N., Dong, S., Adams, D. R., Allan, K., Bakshi, M., Baldwin, E. E., ... Whiffin, N. (2024). De novo variants in the RNU4-2 snRNA cause a frequent neurodevelopmental syndrome. *Nature*, 632(8026), 832–840. <https://doi.org/10.1038/s41586-024-07773-7>
- Cooper, G. M. (2000). *The cell : A molecular approach* (2e éd.). Sinauer Associates. <https://www.ncbi.nlm.nih.gov/books/NBK9864/>

D

- De Jonghe, B., Kim, H. C., Adedeji, A., Leitão, E., Dawes, R., Chen, Y., Blakes, A. J., Simons, C., Rius, R., Alvi, J. R., Amblard, F., Austin-Tse, C., Baer, S., Balton, E. V., Blanc, P., Calame, D. G., Coutton, C., Cunningham, C. A., Dargie, N., ... Findlay, G. M. (2026). Saturation genome editing of RNU4-2 reveals distinct dominant and recessive neurodevelopmental disorders. *Nature*. <https://doi.org/10.1038/s41586-026-10334-9>

E

- Ebrahimi Meimand, S., Amiri, Z., Shobeiri, P., Malekpour, M.-R., Saeedi Moghaddam, S., Ghanbari, A., Tehrani, Y. S., Shokri Varniab, Z., Pourabhari Langroudi, A., Sohrabi, H., Foroutan Mehr, E., Rezaei, N., Moradi-Lakeh, M., H Mokdad, A., & Larijani, B. (2023). Burden of autism spectrum disorders in North Africa

and Middle East from 1990 to 2019 : A systematic analysis for the Global Burden of Disease Study 2019. *Brain and Behavior*, 13(7), e3067. <https://doi.org/10.1002/brb3.3067>

Eliáš, R., & Hoksza, D. (2016). RNA secondary structure visualization using tree edit distance. *International Journal of Bioscience, Biochemistry and Bioinformatics*, 6, 9–17. <https://doi.org/10.17706/ijbbb.2016.6.1.9-17>

ENCODE Project Consortium. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, 489(7414), 57–74. <https://doi.org/10.1038/nature11247>

G

Greene, D., Thys, C., Berry, I. R., Jarvis, J., Ortibus, E., Mumford, A. D., Freson, K., & Turro, E. (2024). Mutations in the U4 snRNA gene RNU4-2 cause one of the most prevalent monogenic neurodevelopmental disorders. *Nature Medicine*, 30(8), 2165–2169. <https://doi.org/10.1038/s41591-024-03085-5>

H

Halvorsen, M., Martin, J. S., Broadaway, S., & Laederach, A. (2010). Disease-associated mutations that alter the RNA structural ensemble. *PLOS Genetics*, 6(8), e1001074. <https://doi.org/10.1371/journal.pgen.1001074>

Hayashi, Y., Kajiwar, K., Mizuno, S., Okamoto, N., Chan, M. Y., Goto, T., Hayakawa, S., Kato, M., Kim, C. A., Lev, D., Moey, L. H., Taylor, J., Gregersen, N., Nezer-Kaner, I., Keng, W. T., Okada, S., Osaka, H., Sagie, T., Sakai, Y., ... Matsumoto, N. (2026). Monoallelic and biallelic RNU4-2 variants in neurodevelopmental disorders. *Journal of Human Genetics*, 71(6), 327–334. <https://doi.org/10.1038/s10038-025-01442-x>

I

Ivry, T., Michal, S., Avihoo, A., Sapiro, G., & Barash, D. (2009). An image processing approach to computing distances between RNA secondary structures dot plots. *Algorithms for Molecular Biology*, 4(1), 4. <https://doi.org/10.1186/1748-7188-4-4>

J

Jiang, T., Lin, G., Ma, B., & Zhang, K. (2002). A general edit distance between RNA structures. *Journal of Computational Biology*, 9(2), 371–388. <https://doi.org/10.1089/10665270252935511>

Jonghe, J. D., Kim, H. C., Adedeji, A., Leitão, E., Dawes, R., Chen, Y., Blakes, A. J., Simons, C., Rius, R., Alvi, J. R., Amblard, F., Austin-Tse, C., Baer, S., Balton, E. V., Blanc, P., Calame, D. G., Coutton, C., Cunningham, C. A., Dargie, N., ... Findlay, G. M. (2025). Saturation genome editing of RNU4-2 reveals distinct dominant and recessive neurodevelopmental disorders [Preprint]. *medRxiv*. <https://doi.org/10.1101/2025.04.08.25325442>

K

Kundakovic, M., & Champagne, F. A. (2015). Early-life experience, epigenetics, and the developing brain. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, 40(1), 141–153. <https://doi.org/10.1038/npp.2014.140>

L

Lorenz, R., Bernhart, S. H., Höner zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P. F., & Hofacker, I. L. (2011). ViennaRNA Package 2.0. *Algorithms for Molecular Biology*, 6(1), 26. <https://doi.org/10.1186/1748-7188-6-26>

Loumi, O., & Andres, C. R. (2025). Contribution of an ambidirectional cohort study on the epidemiology of 186 autism spectrum disorder cases in an Algerian population. *International Journal of Developmental Neuroscience*, 85(5), e70036. <https://doi.org/10.1002/jdn.70036>

M

Matera, A. G., Terns, R. M., & Terns, M. P. (2007). Non-coding RNAs : Lessons from the small nuclear and small nucleolar RNAs. *Nature Reviews. Molecular Cell Biology*, 8(3), 209–220. <https://doi.org/10.1038/nrm2124>

Mathews, D. H., Sabina, J., Zuker, M., & Turner, D. H. (1999). Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *Journal of Molecular Biology*, 288(5), 911–940. <https://doi.org/10.1006/jmbi.1999.2700>

Mattick, J. S., Amaral, P. P., Carninci, P., Carpenter, S., Chang, H. Y., Chen, L.-L., Chen, R., Dean, C., Dinger, M. E., Fitzgerald, K. A., Gingeras, T. R., Guttman, M., Hirose, T., Huarte, M., Johnson, R., Kanduri, C., Kapranov, P., Lawrence, J. B., Lee, J. T., ... Wu, M. (2023). Long non-coding RNAs : Definitions, functions, challenges and recommendations. *Nature Reviews. Molecular Cell Biology*, 24(6), 430–447. <https://doi.org/10.1038/s41580-022-00566-8>

Mehari, A., & Zeleke, S. (2026). Prevalence, trends in, and factors associated with neurodevelopmental disorders (NDDs) in low- and low-middle income countries (LLMICs) : A systematic review and meta-analysis. *Research in Developmental Disabilities*, 168, 105186. <https://doi.org/10.1016/j.ridd.2025.105186>

Miao, Z., & Westhof, E. (2017). RNA structure : Advances and assessment of 3D structure prediction. *Annual Review of Biophysics*, 46, 483–503. <https://doi.org/10.1146/annurev-biophys-070816-034125>

Michetti, C., Falace, A., Benfenati, F., & Fassio, A. (2022). Synaptic genes and neurodevelopmental disorders : From molecular mechanisms to developmental strategies of behavioral testing. *Neurobiology of Disease*, 173, 105856. <https://doi.org/10.1016/j.nbd.2022.105856>

N

Nava, C., Cogne, B., Santini, A., Leitão, E., Lecoquierre, F., Chen, Y., Stenton, S. L., Besnard, T., Heide, S., Baer, S., Jakhar, A., Neuser, S., Keren, B., Faudet, A., Forlani, S., Faucher, M., Uguen, K., Platzer, K., Afenjar, A., ... Depienne, C. (2025). Dominant variants in major spliceosome U4 and U5 small nuclear RNA genes cause neurodevelopmental disorders through splicing disruption. *Nature Genetics*, 57(6), 1374–1388. <https://doi.org/10.1038/s41588-025-02184-4>

Nearest Neighbor Database. (2024). Consulté le 4 mai 2024, à l'adresse <https://ma.urmc.rochester.edu/NNDB/>

Nguyen, T. H. D., Galej, W. P., Fica, S. M., Lin, P.-C., Newman, A. J., & Nagai, K. (2016). CryoEM structures of two spliceosomal complexes : Starter and dessert at the spliceosome feast. *Current Opinion in Structural Biology*, 36, 48–57. <https://doi.org/10.1016/j.sbi.2015.12.005>

Nielsen, K. H., & Staley, J. P. (2012). Spliceosome activation : U4 is the path, stem I is the goal, and Prp8 is the keeper. *Genes & Development*, 26(22), 2461–2467. <https://doi.org/10.1101/gad.207514.112>

Nottrott, S., Hartmuth, K., Fabrizio, P., Urlaub, H., Vidovic, I., Ficner, R., & Lührmann, R. (1999). Functional interaction of a novel 15.5kD [U4/U6·U5] tri-snRNP protein with the 5' stem-loop of U4 snRNA. *The EMBO Journal*, 18(21), 6119–6133. <https://doi.org/10.1093/emboj/18.21.6119>

O

Okamoto, N., Nishi, E., Hasegawa, Y., Higuchi, S., Kuki, I., Yanagi, K., Kaname, T., Uchiyama, Y., & Matsumoto, N. (2025). A clinical study of nine patients with ReNU syndrome. *American Journal of Medical Genetics. Part A*, 197(11), e64151. <https://doi.org/10.1002/ajmg.a.64151>

P

Peñafiel-Sam, J., Valenzuela, I., & Peris, P. (2025). Severe osteoporosis in an adult subject with RNU4-2 gene mutation. *Calcified Tissue International*, 116(1), 1. <https://doi.org/10.1007/s00223-025-01351-3>

R

Redon, R., Ishikawa, S., Fitch, K. R., Feuk, L., Perry, G. H., Andrews, T. D., Fiegler, H., Shapero, M. H., Carson, A. R., Chen, W., Cho, E. K., Dallaire, S., Freeman, J. L., González, J. R., Gratacòs, M., Huang, J., Kalaitzopoulos, D., Komura, D., MacDonald, J. R., ... Hurles, M. E. (2006). Global variation in copy number in the human genome. *Nature*, 444(7118), 444–454. <https://doi.org/10.1038/nature05329>

Romano, J., Kromrey, J. D., Coraggio, J., & Skowronek, J. (2006). Appropriate statistics for ordinal level data : Should we really be using t-test and Cohen's d for evaluating group differences on the NSSE and other surveys? *Annual Meeting of the Florida Association of Institutional Research*, 1–33.

Rosenblum, J., Beysen, D., Jansen, A. C., De Rademaeker, M., Reyniers, E., Janssens, K., & Meuwissen, M. (2025). RNU4-2-related neurodevelopmental disorder is associated with a recognisable facial gestalt. *Clinical Genetics*, 107(1), 104–112. <https://doi.org/10.1111/cge.14628>

S

Sander, B., Golas, M. M., Makarov, E. M., Brahms, H., Kastner, B., Lührmann, R., & Stark, H. (2006). Organization of core spliceosomal components U5 snRNA loop I and U4/U6 di-snRNP within U4/U6·U5 tri-snRNP as revealed by electron cryomicroscopy. *Molecular Cell*, 24(2), 267–278. <https://doi.org/10.1016/j.molcel.2006.08.021>

U

UTC, G. (2021/2026). GEC-git/AptaMat [Logiciel Python]. GitHub. <https://github.com/GEC-git/AptaMat>

V

Valenzuela, I., Codina-Solà, M., Vazquez, E., Cueto-González, A., Leno-Colorado, J., Lasa-Aranzasti, A., Trujillano, L., Masotto, B., Masas, M., Escobar, M., García-Arumí, E., & Tizzano, E. F. (2024). Deep phenotyping of 11 individuals with pathogenic variants in RNU4-2 reveals a clinically recognizable syndrome. *Genetics in Medicine*, 26(12), 101288. <https://doi.org/10.1016/j.gim.2024.101288>

Vidovic, I., Nottrott, S., Hartmuth, K., Lührmann, R., & Ficner, R. (2000). Crystal structure of the spliceosomal 15.5kD protein bound to a U4 snRNA fragment. *Molecular Cell*, 6(6), 1331–1342. [https://doi.org/10.1016/s1097-2765\(00\)00131-3](https://doi.org/10.1016/s1097-2765(00)00131-3)

W

Wahl, M. C., Will, C. L., & Lührmann, R. (2009). The spliceosome : Design principles of a dynamic RNP machine. *Cell*, 136(4), 701–718. <https://doi.org/10.1016/j.cell.2009.02.009>

Washbourne, P. (2015). Synapse assembly and neurodevelopmental disorders. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, 40(1), 4–15. <https://doi.org/10.1038/npp.2014.163>

Will, C. L., & Lührmann, R. (2011). Spliceosome structure and function. *Cold Spring Harbor Perspectives in Biology*, 3(7), a003707. <https://doi.org/10.1101/cshperspect.a003707>

Z

Zampetaki, A., Albrecht, A., & Steinhofel, K. (2018). Long non-coding RNA structure and function : Is there a link? *Frontiers in Physiology*, 9, 1201. <https://doi.org/10.3389/fphys.2018.01201>

Zuker, M., & Stiegler, P. (1981). Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Research*, 9(1), 133–148. <https://doi.org/10.1093/nar/9.1.133>

Sites web et ressources en ligne

Bases de données génomiques

ClinVar. (2025). ClinVar : Public archive of relationships among sequence variation and human phenotype. National Center for Biotechnology Information (NCBI). Consulté en mai 2025, à l'adresse <https://www.ncbi.nlm.nih.gov/clinvar/> Requête effectuée : gène RNU4-2, variants classés Pathogenic, Likely pathogenic et Uncertain significance.

Ensembl Genome Browser. (2024). Release 115. European Bioinformatics Institute (EMBL-EBI) & Wellcome Sanger Institute. Consulté en 2025, à l'adresse <https://www.ensembl.org> Séquences extraites : RNU4-2 (ENST00000365668.2, chr12:120 291 759–120 291 903, GRCh38) et RNU6-1 (ENST00000383898.1, chr15:67 839 939–67 840 045, GRCh38).

Outils bioinformatiques

GEC-git. (2021/2026). AptaMat Matrix-based RNA structure comparison tool [Logiciel Python]. GitHub. <https://github.com/GEC-git/AptaMat>

Lorenz, R., Bernhart, S. H., Hofacker, I. L., & al. (2011). ViennaRNA Web Services RNAfold : Prédiction de la structure secondaire à énergie libre minimale (MFE). Institut für Theoretische Chemie, Université de Vienne. Consulté en 2025, à l'adresse <https://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi> — Paramètres utilisés : température 37 °C, paramètres par défaut, sans matrice de partition.

Lorenz, R., Bernhart, S. H., Hofacker, I. L., & al. (2011). ViennaRNA Web Services — RNAcifold : Prédiction de la structure hybride et de l'énergie d'interaction U4/U6. Institut für Theoretische Chemie, Université de Vienne. Consulté en 2025, à l'adresse <https://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAcifold.cgi> Format d'entrée : U4_variant&U6_WT ; température 37 °C, paramètres par défaut.

Lorenz, R., Bernhart, S. H., Hofacker, I. L., & al. (2011). ViennaRNA Web Services RNAdistance : Calcul de la distance structurale (Tree Edit Distance). Institut für Theoretische Chemie, Université de Vienne. Consulté en 2025, à l'adresse <https://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAdistance.cgi>

Nearest Neighbor Database (NNDB). (2024). Thermodynamic parameters for RNA secondary structure prediction (Turner rules). Université de Rochester. Consulté le 4 mai 2024, à l'adresse <https://ma.urmc.rochester.edu/NNDB/>

Environnements de calcul

Google LLC. (2025). Google Colaboratory (Colab) Environnement Python en ligne utilisé pour les analyses statistiques (SciPy, statsmodels). Consulté en 2025, à l'adresse <https://colab.research.google.com>

Documentation ViennaRNA

ViennaRNA Package Documentation. (s. d.). Dot-bracket notation of secondary structures. Consulté le 17 mai 2026, à l'adresse <https://viennarna.readthedocs.io/en/latest/utis/struct/dotbracket.html>

ViennaRNA Package Documentation. (s. d.). RNA structure notations (RNAlib 2.4.14). Consulté le 17 mai 2026, à l'adresse https://gensoft.pasteur.fr/docs/ViennaRNA/2.4.14/rna_structure_notations.html