HAL
open science

Approche intégrative de l'analyse des variations faux-sens dans le syndrome MED13L, multiples mécanismes impliqués

Thomas Smol

To cite this version:

Thomas Smol. Approche intégrative de l'analyse des variations faux-sens dans le syndrome MED13L, multiples mécanismes impliqués. Médecine humaine et pathologie. Université de Lille, 2021. Français. NNT : 2021LILUS066 . tel-03714135

HAL Id: tel-03714135

https://theses.hal.science/tel-03714135

Submitted on 5 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE LILLE

École Doctorale Biologie-Santé - ED 446

THESE DE DOCTORAT

En vue de l'obtention du grade de

DOCTEUR DE L'UNIVERSITE

Discipline : Génétique

« APPROCHE INTEGRATIVE DE L'ANALYSE DES VARIATIONS FAUX-SENS DANS LE SYNDROME *MED13L, MULTIPLES MECANISMES IMPLIQUES* »

Présentée et soutenue publiquement par

Thomas Smol

A Lille, le 13 décembre 2021

Thèse dirigée par

le Pr Sylvie MANOUVRIER-HANU et le Dr Jamal GHOUMID

ULR7364 RADEME

Soutenance devant le jury composé de :

Pr Andrée DELAHAYE-DURIEZ	Université de Paris 13	**Rapporteure**
Dr Jean MULLER	Université de Strasbourg	**Rapporteur**
Pr Sylvie NGUYEN THE TICH	Université de Lille	**Examinatrice, présidente du jury**
Pr Gaétan LESCA	Université de Lyon	**Examinateur**
Dr Gaël NICOLAS	Université de Rouen	**Examinateur**
Pr Sylvie MANOUVRIER-HANU	Université de Lille	**Directrice de thèse**
Dr Jamal GHOUMID	Université de Lille	**Co-directeur de thèse**

RESUME

« APPROCHE INTEGRATIVE DE L'ANALYSE DES VARIATIONS FAUX-SENS DANS LE SYNDROME *MED13L, MULTIPLES MECANISMES IMPLIQUES* »

Le syndrome *MED13L* est une cause de déficience intellectuelle due aux variants du gène. Initialement seuls les variants tronquants étaient connues comme responsables, mais le recours grandissant au séquençage non ciblé a conduit à l'identification d'un nombre croissant de variations faux-sens candidates.

Par une collaboration internationale, nous avons rapporté 36 nouveaux individus, atteints de la pathologie, dont 9 porteurs d'un variant faux-sens. Les variants faux-sens sont regroupés majoritairement dans les exons 15-17 et 25-31 et les patients porteurs ce ces variants avaient un phénotype plus sévère.

Afin de comprendre les causes de cette corrélation génotype-phénotype, cinq variants faux-sens pathogènes provoquant soit un phénotype sévère (p.Pro866Leu, p.Pro869Ser, et p.Cys1131Tyr) soit un phénotype typique (p.Gly1899Arg et p.Thr2162Met) ont été sélectionnées et étudiées fonctionnellement. Parallèlement, une comparaison des paramètres *in silico* associés aux faux-sens pathogènes précédemment publiés ou rapportées dans les bases de données témoins a été réalisée.

Des différences significatives ont été identifiées entre les faux-sens pathogènes et neutres pour les scores de conservation, les différences physico-chimiques entre acides aminés suivant le score de Grantham, la conservation de la séquence dans la protéine MED13 et la proximité de sites de phosphorylation. Fonctionnement, les variations p.Pro866Leu et p.Pro869Ser étaient susceptibles d'induire une hyper-phosphorylation du domaine $MED13L^{823-930}$. Les résidus Cys1131 et Thr2162 pourraient impacter la conformation de MED13L avec une modification de la structure 3D. Une dégradation accrue médiée par le protéasome ainsi qu'une relocalisation cytoplasmique et une modification de l'intégration dans le module kinase CDK8 ont été observées uniquement pour le faux-sens p.Gly1899Arg.

Cette étude confirme que les variations faux-sens de *MED13L* sont une cause du Syndrome MED13L avec des particularités cliniques et moléculaires. Les faux-sens de *MED13L* associées aux phénotypes typiques (p.Gly1899Arg et p.Thr2162Met) induisent probablement une perte de fonction, tandis que les variations associées aux phénotypes sévères provoqueraient probablement un effet dominant-négatif.

MOTS-CLES : MED13L ; COMPLEXE MEDIATEUR ; DEFICIENCE INTELLECTUELLE

ABSTRACT

« INTEGRATIVE APPROACH TO DEFINE PATHOGENIC MED13L MISSENSE VARIATIONS, INVOLVEMENT OF MULTIPLE MECHANISMS »

Pathogenic heterozygous *MED13L* variants cause a rare neurodevelopmental disorder characterized by a moderate-to-severe intellectual disability. Whole exome or genome sequencing in patients with non-specific neurodevelopmental disorders leads to the identification of an increasing number of *MED13L* candidate missense variations.

Through international collaboration between centers of expertise for developmental anomalies, we report on 36 patients with *MED13L* molecular anomaly to expand the phenotype. Five *MED13L* pathogenic missense variants causing either a severe phenotype (p.Pro866Leu, p.Pro869Ser, and p.Cys1131Tyr) or a typical phenotype (p.Gly1899Arg and p.Thr2162Met) were selected and functionally studied. Along with the functional analysis, a comparison of *in silico* parameters associated either with previously published pathogenic missense or neutral missenses was performed.

Missense variants clustered in two mutation hot-spots, exons 15–17 and 25–31. We found that patients carrying missense variations showed a more severe phenotype with frequent epilepsy. Significant differences were identified between pathogenic and neutral missenses for global conservation, physico-chemical difference between amino acids by Grantham scores, conservation of sequence in MED13 protein, and proximity to phosphorylation sites. Through functional evaluations, the p.Pro866Leu and p.Pro869Ser variants were likely to induce a hyper-phosphorylated status of the $MED13L^{823-930}$ domain. Residues Cys1131 and Thr2162 are predicted to be critical for MED13L proper folding, and substitutions are likely to induce modification of the 3D structure. Increased proteasome-mediated degradation was observed only with the p.Gly1899Arg variant, as well as cytoplasmic relocalization and reduced integration into kinase module and core Mediator complex.

This study shows evidence of missense variations in *MED13L* as a cause for *MED13L*-related intellectual disability and improves the clinical and molecular delineations of the condition. *MED13L* variants associated with typical phenotypes (p.Gly1899Arg and p.Thr2162Met) probably induce a loss of function, while variants associated with severe phenotypes probably lead to a dominant-negative effect.

KEY-WORDS : MED13L ; MEDIATOR COMPLEX ; INTELLECTUAL DISABILITY

ENCADREMENT

ULR7364 RADEME

Maladies rares du développement et du métabolisme

Dirigée par le Pr Sylvie MANOUVRIER-HANU, PU-PH

et

le Dr Jamal GHOUMID, MCU-PH

Faculté de Médecine – Pôle Recherche

1, Place de Verduin

F59000 LILLE

REMERCIEMENTS

Je remercie chaleureusement l'ensemble des personnes ayant contribué à la réalisation de ces travaux et qui m'ont accompagné au long de ces années de thèse.

Merci aux membres du jury, **Andrée Delahaye-Duriez** et **Jean Muller** d'avoir accepté d'être rapporteurs de mon travail, **Sylvie Nguyen The Tich**, **Gaétan Lesca** et **Gaël Nicolas** d'avoir accepté d'être examinateurs. Veuillez trouver ici l'expression de mon plus profond respect ainsi que le témoignage de ma reconnaissance pour avoir consacré du temps à juger ce travail.

A mes directeurs de thèse, **Sylvie Manouvrier** et **Jamal Ghoumid**.

Sylvie, un grand merci de m'avoir permis de réaliser cette thèse dans l'unité RADEME, d'avoir accepté son encadrement et de m'avoir soutenu dans mes projets. Soyez assurée de ma profonde reconnaissance.

Jamal, merci de m'avoir proposé ce sujet pour ta première direction de thèse. Je te remercie pour ta disponibilité, ton encadrement, et tes idées toujours enrichissantes ! Merci de ton soutien pendant ces années de thèse.

J'adresse mes remerciements à mes collègues de l'EA7364 RADEME qui suivent les péripéties de cette thèse depuis 2016 : **Anne-Sophie Jourdain**, **Clémence Vanlerberghe**, **Fabienne Escande**, **Florence Petit**, **Frédéric Frenois**, **Malika Balduyck**, **Perrine Brunelle**, **Simon Boussion** et **William Dufour**. Merci pour vos conseils, votre soutien et nos échanges scientifiques conviviaux.

Je remercie particulièrement **Fréderic Frenois**, de l'EA7364 RADEME, très impliqué dans ce projet, et sans qui de nombreuses étapes d'analyses fonctionnelles n'auraient pu être menées à terme. Merci également à **Samra Holuigue** pour son aide sur la culture cellulaire et dans les études de dégradation de MED13L.

Je remercie vivement les étudiants de M1 et de licence qui se sont succédés dans l'EA7364 et ont participé aux différentes étapes des protocoles, dont la pénible étape de mutagénèse dirigée, **Marilyne**, **Emma**, **Rim**, **Camille** et **Fréderic**. Merci à **Morgane Billotte**, dont le travail de M2 encadré par Jamal, a posé les bases de réflexion sur les explorations fonctionnelles.

Merci à **Alexis Verger** et **Didier Monté** pour leur soutien technique majeur sur les études de co-immunoprécipitation et leurs retours d'expériences sur le complexe médiateur. J'espère que ce travail se poursuivra par d'autres collaborations fructueuses.

Merci à l'équipe des « biopuces » **Caroline**, **Alexis**, **Delphine**, **Heidi** et **Pauline** pour votre réactivité pour les designs et votre aide pour les séquençages.

Merci à **Damien Sanlaville**, **Marie-Pierre Buisine** et **Martin Figeac**, de m'avoir accompagné, écouté et guidé lors de nos échanges toujours très constructifs sur l'avancée de la thèse.

Merci à mes collègues de l'institut de génétique médicale, au Pr **Catherine Roche-Lestienne** qui m'a accueilli et responsabilisé au sein de l'institut, et aux Dr **Agnès Daudignon**, **Elise Boudry**, **Sonia Bouquillon**, **Mélanie Rama** et **Perrine Brunelle**. Merci particulièrement aux Dr **Mélanie Rama** et **Perrine Brunelle** pour la gestion et le suivi des dossiers lors de mon temps universitaire. Merci également à **Clémence Carré** et **Fadima Touré** pour la prise en charge et la préparation des consultations de cytogénétique lors de mes absences.

Merci également aux très nombreux collaborateurs cliniciens et biologistes moléculaires que j'ai côtoyés de très près à la clinique de génétique, **Anne Dieux**, **Catherine Vincent-Delorme**, **Cindy Colson**, **Clémence Vanlerberghe** (re), **Florence Petit** (re), **Luisa Marsili**, **Odile Boute**, **Roseline Caumes**, et de beaucoup plus loin grâce aux collaborations.

Je tiens également à remercier les familles de patients, les patients, l'association Syndrome MED13L, et GenIDA (**Jean Louis Mandel**, **Florent Colin**, **Pauline Burger**) qui suivent les avancées de ces travaux depuis plusieurs années.

Je terminerai par des remerciements plus personnels à mes proches toujours à mes côtés quelles que soient les épreuves.

AVANT-PROPOS

Plusieurs articles originaux publiés ou en cours de publication en premier auteur dans des revues scientifiques internationales à comité de lecture structurent ce travail de thèse :

- **Smol T, *et al*. *MED13L*-related intellectual disability: involvement of missense variants and delineation of the phenotype. Neurogenetics. 2018 May;19(2):93-103.**
- **Smol T, *et al*. Performance of meta-predictors for the classification of *MED13L* missense variations, implication of raw parameters. Eur J Med Genet. 2021.**
- **Smol T, *et al*. *MED13L* missense variations cause multifaceted functional consequences underlying severe phenotypes. In revision in Molecular Psychiatry.**

Ces articles ont également fait l'objet de présentations lors de communications orales et/ou affichées en congrès scientifiques nationaux et internationaux :

- **Phénotype associé au syndrome MED13L : étude d'une cohorte de 30 patients et implication des variations faux-sens. 9e Assises de Génétique Humaine et Médicales, Nantes – France, 2018 (Communication orale) ;**
- **Functional analysis of *MED13L* missense variants. American Society of Human Genetics Congress, 2018, San Diego – Etats-Unis (Communication affichée) ;**
- **Evaluation of 15 *in silico* prediction tools for the classification of *MED13L* missense variation. 18e Journées Ouvertes en Biologie, Informatique et Mathématiques, Lille – France, 2017 (Communication affichée) ;**
- **Genotype and phenotype features of 22 patients with intellectual disability caused by *MED13L* variations. European Human Genetics Conference, 2017, Copenhague - Danemark (Communication affichée) ;**

La thématique du syndrome MED13L a également été abordée pour d'autres communications, dont les résultats ne sont pas présentés dans ce travail.

- **Overview of *MED13L* Cohort in GENIDA database. 5TH Genida Scientific Advisory Board – OnLine, 2020 (Communication orale) ;**
- **Implication des familles de patients dans la description du syndrome *MED13L*, apport de la base de données GENIDA pour la caractérisation d'un syndrome rare. 10èm Assises de Génétique Humaine et Médicales, Tours – France, 2020 (Communication affichée) ;**
- **Syndrome MED13L – une pathologie du complexe médiateur. 2e réunion de l'association française « SYNDROME MED13L », Paris – France, 2019 (Communication orale).**

TABLE DES MATIÈRES

TABLES & FIGURES ____ *15*

ABBRÉVIATIONS ____ *19*

LEXIQUE DES GÈNES – NOMENCLATURE HUGO ____ *20*

1. INTRODUCTION ____ **21**

1.1. Le complexe médiateur ____ **22**

1.1.1. Organisation du complexe ____ **22**

1.1.1.1. Corps du complexe ____ **22**

1.1.1.2. Module kinase CDK8 ____ **23**

1.1.2. Rôles du complexe médiateur ____ **25**

1.1.2.1. Régulation de la transcription ____ **25**

1.1.2.2. Conformation chromatinienne ____ **26**

1.1.2.3. Réparation de l'ADN ____ **27**

1.2. Structure et rôles de MED13L ____ **28**

1.2.1. Localisation ____ **28**

1.2.2. Structure protéique ____ **29**

1.2.2.1. Domaines et régions connues ____ **29**

1.2.3. Régulation ____ **32**

1.2.4. Interactions ____ **34**

1.3. Pathologies associées au complexe médiateur 35

1.3.1. Module kinase CDK8 36

1.3.1.1. Déficience intellectuelle liée à MED13 37

1.3.1.2. Déficiences intellectuelles associées à *MED12* 38

1.3.1.3. Déficience intellectuelle associée à *MED12L* 38

1.3.1.4. Déficience intellectuelle en lien avec *CDK8* 39

1.3.1.5. Déficience intellectuelle liée à *CDK19* 39

1.3.2. Head Module 40

1.3.2.1. Microcéphalie postnatale progressive en lien avec *MED17* 40

1.3.2.2. Troubles neuro-évolutifs avec mouvements anormaux en lien avec *MED20* 40

1.3.3. Tail Module 41

1.3.3.1. Déficience intellectuelle en lien avec *MED23* 41

1.3.3.2. Syndrome Basel-Vanagaite-Smirin-Yosef associé à *MED25* 41

1.3.3.3. Troubles du neuro-développement en lien avec *MED27* 42

1.3.3.4. Spectre des pathologies du *tail module* 42

1.4. Syndrome MED13L 43

1.4.1. Historique du syndrome 43

1.4.2. Phénotypes associés 44

1.4.3. Variations pathogènes 45

1.4.4. Survenue 49

1.4.5. Données fonctionnelles 49

1.4.6. Données *in silico* 51

2. OBJECTIFS 53

3. RESULTATS 56

ARTICLE 1 : PHÉNOTYPE ASSOCIÉ AU SYNDROME *MED13L* – DESCRIPTION D'UNE SERIE DE 36 PATIENTS 59

ARTICLE 2 : PROFIL *IN SILICO* DES VARIATIONS FAUX-SENS PATHOGENES DU GÈNE *MED13L* 89

108

ARTICLE 3 : ETUDES FONCTIONNELLES DE VARIATIONS FAUX-SENS PATHOGENES DU GENE *MED13L* 109

4. DISCUSSION ET CONCLUSION 139

REFERENCES *150*

ANNEXES *169*

TABLES

Tableau 1. **Pathologies constitutionnelles en lien avec des sous-unités du complexe médiateur** 36

Tableau 2. **Synthèse du phénotype rapporté dans le syndrome MED13L.** Adapté d'après (105) 44

Tableau 3. Variations pathogènes SNV/INDEL de MED13L rapportées en pathologies 45

Tableau 4 **Variations faux-sens pathogènes et possiblement pathogènes rapportées dans PubMed, ClinVar, DECIPHER et LOVD étudiées dans l'article 2 de la thèse**..... 170

FIGURES

Figure 1. **Représentation schématique du complexe médiateur** (figure adaptée d'après (3)). Le module kinase CDK8 d'association transitoire est représenté en orange. Le « head module » est représenté en vert, le module centrale ou « middle module » en jaune et le module de queue ou « tail module » en bleu. La protéine MED14 en gris possède un rôle transversal de soutien des différentes sous structures. 22

Figure 2. **Associations et combinaisons des sous-unités du module kinase CDK8.** Les interactions positives sont marquées par des traits noirs, les exclusions d'associations par des traits rouges (7)..... 24

Figure 3. **Différents rôles du complexe médiateur**. Adaptée d'après (13). 25

Figure 4. **Modèle de boucles chromatiniennes, transcription et complexe médiateur**. Figure adaptée d'après (19). Les régions régulatrices « enhancers » et « promoters » sont représentées par des rectangles bleus. Elles se lient respectivement aux facteurs de

transcriptions (TF, en violet) et au complexe ARN polymérase II (en gris). Le complexe médiateur pourrait avoir un rôle facilitateur dans l'assemblage, en lien avec le complexe cohésine (boucle en orange). 26

Figure 5. **Visualisation de la région génome Chr12(GRCh37):114000000-119000000 comprenant le gène MED13L (Chr12 :116396381-116714991)**. De haut en bas sont représentés les locus cytogénétiques, les profils de Hi-C déterminés par Rao et collaborateurs (26), les gènes impliqués, les interactions entre gènes et régions régulatrices (27) et la conservation inter-espèces de la région. Alignements des régions réalisées grâce à l'outil en ligne UCSC Genome Browser (28). 28

Figure 6. **Domaines connus des protéines MED13L Hsa, MED13 Hsa et Med13 Sce.** Les domaines N et C-terminaux sont représentés en vert, les régions intrinsèquement désorganisées en jaune. La représentation des domaines PAZ (rose), Mid (orange) et PIWI (cyan) chez Saccharomyces cerevisiae est adaptée d'après (33). Les motifs LXXLL de MED13 et MED13L sont en bleu, la région leucine zipper en violet. Les représentations de MED13 et MED13L sont respectivement basées sur les données UniProt, modèles Q9UHV7 et Q71F56. Les points rouges représentent les localisations des motifs phosphodegron. Illustration réalisée avec l'outil IBS (34). 29

Figure 7. **Alignements des séquences protéiques de MED13L Hsa (complète), MED13 Hsa (complète) et SSN2 Sce (partielle). (A)** Alignement de la séquence de MED13L Hsa contre la séquence de MED13 Hsa ; **(B-C-D)** Alignements des motifs candidats PAZ, Mid et PIWI de SSN2 Sce contre MED13L Hsa. Alignement réalisé grâce à l'outil en ligne « SIM – Alignment Tool for protein sequences » et illustré par l'outil LALNVIEW (39). 31

Figure 8. **Modélisation de MED13L proposée par la base AlphaFold** (40). (A) Prédiction de la structure 3D de MED13L par AlphaFold : une grande partie de la structure (orange) reste non caractérisée ; (B) Prédiction des erreurs d'alignements théoriques. Plus la distance en Ångströms est petite (en vert foncé), plus la probabilité d'erreur est faible dans la prédiction ; plus la distance est importante (vert clair), plus la probabilité d'erreur est importante. Les blocs en vert foncé correspondent aux domaines figurant les prédictions structurales les plus probables. 32

Figure 9. **Motif phosphodegron des protéines MED13/MED13L** centré sur l'acide aminé thréonine en position protéique p.326. Le motif « TPPTSPE » présente une homologie de 100% entre MED13 et MED13L. 33

Figure 10. **Régulation de MED13/MED13L au sein du complexe médiateur**. Schéma adapté d'après (30,42,43). Le module kinase CDK8 d'association transitoire est représenté en orange. Le « head module » est représenté en vert, le « middle module » en jaune et le « tail module » en bleu. La régulation de MED13L fait intervenir dans un premier temps la phosphorylation du motif phosphodegron, permettant la reconnaissance et l'ubiquitination de cibles proches du phosphodegron et l'adressage de la protéine pour une dégradation via la voie de l'ubiquitine protéasome. 33

Figure 11. **Sous-unités du complexe médiateur décrites en pathologie constitutionnelle**. Le module kinase CDK8 d'association transitoire est représenté en orange. Le module de tête ou « head module » est représenté en vert, le module central ou « middle module » en jaune et le module de queue ou « tail module » en bleu. La protéine MED14, en gris, possède un rôle transversal de soutien des différentes sous structures. Les sous-unités altérées sont marquées par un éclair rouge. 35

Figure 12. **Variations pathogènes rapportées dans MED13, paralogue de MED13L.** Les variations pathogènes décrites dans MED13 sont en haut (carré orange : faux-sens ; pentagone rouge : non-sens ; triangle rouge : décalage de cadre de lecture). Illustration réalisée avec l'outil IBS (34). 37

Figure 13. **Diagramme de Venn présentant une synthèse des phénotypes associés à MED17-MED20-MED23-MED25-MED27**. 42

Figure 14. **Localisation des variations pathogènes et candidates SNV/INDEL de MED13L rapportées dans la littérature.** Les variations pathogènes décrites dans le syndrome MED13L sont en haut (carré orange : faux-sens ; pentagone rouge : non-sens ; triangle rouge : décalage de cadre de lecture), les variations faux-sens décrites dans les cardiopathies et les troubles du spectre autistique sont en bas (respectivement carré rose et carré bleu). Illustration réalisée avec l'outil en ligne IBS (34). 48

Figure 15. **Prédictions des régions critiques pour les substitutions dans la protéine MED13L** (A) Base MTR, les régions protéiques identifiées en rouge sur le graphique présentent un éloignement significatif par rapport aux variations neutres. Graphique adapté de (167) ; (B) Base Metadome, les régions identifiées en orange/rouge, sont associées à un score élevé d'intolérance à la substitution. Graphique adapté de (168). 52

Figure 16. **Nombre de variations candidates pathogènes dans MED13L rapportées en publication depuis 2003**. Les variations faux-sens sont en orange, les variations associées à une possible haplo-insuffisance sont en bleu (non-sens, décalage de cadre de lecture, microdélétions). LoF = variations « pertes de fonction ». 140

Figure 17. **Comparaison des phénotypes associés aux variations faux-sens et pertes-de-fonction fondées sur une étude de 32 patients** (Communication personnelle Thomas SMOL – Assises de génétique 2018). DI/RD = Déficience intellectuelle – Retard de développement ; H/D = haut/dehors ; TC = troubles du comportement ; TSA = Troubles du spectre autistique. 141

Figure 18. **Localisation des variations faux-sens de l'exon 15**. Les sites connus de phosphorylation sont identifiés par la lettre p minuscule en vert, les acides aminés rapportés en pathologie sont en rouge et en gras, une flèche indique les substitutions décrites. 142

Figure 19. **Localisation des sites potentiels de phosphorylation** (en bleu) **et d'ubiquitination** (en rose) sur MED13L selon la base PhosphoSite. Figure adaptée d'après (174). 146

Figure 20. **Comparaison des paramètres in silico (A) associés aux variations faux-sens et des valeurs médianes associées aux différents outils de prédictions in silico pour les substitutions pathogènes et neutres (B).** 148

ABBRÉVIATIONS

ADN	:	Acide Désoxyribonucléique
ARN	:	Acide Ribonucléique
CKM	:	Module kinase CDK8/CDK19
DI	:	Déficience intellectuelle
FGF	:	*Fibroblast Growth Factor*
H.s	:	*Homo sapiens*
IDR	:	*Intrinsically Disordered Region*
KO	:	*Knock Out*
M.m	:	*Mus musculus*
MG132	:	Carbobenzoxy-l-leucyl-l-leucyl-l-leucinal
PIC	:	Complexe de Pré-Initiation de la transcription
S. c	:	*Saccharomyces cerevisiae*
SCF	:	*Skp1-Cul1-F-box protein ubiquitin ligase*
TAD	:	Domaine Topologiquement Associé
TGV	:	Transposition des Gros Vaisseaux

LEXIQUE DES GÈNES – NOMENCLATURE HUGO

HGNC	**Approved name**	**Locus**	**OMIM**
C12orf49	Chromosome 12 open reading frame 49	12q24.22	N/A
CDK8	cyclin dependent kinase 8	13q12.13	#603164
CDK19	cyclin dependent kinase 19	6q21	#614720
FBXO21	F-box protein 21	12q24.22	#609095
FBXW7	F-box and WD repeat domain containing 7	4q31.3	#606276
FBXW8	F-box and WD repeat domain containing 8	12q24.22	#609073
FGFR3	fibroblast growth factor receptor 3	4p16.3	#134934
FOS	Fos proto-oncogene, AP-1 transcription factor subunit	14q24.3	#164810
JUN	Jun proto-oncogene, AP-1 transcription factor subunit	1p32.1	#165160
MAP1LC3B2	microtubule associated protein1 light chain 3 beta 2	12q24.22	N/A
MED6	mediator complex subunit 6	14q24.2	#602984
MED12	mediator complex subunit 12	Xq13.1	#300188
MED12L	mediator complex subunit 12L	3q25.1	#611318
MED13	mediator complex subunit 13	17q23.2	#603808
MED13L	mediator complex subunit 13L	12q24.21	#608771
MED15	mediator complex subunit 15	22q11.21	#607372
MED17	mediator complex subunit 17	11q21	#603810
MED20	mediator complex subunit 20	6p21.1	#612915
MED23	mediator complex subunit 23	6q23.2	#605042
MED25	mediator complex subunit 25	19q13.33	#610197
MED27	mediator complex subunit 27	9q34.13	#605044
POLR2C	RNA polymerase II subunit C	16q21	#180663
RNFT2	ring finger protein, transmembrane 2	12q24.22	N/A
SP8	Sp8 transcription factor	7p21.1	#608306
SPRING1	SREBF pathway regulator in golgi 1	12q24.22	N/A
STAT1	Signal transducer and activator of transcription 1	2q32.2	#600555
TBX3	T-box transcription factor 3	12q24.21	#601621
TESC	tescalcin	12q24.22	#611585

1. INTRODUCTION

1.1.Le complexe médiateur

1.1.1. Organisation du complexe

1.1.1.1. Corps du complexe

Identifié il y a environ 25 ans, le complexe médiateur, appelé également « médiateur du complexe ARN polymérase II » possède une place centrale dans la régulation de la transcription par ses interactions avec l'ARN polymérase II, dans le recrutement des protéines d'architecture chromatinienne, ainsi que dans la régulation des protéines d'assemblage au promoteur (1). Son organisation est complexe, dynamique et variable selon les tissus. Le complexe médiateur est une structure multi-protéique de plus d'1 MDa, composé d'une trentaine de protéines d'expression ubiquitaire et reparties en 4 modules très conservés (Figure 1) (2).

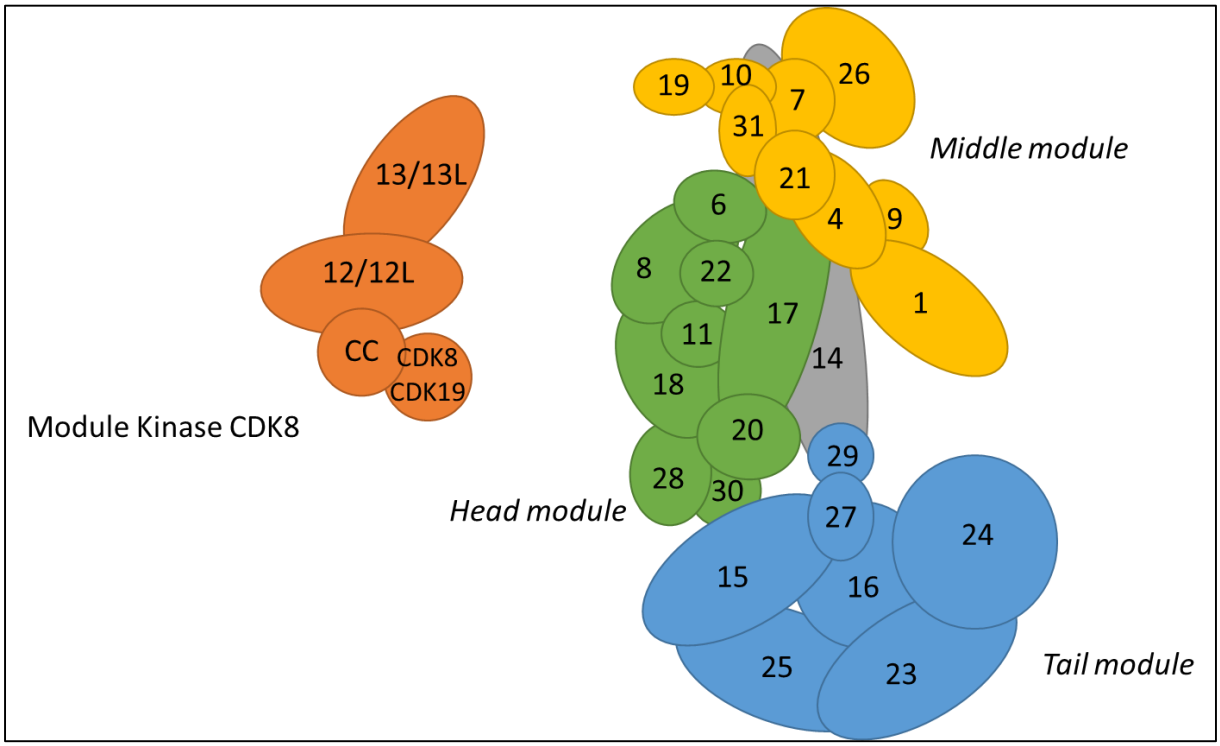


Figure 1. ***Représentation schématique du complexe médiateur*** *(figure adaptée d'après (3)). Le module kinase CDK8 d'association transitoire est représenté en orange. Le « head module » est représenté en vert, le module centrale ou « middle module » en jaune et le module de queue ou « tail module » en bleu. La protéine MED14 en gris possède un rôle transversal de soutien des différentes sous structures.*

Le corps du complexe est divisé en 3 sous structures appelées module de tête ou « *head module* », composé de 7 à 9 sous-unités, module centrale ou « *middle module* », de 8 à 9 sous-

unités et module de queue ou « *tail module* », de 2 à 9 sous unités (4). Par commodités pour la suite de l'exposé, nous utiliserons la terminologie anglophone pour décrire les structures du complexe médiateur. L'assemblage de ces 3 structures est réalisé en partie autour de la sous-unité MED14, très conservé de la levure à l'Humain, qui possède un rôle pivot dans l'assemblage de l'ensemble des sous-unités (5). Des travaux réalisés chez *Saccharomyces cerevisiae* (*Sce*) ont permis d'identifier des interactions fortes entre SRB4 (MED17), MED6 et SRB2 (MED20) du « *head module* » et RGR1 (MED14), ainsi qu'entre RGR1 (MED14) et GAL11 (MED15) du « *tail module* » (6).

La structure du corps du complexe médiateur est très conservée entre les espèces ainsi qu'au sein de l'évolution. Néanmoins, on peut noter l'émergence de sous-unités chez les eucaryotes, dont MED23, MED25, MED26, MED28 et MED30 retrouvés chez les métazoaires, animaux pluricellulaires (3).

1.1.1.2. Module kinase CDK8

Les trois sous-structures permanentes du complexe médiateur forment son « corps ». Elles sont associées à une 4ème sous-structure, appelée module kinase CDK8 (CKM) d'une taille de 600 kDa, et qui interagit avec le corps du complexe médiateur (Figure 1) (2). Le CKM est caractérisé par sa capacité d'association transitoire et réversible avec le complexe (1). Il est composé de 4 sous-unités chez *Saccharomyces cerevisiae*: SRB8 (MED12), SSN2 (MED13), SSN3 (CDK8) et la cycline C. Chez les vertébrés dont *Homo sapiens* (*Hsa*), la structure du CKM s'est complexifiée et diversifiée avec l'émergence de paralogues des sous-unités MED12, MED13 et CDK8, respectivement MED12L, MED13L et CDK19 (7).

Ces structures paralogues possèdent notamment comme propriétés d'être mutuellement exclusives, *i.e.*, lorsque la protéine MED13L est présente au sein du CKM, la protéine MED13 ne peut être présente. Ce constat est également fait pour MED12 et MED12L, ainsi que pour CDK8 et CDK19. Si les différents paralogues peuvent présenter un structure proche permettant une incorporation dans le CKM, les fonctions associées sont différentes, leur conférant une spécificité propre (1). L'existence de ces paralogues permet l'assemblage du CKM sous 8 combinaisons différentes chez les vertébrés, en lien avec des fonctions et des tissus différents (Figure 2) (7). CDK8 et CDK19 sont des protéines kinases dépendantes des

cyclines impliquées dans la régulation du cycle cellulaire et de la transcription. L'activité kinase de la protéine CDK8 est fonction de sa capacité de liaisons avec ses protéines partenaires du CKM, cycline C et MED12/MED12L (8).

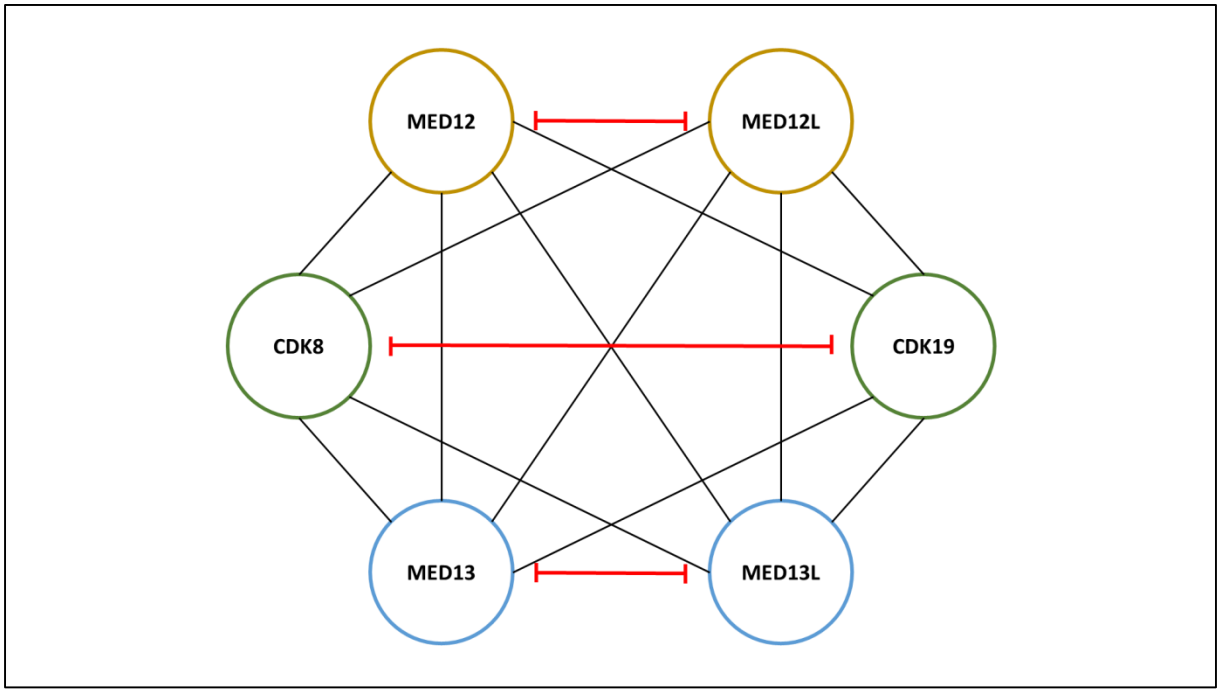


Figure 2. ***Associations et combinaisons des sous-unités du module kinase CDK8.*** *Les interactions positives sont marquées par des traits noirs, les associations mutuellement exclusives par des traits rouges (7).*

Les différents paralogues possèdent une répartition tissulaire spécifique. La sous-unité MED13 est associée à une forte expression au niveau des tissus musculo-squelettiques et cardiaques (9,10), tandis que MED13L possède une expression cérébrale importante, notamment au cours du développement embryonnaire, ainsi que dans les tissus hépatiques et rénaux (11,12).

Au sein du CKM, les paralogues MED13/MED13L agissent comme une ancre du module kinase. Ils interagissent avec le complexe médiateur en s'associant notamment avec la sous-unité MED19 située dans le « *middle module* » et MED14, sous-unité transversale, rattachée fonctionnellement au « *middle module* » (5). Si leurs fonctions structurales dans le CKM sont proches, on peut noter des différences dans certaines associations. Par exemple, MED13L, à la différence de MED13, peut s'associer de manière spécifique avec la sous-unité MED26 du « *middle module* » (7).

1.1.2. Rôles du complexe médiateur

Résumer les rôles et les interactions du complexe médiateur est un challenge tant celui-ci par sa composition, sa structure dynamique et ses très nombreux réseaux d'interaction est essentiel aux fonctions nucléaires (Figure 3). Différentes études portant sur le complexe médiateur font émerger trois rôles essentiels, à savoir, régulation de la transcription, conformation chromatinienne et réparation de l'ADN.

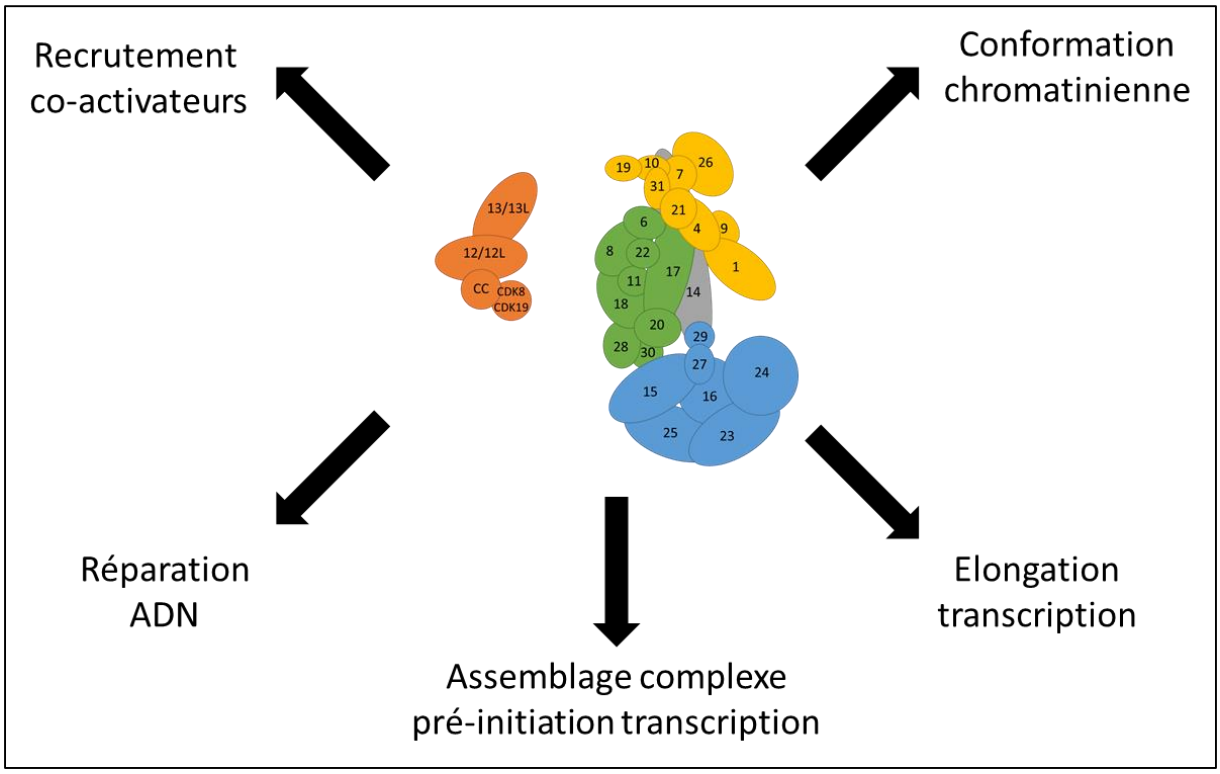


Figure 3. ***Différents rôles du complexe médiateur****. Adaptée d'après (13).*

1.1.2.1. Régulation de la transcription

Les fonctions du complexe médiateur sont reliées à ses capacités de liaisons directes avec des complexes de régulations de la transcription. Le complexe médiateur possède un rôle majeur dans la régulation de la transcription. Le corps du complexe médiateur agit comme stabilisateur dans l'interaction entre les facteurs de transcriptions et l'ARN polymérase II lors de la phase d'assemblage du complexe de pré-initiation de la transcription (*Pre-Initiation* Complex, PIC). Il interagit avec l'ARN polymérase II, et les facteurs TFIIB, TFIID et TFIIH (14). Cette interaction est facilitée par une modification dynamique de la structure du

complexe médiateur, notamment de la conformation du « *head module* » facilitant les contacts structuraux entre le complexe médiateur et le PIC, ainsi qu'entre les régions promotrices et le complexe médiateur (3,15). *In vitro* chez *Saccharomyces cerevisiae*, des interactions sont notées entre SRB4 (MED17), RGR1 (MED14) et TFIIB, et entre NUT2 (MED10) et TFIIB (16). Un assemblage dynamique du PIC aux régions promotrices facilité par le complexe médiateur est possible.

Le complexe médiateur stimule également la phosphorylation du domaine C-terminal de l'ARN polymérase II, déclenchant sa libération après l'étape d'initiation de la transcription, basculant vers l'élongation (17).

1.1.2.2. Conformation chromatinienne

Le complexe médiateur a été initialement décrit comme facilitateur dans la transmission des signaux de régulation (3). On suspectait une interaction d'une part, entre les régions activatrices (*enhancers*) et des facteurs de transcription recrutés par le complexe médiateur et d'autre part, entre le complexe médiateur et le complexe ARN polymérase II (Figure 4). Ces interactions croisées seraient ainsi facilitées par la capacité d'association entre le complexe médiateur et les membres du complexe cohésine pour la formation de boucles chromatiniennes chez les métazoaires (18).

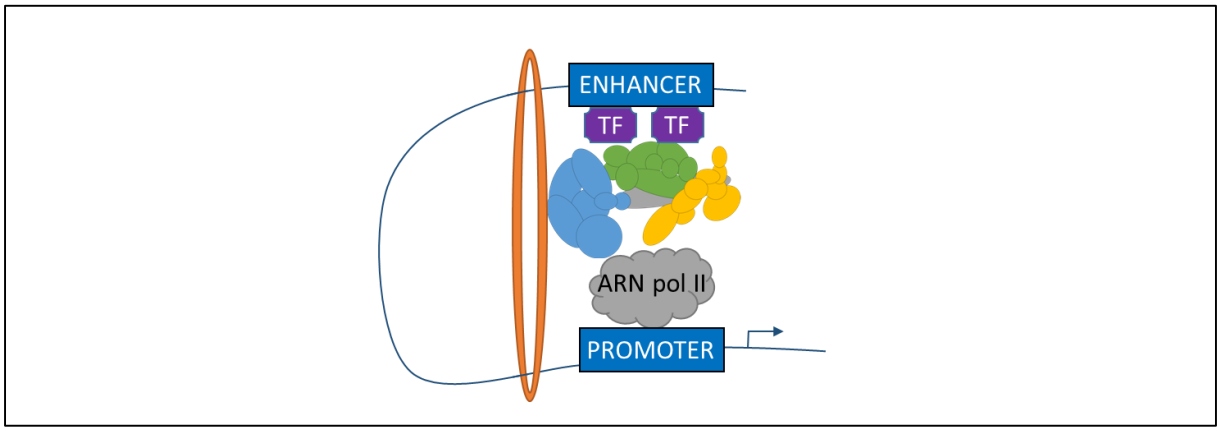


Figure 4. ***Modèle de boucles chromatiniennes, transcription et complexe médiateur****. Figure adaptée d'après (19). Les régions régulatrices « enhancers » et « promoters » sont représentées par des rectangles bleus. Elles se lient respectivement aux facteurs de transcriptions (TF, en violet) et au complexe ARN polymérase II (en gris). Le complexe médiateur pourrait avoir un rôle facilitateur dans l'assemblage, en lien avec le complexe cohésine (boucle en orange).*

Ce modèle a récemment été remis en question par les études de déplétions en complexes médiateurs dans des cellules souches murines sans altération dans les interactions entre promoteurs et régions activatrices, *a contrario* des déplétions en complexes cohésines (20).

Un rôle du complexe médiateur apparaît probable dans l'organisation de l'architecture globale chromatinienne *via* les domaines topologiquement associés (TAD). Ces unités fonctionnelles du génome, d'une taille d'environ 800 kb, sont enrichies à leurs extrémités frontières 5' et 3' par des marques insulatrices, dont les marques CTCF, ZNF143 et YY1 (21). Une étude dans le modèle levure par ChIP-Seq sur les domaines d'interactions chromatiniens, équivalent des TAD chez les mammifères, a permis la caractérisation d'un co-enrichissement des marques associées au médiateur complexe dans les régions frontières, notamment les frontières aux propriétés insulatrices les plus fortes (22). S'il s'avère possible que le complexe médiateur n'agisse pas directement dans la formation de boucle chromatinienne, il pourrait cependant avoir un rôle dans le recrutement de protéines d'architectures impliquées dans les boucles chromatinienne.

1.1.2.3. Réparation de l'ADN

En tant que co-activateur du complexe ARN polymérase II, le complexe médiateur serait impliqué dans le processus de réparation de l'ADN par excision de nucléotides. En effet, chez *Saccharomyces cerevisiae*, un lien direct a été identifié entre RAD2, l'orthologue de l'endonucléase XPG, et SRB4 (MED17), l'orthologue de la sous-unité MED17 située dans le « *head module* » (23). Ce lien provient de l'observation d'une diminution de la liaison de RAD2 *Sce* aux régions transcrites en présence de mutations possiblement pathogènes de la sous-unité SRB4 (MED17) *Sce*, ainsi qu'une augmentation cellulaire de la sensibilité aux ultraviolets (UV) (23). Chez l'Homme, une augmentation de la sensibilité aux UV a notamment été observée, *in vitro*, pour des lignées cellulaires porteuses de variations candidates dans les sous-unités *MED12L* et *MED13L* (24).

1.2. Structure et rôles de MED13L

1.2.1. Localisation

Le gène *MED13L*, initialement décrit sous les terminologies *PROSIT240* et *THRAP2* (*Thyroid Hormone Receptor Associated Protein 2*), code pour le paralogue de la 13ème sous-unité du complexe médiateur. Au *locus* 12q24.21, *MED13L* est situé dans le TAD des gènes *MAP1LC3B2*, *C12orf49/SPRING1*, *RNFT2*, *FBXW8*, *TESC* et *FBXO21*.

Une large région dépourvue de régions codantes d'une taille de 1,3 Mb est décrite entre *MED13L* Chr12(GRCh37):116396381-116714991 et le gène *TBX3* Chr12(GRCh37):115108059-115121969 situé dans le TAD en 5'(Figure 5) (25).

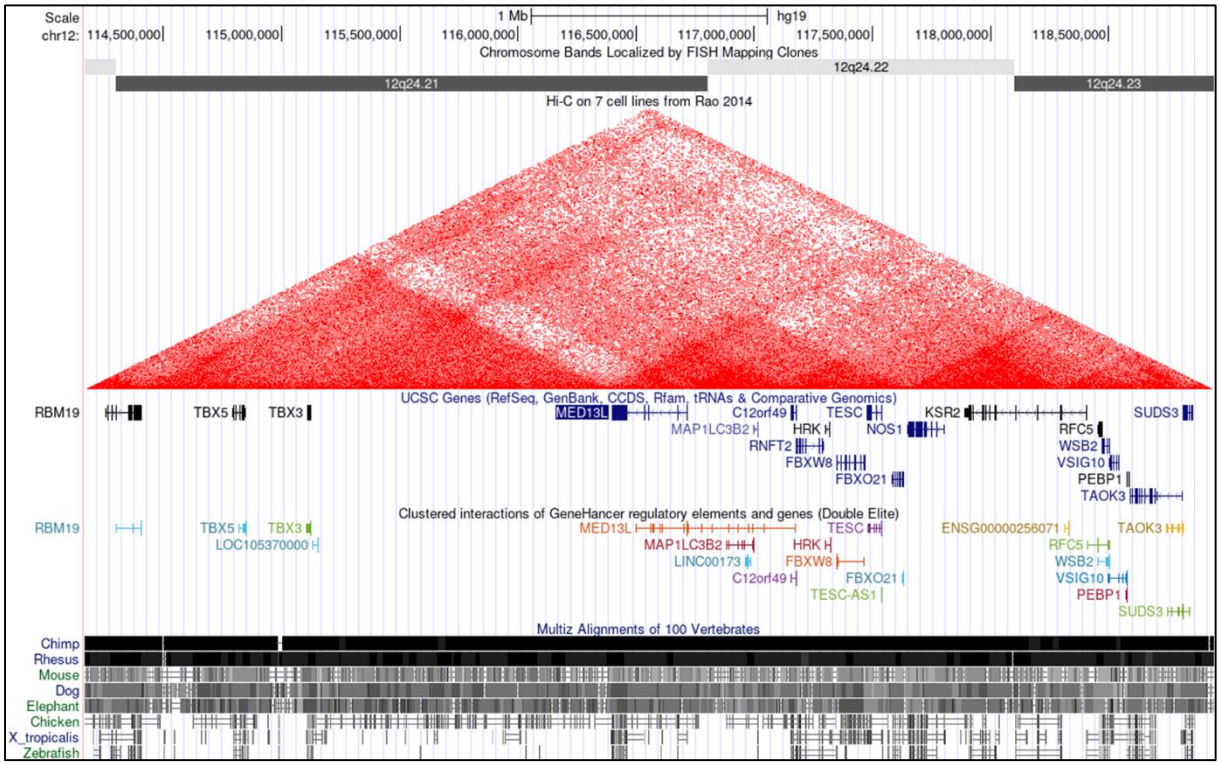


Figure 5. ***Visualisation de la région génome Chr12(GRCh37):114000000-119000000 comprenant le gène MED13L (Chr12 :116396381-116714991)****. De haut en bas sont représentés les locus cytogénétiques, les profils de Hi-C déterminés par Rao et collaborateurs (26), les gènes impliqués, les interactions entre gènes et régions régulatrices (27) et la conservation inter-espèces de la région. Alignements des régions réalisées grâce à l'outil en ligne UCSC Genome Browser (28).*

1.2.2. Structure protéique

1.2.2.1. Domaines et régions connues

Sa séquence de 2210 acides aminés présente 53% de similarité avec son paralogue MED13, localisé en 17q23.2 (29). L'émergence de MED13L chez les métazoaires est probablement en lien avec un mécanisme de duplication génique suivi d'un processus de sous-fonctionnalisation (7).

La protéine MED13L est située à l'interface entre le CKM et le complexe médiateur. Elle permet notamment la liaison du CKM au complexe médiateur (30). Des expérimentations de d'inhibition de *MED13L* par siRNA dans des lignées cellulaires NSCLC (*non-small cancer cell line*) ne montrent pas de modification de l'intégrité des interactions entre les composants MED8 *(head module*), MED1 (*middle module*) et MED23 (*tail module*), ni entre les composants du CKM : MED12-CDK8-Cycline C. En revanche, l'inhibition de *MED13L* altère les interaction entre le CKM et la structure du complexe médiateur (31).

La composition, l'organisation, et la structure de MED13L chez l'*Homo sapiens* restent peu caractérisées (Figure 6). Par similarité de séquence avec SSN2 (MED13) *Sce*, une région centrale intrinsèquement désorganisée (IDR) est suspectée pour MED13L. La région IDR est caractérisée par une succession d'acides aminés non structurés, en lien avec des propriétés élevées de plasticité permettant la transition d'un état à un autre. Cette région confère une flexibilité élevée aux protéines MED13L et MED13 (32).

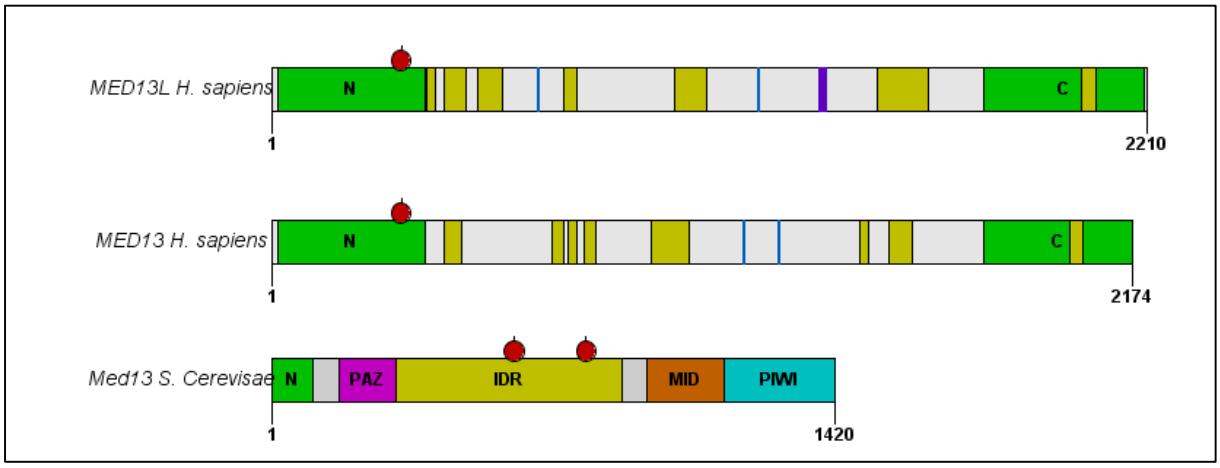


Figure 6. ***Domaines connus des protéines MED13L Hsa, MED13 Hsa et Med13 Sce.*** *Les domaines N et C-terminaux sont représentés en vert, les régions intrinsèquement désorganisées en jaune. La représentation des domaines PAZ (rose), Mid (orange) et PIWI (cyan) chez Saccharomyces cerevisiae est adaptée d'après (33). Les motifs LXXLL de MED13 et MED13L sont en bleu, la région leucine*

zipper en violet. Les représentations de MED13 et MED13L sont respectivement basées sur les données UniProt, modèles Q9UHV7 et Q71F56. Les points rouges représentent les localisations des motifs phosphodegron. Illustration réalisée avec l'outil IBS (34).

Deux motifs signatures « LXXLL » de liaison aux récepteurs nucléaires, où « L » représente une leucine et « X » n'importe quel acide aminé, conservés entre MED13 et MED13L et situés en 669-673 et 1225-1229 pour MED13L sont également décrits (Figure 6) (35). La séquence aminoacide en 669-673, « TALQRLL », correspond à une séquence de type « *nuclear receptor boxes* » (NRB) de type III : « (T/S)ΦLXXLL » où « Φ » représente un acide aminé hydrophobe et T/S une sérine ou une thréonine (36). Les NRB de classe III sont notamment retrouvées parmi les co-activateurs transcriptionnels. La séquence en 1225-1229 « ISLLLLL » pourrait également ressembler à un motif NRB de classe III (36). Une région « *Leucine-zipper* » est suspectée aux coordonnées protéiques 1331-1352 (Figure 6). Ce motif habituellement constitué de 60 à 80 acides aminés est souvent identifié parmi les facteurs de transcription et les protéines régulatrices. Il possède le plus souvent une organisation secondaire α-hélicoïdale (37). Les fonctionnalités associées à ces motifs, identifiés par similarité de séquences, restent peu caractérisées pour MED13L.

1.2.2.2. Domaines et régions hypothétiques

La caractérisation des interactions entre MED13L, ses partenaires du CKM et ceux du complexe médiateur présente un défi majeur aux équipes, lié à la fois à la taille du complexe et à la stabilité de ses sous-unités. Une évaluation de la structure du CKM par cryo-électro-microscopie a pu être réalisée sur *Saccharomyces cerevisiae* (33). MED13L n'existant pas chez la levure, la modélisation de structure concernait son orthologue SSN2 (MED13), présentant 20.09% de similarités de séquences. Les auteurs ont pu localiser SSN2 (MED13) *Sce* au centre du CKM, interagissant directement avec SRB8 (MED12) et la Cycline C *Sce*.

Plusieurs hypothèses concernant les domaines protéiques de SSN2 (MED13) *Sce* ont été posées postulant de la présence d'un domaine N-terminal, d'un domaine PAZ, d'un domaine Mid, d'un domaine PIWI, ainsi que de deux domaines de liaisons L1-L2 (33). La présence de ces différents domaines dans SSN2 (MED13) rapprocherait sa structure des protéines de la famille Argonaute notamment impliquées dans la régulation par petits ARN non codants (38). Les protéines de la famille Argonaute appartiennent au complexe *RNA-induced silencing*

complex et peuvent s'associer aux petits ARN non codants pour agir comme régulateurs négatifs. Elles sont composées d'une succession de 4 domaines : un domaine N-terminal, un domaine PAZ d'environ 110 acides aminés, responsables de la liaison avec les petits ARN non-codants, un domaine Mid et un domaine PIWI, associés à des propriétés catalytiques (38).

Si l'on peut supposer que des réminiscences de ces domaines existent chez MED13L *Hsa*, qui la classerait parmi les protéines de la famille Argonaute, l'éloignement des séquences entre MED13L *Hsa* et SSN2 (MED13) *Sce* rend difficile les projections des motifs de l'une à l'autre des protéines. La présence d'une large région IDR semble commune entre MED13 *Hsa* MED13L *Hsa* et SSN2 (MED13) *Sce* (Figure 6). Les alignements des séquences protéiques ne permettent pas de définir des similarités fortes de domaine PAZ, Mid ou PIWI dans MED13L *Hsa* (Figure 7).

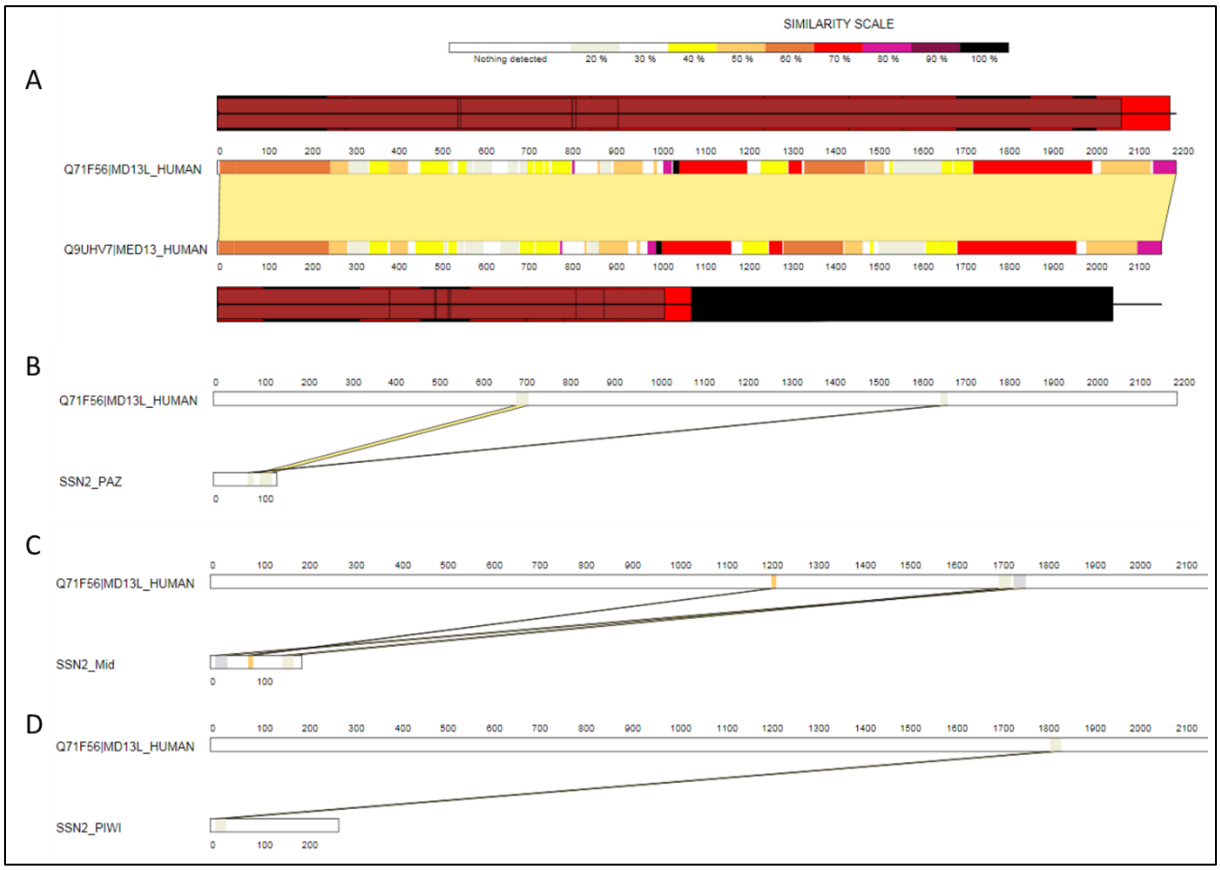


Figure 7. ***Alignements des séquences protéiques de MED13L Hsa (complète), MED13 Hsa (complète) et SSN2 Sce (partielle). (A)*** *Alignement de la séquence de MED13L Hsa contre la séquence de MED13 Hsa ;* ***(B-C-D)*** *Alignements des motifs candidats PAZ, Mid et PIWI de SSN2 Sce contre MED13L Hsa. Alignement réalisé grâce à l'outil en ligne « SIM – Alignment Tool for protein sequences » et illustré par l'outil LALNVIEW (39).*

La structure 3D de MED13L *Hsa* reste ainsi très majoritairement non déterminée. Selon les modélisations proposées par la base AlphaFold de la séquence Q71F56 d'UniProt, les régions situées entre les acides aminés 1 et 250, 1100 et 1500, 1650 et 2000, puis 2100 à 2210, semblent être structurées avec la présence de structures protéiques secondaires en hélices **α** et feuillets ß (Figure 8A). Une grande partie de la structure qui reste non établie, correspond au domaine intrinsèque désorganisé (IDR) (Figure 6 et Figure 8A-B).

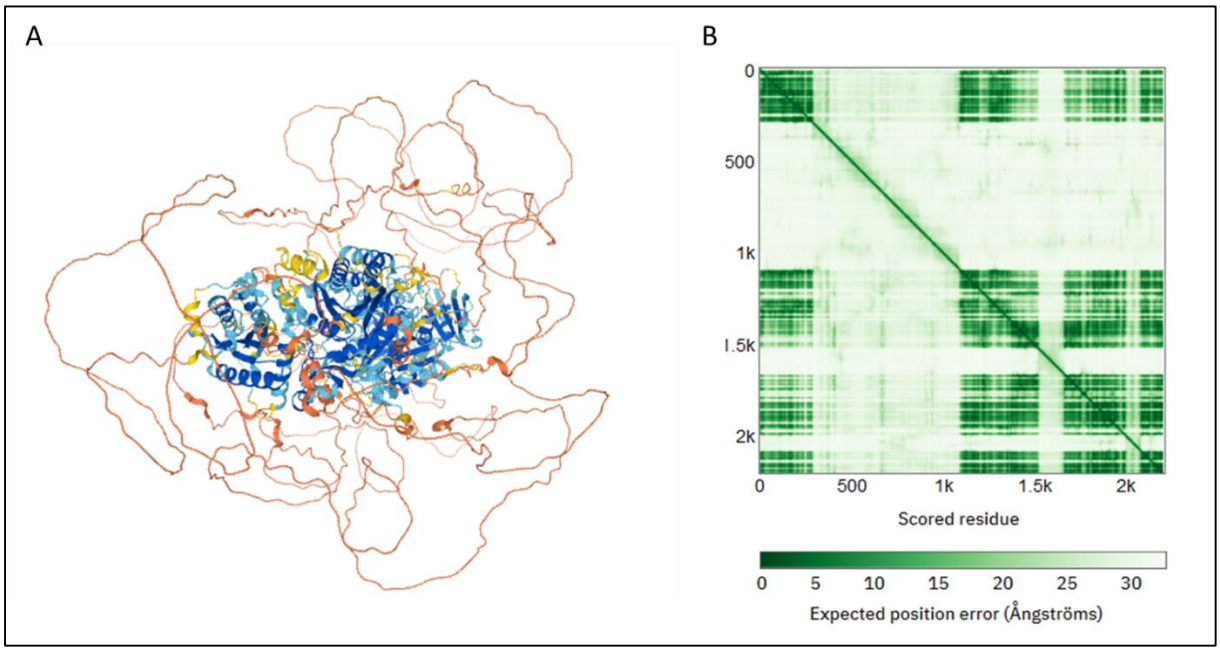


Figure 8. ***Modélisation de MED13L proposée par la base AlphaFold*** *(40). (A) Prédiction de la structure 3D de MED13L par AlphaFold : une grande partie de la structure (orange) reste non caractérisée ; (B) Prédiction des erreurs d'alignements théoriques. Plus la distance en Ångströms est petite (en vert foncé), plus la probabilité d'erreur est faible dans la prédiction ; plus la distance est importante (vert clair), plus la probabilité d'erreur est importante. Les blocs en vert foncé correspondent aux domaines figurant les prédictions structurales les plus probables.*

1.2.3. Régulation

Comme MED13, MED13L possède un motif phosphodegron « ΦXΦΦΦ(T/S)PPX(T/S/E/D) » centré sur la Thréonine 326 dans les deux paralogues (Figure 9). Le motif comprend une succession d'acides aminés hydrophobes « Φ », suivie d'une phospho-sérine ou d'une phospho-thréonine « T/S » : « TPPTSPE ». La phosphorylation du phosphodegron par une kinase permet l'identification du motif par la protéine FBXW7, appelé également FBW7 (30). La protéine suppresseur de tumeur FBXW7, substrat de reconnaissance du complexe ubiquitine E3 SKP-cullin-F-box (SCF) ligase,

reconnait et se lie aux motifs protéiques phospho-thréonines et phospho-sérines. Cette liaison permet secondairement l'ubiquitination, puis la dégradation protéique via la voie du protéasome (41).

```
Q9UHV7|MED13_HUMAN      STHCSSSCLGVHQVPASTRDPAMSSVTLTPPTSPEEVQTVDPQSVQKWV-KFSSVSDGFN      356
Q71F56|MD13L_HUMAN      SAGGHIA-VG-QQGLGSVKDPSNCGMPLTPPTSPEQAILGESGGMQSAASHLVSQDGGMI      357
                        *:    : :* :*  .*.:**: ..: ********:.   :  .:*. . :: * ..*:
```

Figure 9. ***Motif phosphodegron des protéines MED13/MED13L*** *centré sur l'acide aminé thréonine en position protéique p.326. Le motif « TPPTSPE » présente une similarité de 100% entre MED13 et MED13L.*

Dans le contexte du complexe médiateur, la reconnaissance du motif phosphodegron de MED13/MED13L par FBXW7 présente une importance majeure. Après une première étape de liaison au CKM, le complexe SCF^{FBXW7} cible la région phosphodegron phosphorylée de la protéine MED13 ou MED13L, promouvant sa dégradation par la voie de l'ubiquitine protéasome. La dégradation de MED13L ou MED13 entraine indirectement la libération du CKM du complexe médiateur. L'interaction du complexe ARN polymérase II avec le complexe médiateur peut ainsi s'enclencher (Figure 10) (30). Ainsi, on peut considérer que le CKM peut agir comme un répresseur de l'activité transcriptionnelle du complexe ARN polymérase II, *via* la régulation de la dégradation des protéines MED13 et MED13L (2).

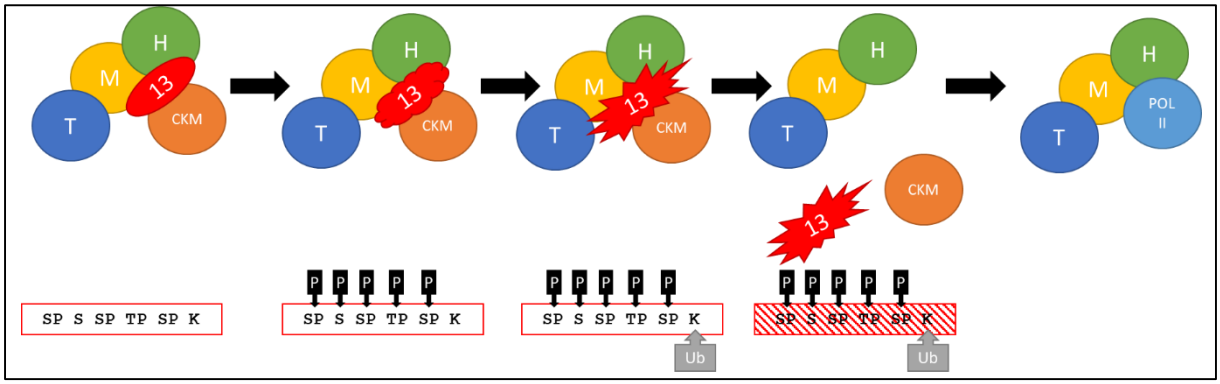


Figure 10. ***Régulation de MED13/MED13L au sein du complexe médiateur****. Schéma adapté d'après (30,42,43). Le module kinase CDK8 d'association transitoire est représenté en orange. Le « head module » est représenté en vert, le « middle module » en jaune et le « tail module » en bleu. La régulation de MED13L fait intervenir dans un premier temps la phosphorylation du motif phosphodegron, permettant la reconnaissance et l'ubiquitination de cibles proches du phosphodegron et l'adressage de la protéine pour une dégradation via la voie de l'ubiquitine protéasome.*

L'implication de la voie ubiquitine protéasome dans la dégradation de MED13/MED13L a été démontrée grâce à des expérimentations d'inhibition du protéasome par le MG132. Le

Carbobenzoxy-l-leucyl-l-leucyl-l-leucinal ou MG132 est un aldéhyde de synthèse réduisant la dégradation des protéines conjuguées à l'ubiquitine chez le eucaryotes (44). Son utilisation dans la lignée cellulaire 293A entraine une rétention partielle du CKM dans le complexe médiateur, une réduction de l'activité transcriptionnelle, ainsi qu'une hausse des niveaux protéiques de MED13 et MED13L (30,45).

La dégradation de MED13/MED13L chez l'Homme est également impliquée dans la réponse cellulaire au stress oxydatif. En effet, dans la levure, la libération de la Cycline C vers le cytoplasme, consécutive à un stress oxydatif, est corrélée aux propriétés de dégradation de SSN2 (MED13) *Sce* via la voie de l'ubiquitine protéasome (32). Cette hypothèse est corroborée par l'étude des niveaux d'expression de MED13/MED13L en western blot après exposition de lignées cellulaires MEF, fibroblastes embryonnaires murins, au H_2O_2. Stieg et collaborateurs ont pu observer une réduction de 20% de la quantité de Med13 et Med13L *Mus musculus* (*Mmu*) après exposition au stress oxydatif (45).

1.2.4. Interactions

Parallèlement à son rôle structural dans le complexe médiateur, MED13L est potiellement impliqué dans la régulation des voies de signalisations « Wnt », du « *fibroblast growth factor* » (FGF) et du « rétinoblastome (Rb) – E2F » (46,47). Concernant les autres partenaires directs ou indirects de MED13L, des études par Chromatin ImmunoPrecipitation Sequencing (ChIP-Seq) ont permis d'observer une réduction du signal d'H3K27ac après inhibition de MED13L dans des lignées cellulaires NSCLS (*non-small cancer cell line*) (31). L'acétylation de l'histone H3K27 est considérée comme une marque indicatrice de l'activité des régions régulatrices *enhancers* (48). Une interaction directe entre MED13L et l'histone acétyltransférase P300 ainsi qu'entre MED13L et le facteur de transcription activateur 4 (AFT4) ont été notées. MED13L pourrait agir en tant que régulateur de protéines associées à la modification de marques histoniques (31). La protéine MED13L pourrait également être impliquée dans une boucle de régulation de l'expression d'AFT4, facteur de transcription essentiel dans la réponse au stress, la résorption osseuse et l'hématopoïèse (49).

1.3. Pathologies associées au complexe médiateur

Les altérations de plusieurs membres du complexe médiateur ont été identifiées en pathologie humaine constitutionnelle (Tableau 1). Cela concerne l'ensemble des paralogues du CKM : MED13L, MED13, MED12, MED12L, CDK8 et CDK19. Mais également, les membres du « *tail module* » MED23, MED25 et MED27, ainsi que MED17 et MED20 du « *head module* ».

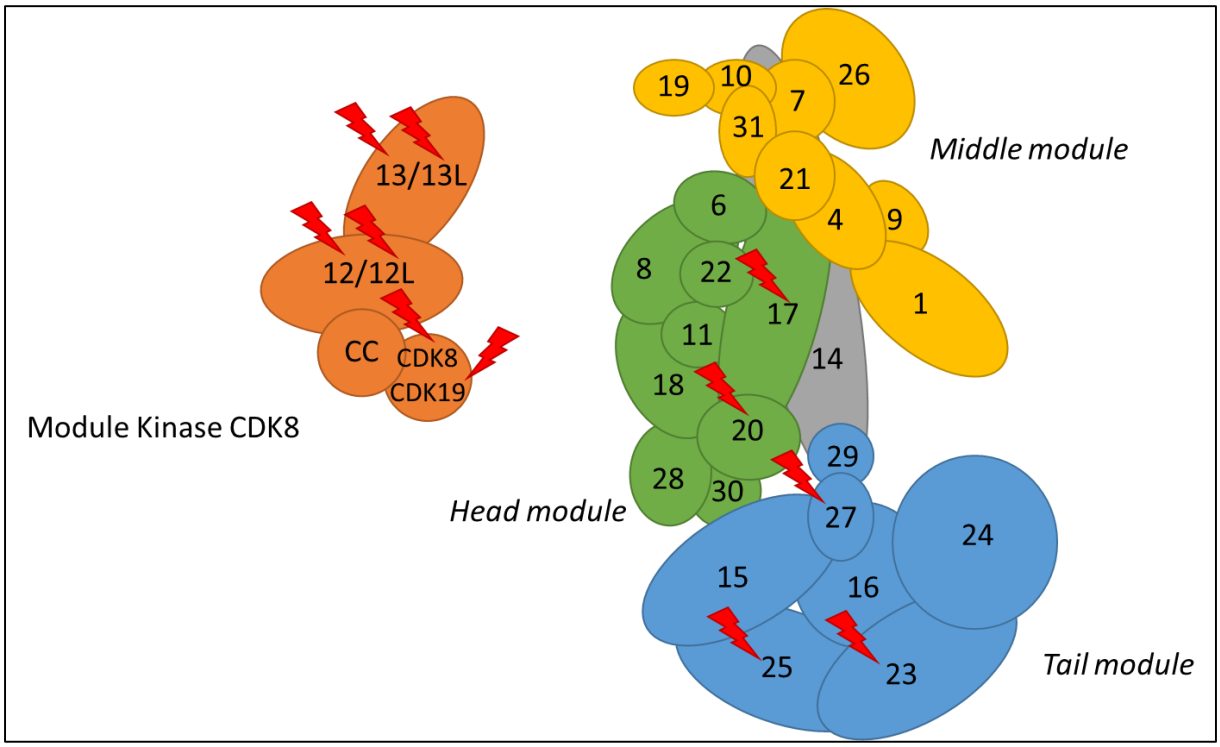


Figure 11. ***Sous-unités du complexe médiateur décrites en pathologie constitutionnelle****. Le module kinase CDK8 d'association transitoire est représenté en orange. Le module de tête ou « head module » est représenté en vert, le module central ou « middle module » en jaune et le module de queue ou « tail module » en bleu. La protéine MED14, en gris, possède un rôle transversal de soutien des différentes sous structures. Les sous-unités altérées sont marquées par un éclair rouge.*

A ce jour, il ne semble pas y avoir d'implication de sous-unités du « *middle module* » en pathologie constitutionnelle (Figure 11). Néanmoins, des variations candidates *de novo* ont été rapportées dans des études de cohortes de patients qui présentaient un trouble du neuro-développement ou un trouble du spectre autistique dans *MED1* (n=1), *MED4* (n=1), *MED7* (n=1), *MED9* (n=1), *MED10* (n=1), *MED14* (n=1), *MED15* (n=3), *MED16* (n=1), *MED21* (n=1), *MED22* (n=1), *MED24* (n=4), *MED26* (n=2), *MED28* (n=2), *MED29* (n=2) et *MED30* (n=1) sans étude, à ce jour, de corrélation génotype/phénotype permettant de confirmer leur implication (50–60).

Tableau 1. ***Pathologies constitutionnelles en lien avec des sous-unités du complexe médiateur***

Structure	Sous-Unités	Pathologies	Transmissions	Références
CKM	MED13L	*Impaired intellectual development and distinctive facial features with or without cardiac defects (MIM #616789)*	AD	(61,62)
	MED13	*Intellectual developmental disorder (MIM #618009)*	AD	(63)
	MED12	*Lujan-Fryns syndrome (MIM #309520)*	XL	(64)
		Ohdo syndrome (MIM #300895)	XL	(65)
		Opitz-Kaveggia syndrome (MIM #305450)	XL	(66)
	MED12L	*Nizon-Isidor syndrome(MIM #618872)*	AD	(24)
	CDK8	*Intellectual developmental disorder with hypotonia and behavioral abnormalities (MIM #618748)*	AD	(67)
	CDK19	*Developmental and epileptic encephalopathy (MIM #618916)*	AD	(68,69)
Head Module	MED17	*Microcephaly, postnatal progressive, with seizures and brain atrophy (MIM #613668)*	AR	(70)
	MED20	*Infantile basal ganglia degeneration and brain atrophy*	AR	(71)
Tail Module	MED23	*Mental retardation, autosomal recessive (MIM #614249)*	AR	(72)
	MED25	*Basel-Vanagait-Smirin-Yosef syndrome (MIM #616449)*	AR	(73,74)
	MED27	*Neurodevelopmental disorder with spasticity, cataracts, and cerebellar hypoplasia (MIM #619286)*	AR	(75)

AD : autosomique dominant ; AR : autosomique récessif ; CKM : module kinase ; XL : transmission liée à l'X

1.3.1. Module kinase CDK8

L'ensemble des composants du CKM ont été décrits en pathologie. Outre le syndrome MED13L qui sera développé dans le chapitre 3 de l'introduction, des variations pathogènes ont été rapportées pour *MED13*, *MED12*, *MED12L*, *CDK8* et *CDK19*.

1.3.1.1. Déficience intellectuelle liée à MED13

MED13, paralogue de *MED13L*, dont il partage 53% d'identité de séquence (29), a récemment été décrit en pathologie (Figure 12). Une approche collaborative de partage de données clinico-biologiques a permis la caractérisation du syndrome MED13 (MIM #618009) de forme autosomique dominante. Les 13 patients rapportés, dans la publication princeps, présentaient notamment une déficience intellectuelle légère à modérée, voire liminale, avec des troubles du spectre autistique. Un retard du langage était noté pour l'ensemble des patients (63).

Parmi les variations rapportées, six étaient en lien avec la survenue d'un codon stop prématuré et quatre étaient des faux-sens identifiés dans le domaine protéine N-terminal de MED13, autour de l'acide aminé Thréonine 326 dans le motif phosphodegron de MED13 : p.(Thr326Ile), p.(Thr326del), p.(Pro327Ser) et p.(Pro327Gln). La variation p.(Pro327Ser) a également été décrite chez un patient suspect de syndrome Kabuki (76). Les deux autres faux-sens étaient localisés dans le domaine C-terminal de la protéine : p.(Gln2060Lys) et p.(Ala2064Val) (Figure 12) (63).

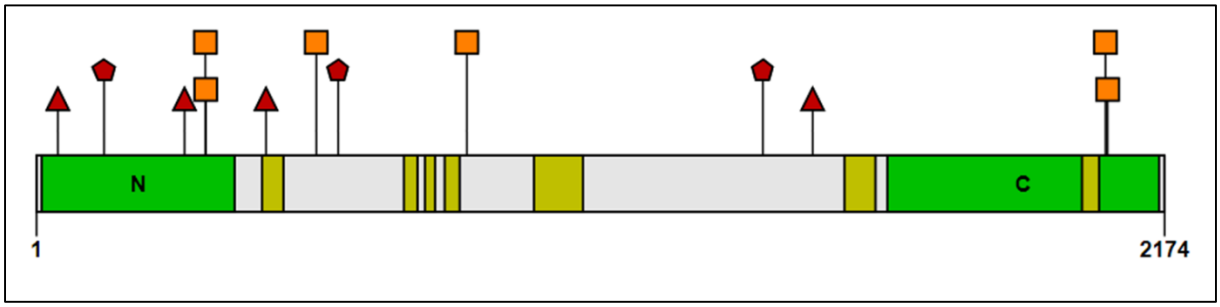


Figure 12. ***Variations pathogènes rapportées dans MED13, paralogue de MED13L.*** *Les variations pathogènes décrites dans MED13 sont en haut (carré orange : faux-sens ; pentagone rouge : non-sens ; triangle rouge : décalage de cadre de lecture). Illustration réalisée avec l'outil IBS (34).*

Les auteurs avaient notamment évalué la quantité de transcrits *MED13* associés à la variation pathogène non-sens NM_005121.2:c.4198C>T; p.(Arg1400*), sans identifier de diminution dans les quantités de transcrits mais avec une absence d'expression protéique, laissant supposer un mécanisme d'échappement à la dégradation des ARNm non-sens pour *MED13* (63).

1.3.1.2. Déficiences intellectuelles associées à *MED12*

MED12 est le premier membre du complexe médiateur dont les altérations ont été rapportées en pathologie. Quatre spectres phénotypiques chevauchant ont été décrits en lien avec *MED12* : le syndrome d'Opitz-Kaveggia ou FG syndrome, le syndrome de Lujan-Fryns, le syndrome d'Ohdo, et des déficiences intellectuelles non syndromiques liées à l'X chez les garçons (64–66,77). Dans le FG syndrome, les patients présentent notamment une déficience intellectuelle associée à une hypotonie, une macrocéphalie et une agénésie partielle du corps calleux (66). Le syndrome de Lujan-Fryns est, en plus, marqué par une grande taille et un habitus marfanoïde (64). Le phénotype associé au syndrome d'Ohdo est caractérisé par une déficience intellectuelle, des traits particuliers du visage dont un blépharo-phimosis et un visage triangulaire (65). Plus récemment, un spectre phénotypique chez les filles a été rapporté chez des patients présentant des variations perte-de-fonctions hétérozygotes *de novo* dans *MED12* (78).

Les variations pathogènes de *MED12* sont responsables d'une modulation de l'expression des gènes associés à une expression précoce, tels que les facteurs de transcriptions *JUN*, *FOS* et *ERG1*. Le profil de dérégulation est fonction de la variation de *MED12* et pourrait expliquer la variabilité des phénotypes identifiées (79).

1.3.1.3. Déficience intellectuelle associée à *MED12L*

Par une approche collaborative de partage de données *via* la plateforme GeneMatcher (80), Nizon et collaborateurs ont ainsi pu décrire l'implication de variations hétérozygotes de *MED12L* en pathologie. Les patients présentaient une déficience intellectuelle légère à modérée, un retard de langage, des troubles du comportement, du sommeil et une hypotonie (24). Une approche fonctionnelle de « *recovery of RNA synthesis* » a été réalisée avec l'observation d'une réduction significative de la synthèse d'ARN après 24h d'exposition aux radiations UV de fibroblastes de patients porteurs de variations candidates dans *MED12L* (24).

1.3.1.4. Déficience intellectuelle en lien avec *CDK8*

Les altérations pathogènes hétérozygotes faux-sens du gène *CDK8* ont été localisées dans la poche de liaison à l'ATP du domaine kinase de CDK8, touchant notamment la boucle riche en glycine de la protéine et impliquant des résidus conservés dans le paralogue CDK19 (67). Le phénotype, décrit par les auteurs, comprend notamment une déficience intellectuelle légère à modérée, une hypotonie, des troubles du comportement et des particularités du visage (67).

L'étude de l'impact des variations faux-sens a été menée sur une cible de phosphorylation de CDK8, la phosphorylation de la Sérine 727 de la protéine STAT1 « STAT1-Ser727 » (*signal transducer and activator of transcription*) dont les modifications de l'état de phosphorylation participent à la régulation de la réponse interféron (81). la déficience intellectuelle liée à des altérations de *CDK8*, Calpena et collaborateurs ont constaté une diminution significative de la phosphorylation de STAT1 après introduction de plasmides porteurs des différentes variations faux-sens candidates dans un lignée cellulaire double KO pour CDK8 et son paralogue CDK19 (67).

1.3.1.5. Déficience intellectuelle liée à *CDK19*

Le paralogue de CDK8, CDK19, qui partage 77% d'identité de séquence dont 97% d'identité au domaine kinase a été très récemment impliqué en pathologie (69). Huit variations faux-sens hétérozygotes *de novo* situées dans le domaine kinase ont été rapportées chez des patients présentant une déficience intellectuelle sévère, un retard de langage, une hypotonie et des particularités du visage. Une approche fonctionnelle similaire à CDK8 a été proposée pour valider les variations pathogènes faux-sens (67), les auteurs ont identifié une diminution de la phosphorylation de la cible « STAT1-Ser727 ». A partir d'un modèle animal, poisson zèbre, ils ont pu noter des variations morphologiques avec une petite taille et des structures antérieurs hypoplasiques (69). Parmi les variations étudiées, la variation pathogène faux-sens p.(Tyr32His) ne semble pas entrainer de perte d'interaction avec les protéines partenaires du CKM : cycline C, MED12 et MED13. Les auteurs posaient l'hypothèse de deux mécanismes à l'origine des altérations de CDK19, par perte ou gain de fonction (69).

1.3.2. Head Module

1.3.2.1. Microcéphalie postnatale progressive en lien avec *MED17*

Les variations homozygotes ou hétérozygotes composites ont été identifiées chez des patients présentant une microcéphalie postnatale progressive, une déficience intellectuelle, une épilepsie, une atrophie cérébrale et une ataxie (70,82,83). Les variations candidates décrites n'ont pas fait l'objet d'une évaluation fonctionnelle. MED17 possède un rôle central au sein du complexe médiateur. Il intervient notamment dans la formation et dans le contrôle de l'intégrité du complexe ARN Polymérase II par son interaction avec la sous-unité POLR2C (dénommée également RPB3) du complexe Pol II. Cette interaction permet la stabilisation *in vivo* de l'association entre le complexe Pol II et le promoteur ciblé (14). A partir de modélisations des variations pathogènes dans la structure cristallographie de srb4 (med17) chez *Schizosaccharomyces pombe,* il semble que les variations faux-sens soient localisées dans les domaines terminaux de srb4 (med17) en lien avec l'interaction directe avec le « *middle module* » (84).

1.3.2.2. Troubles neuro-évolutifs avec mouvements anormaux en lien avec *MED20*

Le gène *MED20* a fait l'objet d'une publication d'une variation faux-sens homozygote c.341G>C ; p.(Gly114Ala) ségrégant dans une famille avec deux probants présentant une spasticité, une dystonie, une microcéphalie post-natale, une cataracte bilatérale et une dégénérescence progressive des ganglions basaux (71). Cette implication n'a, à ce jour, pas fait l'objet d'une confirmation par une seconde publication. Aucune validation fonctionnelle n'a été réalisée.

1.3.3. Tail Module

1.3.3.1. Déficience intellectuelle en lien avec *MED23*

Comme pour *MED17*, des variations homozygotes et hétérozygotes composites de *MED23* ont été décrites chez des patients présentant une déficience intellectuelle associée à une microcéphalie et une hypotonie (72,85–87). Une épilepsie sensible au régime cétogène est rapportée chez un patient (86). Une validation fonctionnelle a été proposée pour les variations c.[4006C>T];[3656A>G] ; p.[(Arg1336*)];[(His1219Arg)], avec mise en évidence en RT-PCR d'une diminution de l'expression de *JUN* et une surexpression de *FOS* après traitement de fibroblastes à un agent mitogène (72,85).

Grâce à la caractérisation de la structure 3D de MED23, protéine la plus large du « *tail module* », la variation faux-sens homozygote touchant l'arginine Arg617 (72) a pu être localisée sur la face convexe de MED23 dans un pont salin très conservé avec la glutamate Glu565 au centre du corps de MED23. Une rupture de la structure altérant les régions d'interactions de MED23 est possible (88).

1.3.3.2. Syndrome Basel-Vanagaite-Smirin-Yosef associé à *MED25*

Le syndrome de Basel-Vanagaite-Smirin-Yosef a été décrit chez plusieurs patients présentant une déficience intellectuelle sévère, une absence de langage, une épilepsie, une microcéphalie et des malformations oculaires – cardiaques – urogénitales – squelettiques (74,89). Une approche fonctionnelle a été proposée pour valider la variation homozygote récurrente pathogène c.116A>G ; p.(Tyr39Cys). Celle-ci est localisée dans le domaine homologue « facteur von Willebrand de type A » de la protéine MED25 qui permet son recrutement dans le complexe médiateur (90). Une étude de l'association entre MED25 et le complexe médiateur par co-immunoprécipitation avait permis de conclure à une altération de son recrutement, avec une absence de co-précipitation avec les sous-unités MED1, MED6, MED14, MED16, MED23, MED24 et CDK8 (74).

1.3.3.3. Troubles du neuro-développement en lien avec *MED27*

Une très récente étude de cohorte a permis de recruter 16 patients porteurs de variations homozygotes ou hétérozygotes composites dans *MED27*. Le phénotype, décrit dans une publication à ce jour, est celui d'un trouble sévère du neuro-développement avec déficience intellectuelle, hypotonie majeure, hypoplasie cérébelleuse, épilepsie, cataracte et dans 50% des cas une microcéphalie (75). Les auteurs n'ont pas proposé de validation fonctionnelle.

1.3.3.4. Spectre des pathologies du *tail module*

Les phénotypes rapportés pour *MED17*, *MED20*, *MED23*, *MED25* et *MED27* sont caractérisés par leur mode de transmission autosomique récessif et un spectre malformatif commun avec une déficience intellectuelle souvent sévère, une microcéphalie post-natale et des malformations oculaires (Figure 13).

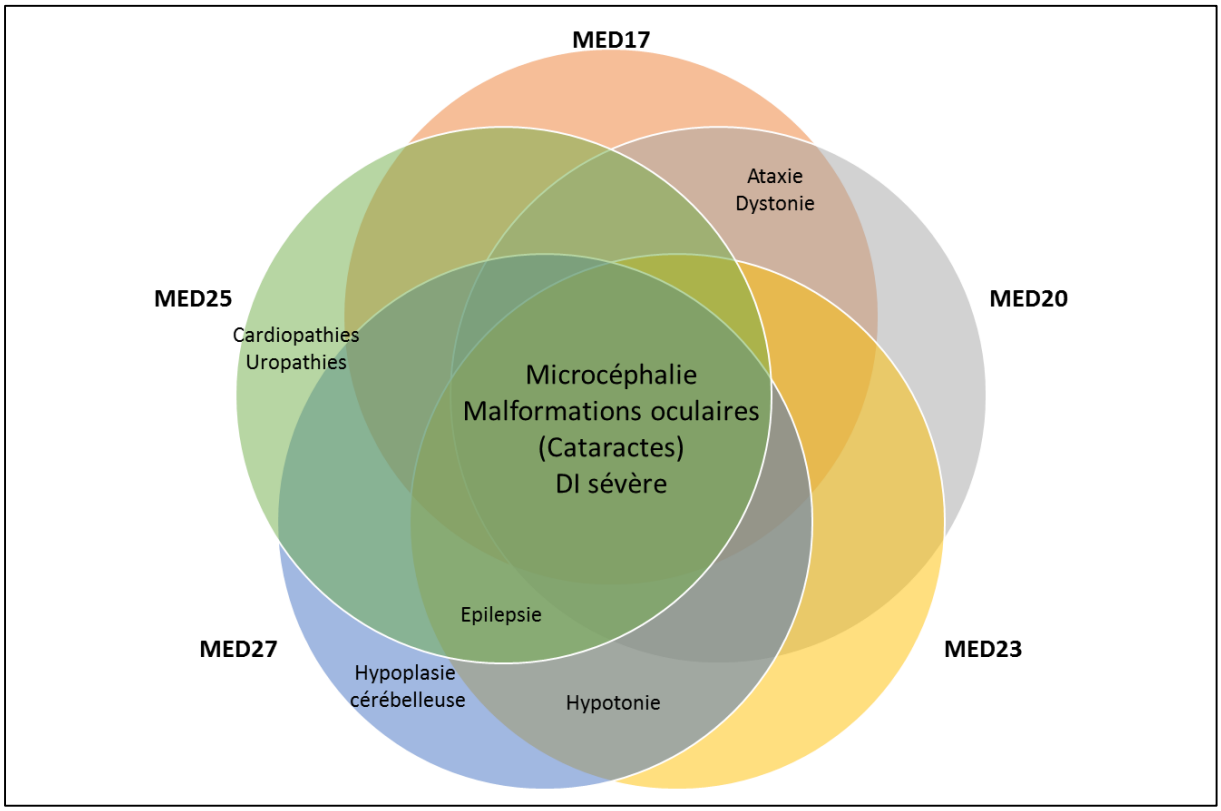


Figure 13. ***Diagramme de Venn présentant une synthèse des phénotypes associés à MED17-MED20-MED23-MED25-MED27.***

1.4. Syndrome MED13L

1.4.1. Historique du syndrome

Initialement, l'implication de *MED13L* en pathologie humaine a été suspectée à travers l'étude d'une cohorte de patients présentant une cardiopathie cono-troncale à type de transposition des gros vaisseaux (TGV). Les auteurs avaient identifié une patiente présentant à la fois une TGV, une déficience intellectuelle sévère, une hypotonie et une absence de langage. Les explorations cytogénétiques conventionnelles, confirmées par *hybridation in situ* en fluorescence avaient permis la caractérisation d'une translocation réciproque cytogénétique équilibrée t(12;17)(q24.1;q21) avec un point de cassure dans l'intron 1 du gène *MED13L*. Partant de ce constat, les auteurs s'étaient orientés sur la recherche de variations rares du gène *MED13L* dans une cohorte de patients présentant une TGV. Ainsi, quatre variations faux-sens hétérozygotes rares avaient été identifiées par séquençage ciblé dans leur cohorte. Le lien entre ces variations et la TGV peut aujourd'hui être ré-évalué devant leurs fréquences dans les bases de données témoins, gnomAD, non disponibles en 2003 lors de la publication (35).

L'association entre altérations de *MED13L* et TGV n'a, par la suite, pas été confirmée, mais le développement des technologies de cytogénétique moléculaire avec le déploiement de la CGH-array sur des cohortes de patients présentant des troubles du neuro-développement a permis dans un premier temps la caractérisation de deux microdélétions hétérozygotes intra-géniques de 17 kb, emportant l'exon 2 de *MED13L*, et de 115 kb, emportant les exons 3 et 4. Le phénotype des patients était alors marqué par une déficience intellectuelle modérée, une hypotonie, des troubles sévères du langage et des particularités faciales (61). Le spectre phénotypique a ensuite pu être confirmée par d'autres descriptions de cas isolés sans cardiopathie cono-troncale (91–93).

A ce jour, les variations pathogènes conduisant à une altération de *MED13L* sont rapportées parmi les principales causes de déficience intellectuelle d'origine génétique. Les larges études de séquençage à haut débit par stratégie d'exome, dans les troubles du neuro-développement, menées notamment par le consortium *Deciphering Developmental Disorders*, ont permis l'identification de l'implication de *MED13L* dans environ 0.45% des diagnostics moléculaires (55,94,95). Les altérations du gène *MED13L* apparaissent ainsi parmi les 10 étiologies génétiques les plus fréquentes dans l'étude menée par le *Deciphering Developmental Disorders* en 2017 (55). Les principales étiologies génétiques rapportées sont les altérations

d'*ANKRD11* dans le syndrome KBG (96), *ARID1B* dans le syndrome de Coffin-Siris de type 1 (97), *KMT2A* dans le syndrome de Wiedemann-Steiner (98), *DDX3X* dans le syndrome Snijders Blok lié à l'X (99), *ADNP* dans le syndrome d'Helsmoortel-van der Aa (100) et *SCN2A* dans les encéphalopathies épileptiques et déficiences intellectuelles (101). A la même fréquence que *MED13L*, on retrouve également *DYRK1A* dans les déficiences intellectuelles autosomiques dominantes de type 7 (102), *EP300* dans le syndrome de Rubinstein-Taybi de type 2 (103) et enfin *SETD5* dans les déficiences intellectuelles autosomiques dominantes de type 23 (104).

1.4.2. Phénotypes associés

Le phénotype des patients présentant un Syndrome MED13L est repris dans le premier article de la thèse (62). Une synthèse de la littérature a été récemment réalisée par Tørring et collaborateurs sur 69 patients reprenant les données rapportées par le l'étude collaborative menée par l'EA7364 RADEME (62,105). Les données sont reprises dans le Tableau 2 et sont discutées dans le premier article de la thèse.

Tableau 2. ***Synthèse du phénotype rapporté dans le syndrome MED13L.*** *Adapté d'après (105)*

Caractéristiques		Nombre de patients	%
Déficience intellectuelle		69/69	100%
Retard moteur		69/69	100%
Retard de langage		68/69	99%
Anomalies des mains/pieds		31/69	45%
Hypotonie		46/66	70%
Anomalies ophtalmologies		23/69	33%
Anomalies IRM cérébrale		26/58	45%
Cardiopathies congénitale		12/64	19%
	Cardiopathies complexes	5/12	-
	Persistance foramen ovale	7/12	-
Troubles du spectre autistique		16/69	23%
Dysmorphie			
	Point du nez large	50/67	75%
	Bouche hypotonique	40/65	62%
	Base du nez élargie	19/33	58%
	Oreilles basses implantées	17/33	52%

	Macrostomie	14/33	42%
	Fissures palpébrales en haut/dehors	26/65	40%
	Macroglossie	9/33	27%
	Pincement bitemporal	8/33	24%
	Brachycéphalie	7/33	21%
	Sourcils horizontaux	7/33	21%
	Oreilles larges	6/33	18%

1.4.3. Variations pathogènes

L'extraction des données concernant les variations pathogènes, probablement pathogènes ou candidates impliquant *MED13L* dans la littérature, a permis l'identification de 134 variations différentes. La très grande majorité concerne le Syndrome MED13L (79.9%, n=107/134), puis des variations candidates identifiées dans des cohortes de patients présentant un trouble du spectre autistique (16.4%, n=22/134), et de rares variations identifiées chez des patients présentant des cardiopathies (3.7%, n=5/134) (Tableau 3).

Tableau 3. Variations pathogènes SNV/INDEL de MED13L rapportées en pathologies

HGVCs	HGVSp	Phénotypes HGMD	Références
c.43G>T	p.Glu15*	*MED13L-related intellectual disability*	(106)
c.173T>A	p.Leu58*	*MED13L-related intellectual disability*	(62)
c.187T>C	p.Cys63Arg	*MED13L-related intellectual disability*	(107)
c.263G>A	p.Trp88*	*MED13L-related intellectual disability*	(108)
c.329G>A	p.Trp110*	*MED13L-related intellectual disability*	(51,94)
c.656C>T	p.Thr219Met	*Autism spectrum disorder*	(109)
c.752A>G	p.Glu251Gly	*Congenital heart defect*	(35)
c.758C>T	p.Ser253Leu	*Autism spectrum disorder*	(109)
c.1361C>G	p.Ser454*	*MED13L-related intellectual disability*	(110)
c.1367C>G	p.Ser456Cys	*Autism spectrum disorder*	(109)
c.1591G>A	p.Val531Met	*MED13L-related intellectual disability*	(111)
c.1690C>T	p.Arg564*	*MED13L-related intellectual disability*	(51,94)
c.1718C>T	p.Pro573Leu	*MED13L-related intellectual disability*	(112)
c.1889A>G	p.Glu630Gly	*Autism spectrum disorder*	(109)
c.1904G>A	p.Ser635Asn	*MED13L-related intellectual disability*	(113)
c.1910G>A	p.Arg637His	*Autism spectrum disorder*	(109)
c.2060A>G	p.Gln687Arg	*Autism spectrum disorder*	(109)

c.2065C>T	p.Gln689*	*MED13L-related intellectual disability*	(62)
c.2071C>T	p.Gln691*	*MED13L-related intellectual disability*	(105)
c.2117G>A	p.Gly706Glu	*MED13L-related intellectual disability*	(114,115)
c.2210C>T	p.Thr737Met	*Autism spectrum disorder*	(109)
c.2312C>G	p.Ala771Gly	*Autism spectrum disorder*	(109)
c.2524C>T	p.Arg842*	*MED13L-related intellectual disability*	(116,117)
c.2530C>A	p.Leu844Ile	*MED13L-related intellectual disability*	(118)
c.2579A>G	p.Glu860Gly	*MED13L-related intellectual disability*	(119,120)
c.2591T>A	p.Met864Lys	*MED13L-related intellectual disability*	(121)
c.2597C>T	p.Pro866Leu	*MED13L-related intellectual disability*	(62)
c.2605C>T	p.Pro869Ser	*MED13L-related intellectual disability*	(62,122)
c.2605C>A	p.Pro869Thr	*MED13L-related intellectual disability*	(106)
c.2928G>A	p.Trp976*	*Autism Spectrum Disorder*	(53)
c.2930C>T	p.Ala977Val	*MED13L-related intellectual disability*	(62)
c.2974A>C	p.Thr992Pro	*Autism spectrum disorder*	(109)
c.2996A>G	p.Asn999Ser	*Autism spectrum disorder*	(109)
c.3392G>A	p.Cys1131Tyr	*MED13L-related intellectual disability*	(105)
c.3521A>G	p.Tyr1174Cys	*Autism spectrum disorder*	(53,123)
c.3656A>C	p.Gln1219Pro	*Autism spectrum disorder*	(109)
c.4076G>A	p.Trp1359*	*MED13L-related intellectual disability*	(51–53,57,123–125)
c.4247G>A	p.Arg1416His	*MED13L-related intellectual disability*	(126)
c.4420A>T	p.Lys1474*	*MED13L-related intellectual disability*	(93)
c.4439T>C	p.Val1480Ala	*Autism spectrum disorder*	(109)
c.5081C>T	p.Thr1694Ile	*Autism spectrum disorder*	(109)
c.5173C>T	p.Gln1725*	*MED13L-related intellectual disability*	(105,127)
c.5278C>T	p.Arg1760*	*MED13L-related intellectual disability*	(62)
c.5282C>T	p.Pro1761Leu	*MED13L-related intellectual disability*	(128)
c.5326C>T	p.Pro1776Ser	*Autism spectrum disorder*	(109)
c.5413A>G	p.Ile1805Val	*Autism spectrum disorder*	(109)
c.5615G>A	p.Arg1872His	*Congenital heart defect*	(35,129)
c.5681G>A	p.Trp1894*	*MED13L-related intellectual disability*	(105)
c.5695G>A	p.Gly1899Arg	*MED13L-related intellectual disability*	(51,55,116,117,130)
c.5698C>T	p.Arg1900*	*MED13L-related intellectual disability*	(131)
c.5764C>T	p.Gln1922*	*MED13L-related intellectual disability*	(62)
c.5768C>T	p.Thr1923Ile	*Autism spectrum disorder*	(109)
c.5797C>T	p.Arg1933Trp	*Autism spectrum disorder*	(109)
c.5849T>C	p.Leu1950Pro	*MED13L-related intellectual disability*	(132)
c.6005C>T	p.Ser2002Leu	*MED13L-related intellectual disability*	(51,62)
c.6068A>G	p.Asp2023Gly	*Congenital heart defect*	(35)
c.6155A>C	p.Asn2052Thr	*MED13L-related intellectual disability*	(113)

c.6199G>A	p.Val2067Met	*Autism spectrum disorder*	(109)
c.6280C>T	p.Pro2094Ser	*Autism spectrum disorder*	(133)
c.6280C>A	p.Pro2094Thr	*MED13L-related intellectual disability*	(110)
c.6392C>T	p.Ser2131Leu	*Congenital heart disease*	(51)
c.6485C>A	p.Thr2162Lys	*MED13L-related intellectual disability*	(134)
c.6485C>T	p.Thr2162Met	*MED13L-related intellectual disability*	(62,135)
c.6488C>T	p.Ser2163Leu	*MED13L-related intellectual disability*	(62)
c.6530C>A	p.Ser2177Tyr	*MED13L-related intellectual disability*	(62)
c.6589G>C	p.Val2197Leu	*Autism spectrum disorder*	(109)
c.5364+1G>T	p.?	*Autism Spectrum Disorder*	(53)
c.480-1G>T	p.?	*MED13L-related intellectual disability*	(91)
c.1009+1G>C	p.?	*MED13L-related intellectual disability*	(62)
c.2345-3C>G	p.?	*MED13L-related intellectual disability*	(62)
c.2570-4_2574del	p.?	*MED13L-related intellectual disability*	(136)
c.4956-2A>C	p.?	*Congenital heart defect*	(137)
c.5588+1G>A	p.?	*MED13L-related intellectual disability*	(62,138)
c.5364+1dup	p.?	*Autism Spectrum Disorder*	(51,53,57,125,139)
c.6225+1G>A	p.?	*MED13L-related intellectual disability*	(62)
c.6226-1G>C	p.?	*MED13L-related intellectual disability*	(140)
c.257del	p.Phe86Serfs*9	*MED13L-related intellectual disability*	(118)
c.500del	p.Phe167Serfs*25	*MED13L-related intellectual disability*	(141)
c.830_845del	p.Arg277Glnfs*5	*MED13L-related intellectual disability*	(62)
c.1077_1093del	p.Met359Ilefs*38	*MED13L-related intellectual disability*	(142)
c.1416_1419del	p.Gln473Lysfs*11	*Autism Spectrum Disorder*	(53)
c.1708_1709del	p.Ser570Phefs*27	*MED13L-related intellectual disability*	(62,143–145)
c.1994_1997del	p.Glu665Alafs*8	*MED13L-related intellectual disability*	(146)
c.1996_1997del	p.Ser666Hisfs*11	*MED13L-related intellectual disability*	(147)
c.2340_2343del	p.Thr781Metfs*19	*MED13L-related intellectual disability*	(62)
c.2393_2396del	p.Thr798Argfs*2	*MED13L-related intellectual disability*	(51,55,94)
c.2395_2396del	p.Gln799Glyfs*10	*Autism spectrum disorder*	(123)
c.2504del	p.Pro835Leufs*46	*MED13L-related intellectual disability*	(120)
c.3765del	p.Cys1256Valfs*2	*MED13L-related intellectual disability*	(93)
c.3911_3912del	p.Val1304Alafs*60	*MED13L-related intellectual disability*	(62)
c.3942_3943del	p.Ile1315Glnfs*49	*MED13L-related intellectual disability*	(62,148,149)
c.4106del	p.Gly1369Glufs*19	*MED13L-related intellectual disability*	(51,55,94)
c.4452del	p.Phe1484Leufs*27	*MED13L-related intellectual disability*	(62)
c.5152_5153del	p.Met1718Glufs*21	*MED13L-related intellectual disability*	(62)
c.5269del	p.Gln1757Serfs*26	*MED13L-related intellectual disability*	(150)
c.5444del	p.Thr1815Serfs*37	*MED13L-related intellectual disability*	(141)
c.5453_5454del	p.Glu1818Glyfs*27	*MED13L-related intellectual disability*	(131)

c.5949_5950del	p.Gln1984Alafs*31	*MED13L-related intellectual disability*	(92)
c.6118_6125del	p.Gly2040Asnfs*32	*MED13L-related intellectual disability*	(62,151)
c.6260del	p.Pro2087Glnfs*4	*MED13L-related intellectual disability*	(152)
c.6579del	p.Val2194Serfs*14	*MED13L-related intellectual disability*	(153)
c.124dup	p.Asp42Glyfs*13	*MED13L-related intellectual disability*	(135)
c.607dup	p.Ser203Phefs*32	*MED13L-related intellectual disability*	(93)
c.1152dup	p.Ile385Hisfs*18	*MED13L-related intellectual disability*	(62)
c.1903dup	p.Ser635Lysfs*8	*MED13L-related intellectual disability*	(62)
c.2399dup	p.Thr801Asnfs*9	*MED13L-related intellectual disability*	(127)
c.3719dup	p.Asn1241Glufs*9	*Autism spectrum disorder*	(109)
c.6274dup	p.Gln2092Profs*16	*MED13L-related intellectual disability*	(105)
c.6284dup	p.Ala2096Glyfs*12	*MED13L-related intellectual disability*	(62)
c.3276delinsAA	p.Asn1093Lysfs*32	*MED13L-related intellectual disability*	(132)

Autism spectrum disorder : Trouble du spectre autistique ; Congenital Heart Defect : Cardiopathie congénitale ; MED13L-related intellectual disability : Syndrome MED13L.

Parmi les variations identifiées chez des patients suspectés du Syndrome MED13L, une très grande majorité de variations conduisant à une haplo-insuffisance potentielle sont rapportées (Figure 14). Nous dénombrons ainsi 31.0% de variations décalage de cadre de lecture (n=34/107), 14.0% de variations non-sens (n=15/107), 9.2% de variations d'épissage (n=10/107) et 22.4% de variations cytogénétiques (microdélétions et microduplications, translocations avec points de cassure dans *MED13L*) (n=24/107). Les variations faux-sens candidates, de description plus récentes, ne représentent que 22.4% de variations rapportées (n=24/107), mais, en revanche, sont très présentes dans les publications les plus récentes (62,105,120).

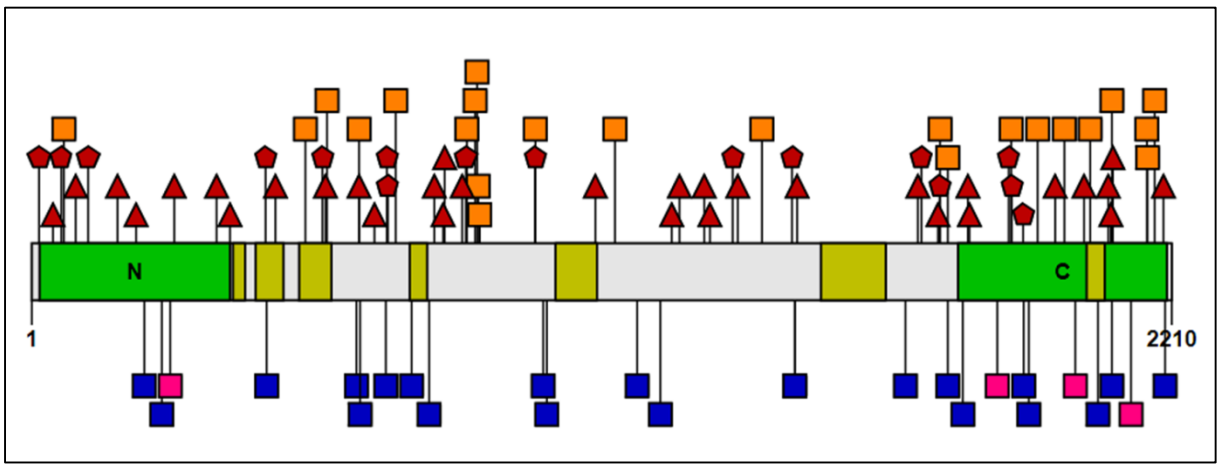


Figure 14. ***Localisation des variations pathogènes et candidates SNV/INDEL de MED13L rapportées dans la littérature.*** *Les variations pathogènes décrites dans le syndrome MED13L sont en haut (carré orange : faux-sens ; pentagone rouge : non-sens ; triangle rouge : décalage de cadre de*

lecture), les variations faux-sens décrites dans les cardiopathies et les troubles du spectre autistique sont en bas (respectivement carré rose et carré bleu). Illustration réalisée avec l'outil en ligne IBS (34).

Aucune des variations faux-sens pathogènes n'est située dans les motifs signatures « LXXLL » ou « Leucine-zipper » de MED13L.

1.4.4. Survenue

La survenue des variations pathogènes de *MED13L* est très majoritairement *de novo*, en lien avec une possible altération du processus méiotique maternelle ou parternelle. Cependant, quelques rares cas sont associés à un mosaïcisme parental germinal. La variation récurrente p.(Gly1899Arg) a récemment été décrite dans une fratrie. Sa transmission était paternelle, avec une mosaïque estimée à 30% dans le tissu spermatique, et une absence de variation dans le tissu sanguin (130). Une délétion intragénique de 125.1 kb emportant les exons 3 à 14 de *MED13L* a également été décrite. La variation était de transmission maternelle avec une balance allélique estimée à 1% en séquençage haut-débit dans le sang maternel (154). Nous avons aussi décrit une variation d'épissage partagée par une fratrie c.5588+1G>A, sans identification de l'origine parentale par séquençage haut-débit sur les prélèvements sanguins des parents (62).

1.4.5. Données fonctionnelles

Une première modélisation a porté sur l'étude combinée d'un KO de *med13b* chez le poisson zèbre et d'un KO de *MED13L* dans un modèle cellulaire de dérivés de progéniteurs neuronaux. La protéine *med13b*, orthologue de MED13L, présente 65% de similarités de séquence avec MED13L. Les auteurs ont observé une altération de la migration des cellules neuronales de la crête neurale associée à une altération de la différenciation neuronale précoce chez le poisson zèbre. L'étude transcriptomique des dérivés neuronaux KO pour *MED13L* mettait en évidence une modification significative de l'expression de 1117 gènes cibles dont les plus dérégulés étaient *SP8*, impliqués dans le développement cranio-facial et *FGFR3*, répresseur de la croissance osseuse (47).

Le *knockdown* de Med13l *Mus musculus* dans des embryons murins entraine une létalité embryonnaire avant pré-implantation. Miao et collaborateurs ont également observé que Med13l *Mmu* peut partiellement compenser le KO de Med13 *Mmu*. Les auteurs posaient l'hypothèse d'une non-redondance fonctionnelle de Med13l et Med13 dans le développement embryonnaire murin (155). Le modèle murin a, par la suite, permis d'établir une cartographie d'expression de med13l au sein du système nerveux central. Hamada et collaborateurs ont démontré une localisation anténatale préférentielle de Med13l *Mmu* dans les noyaux des neurones du cortex cérébral, puis dans le cytoplasme en post-natal, avec une expression importante dans la région ventriculaire du cortex cérébral. Cette constatation de temporalité a également été faite sur les neurones hippocampiques. Les auteurs ont également noté un rôle majeur possible de Med13l *Mmu* lors des étapes précoces du développement embryonnaire cérébral et la formation du réseau synaptique (12).

Très peu d'approches fonctionnelles ont été proposées pour caractériser les variations pathogènes identifiées chez des patients dans le cadre du Syndrome MED13L. La quantification de l'ARN d'un patient porteur d'une micro-délétion de 125kb emportant les exons 3 et 4 de *MED13L* (Chr12(GRCH37):g.116466743_116591383del) a montré la présence d'ARN associé à l'allèle altérée, en lien avec un potentiel mécanisme d'échappement à la dégradation des ARNm non-sens (92). Une diminution de l'expression de *MED13L* mutée chez un patient porteur d'une variation faux-sens hétérozygote c.5282C>T; p.Pro1761Leu avait également été caractérisée. Les auteurs soulignaient une diminution significative de l'ARNm associée à *MED13L* par RT-qPCR chez le patient porteur de la variation (128).

Dans le cadre de la validation fonctionnelle de variations pathogènes de *MED12L*, une lignée cellulaire d'un patient porteur d'une variation pathogène de *MED13L* a été utilisée comme contrôle dans une étude de l'activité transcriptionnelle post-exposition aux radiations UV. Cette lignée contrôle a été associée au pourcentage le plus faible de récupération de l'activité de synthèse d'ARN après expositions UV à 20h par rapport aux lignées contrôles de *MED12* et aux fibroblastes de patients *MED12L*, en lien avec une altération sévère de la transcription (24).

Les approches fonctionnelles proposées mettent en avant deux difficultés. D'une part, les explorations de l'ARN par RT-qPCR, même si elles sont simples à proposer, nécessitent de nouveaux prélèvements et partent de l'hypothèse que les variations pathogènes conduisent à une haplo-insuffisance de *MED13L*. D'autre part, les explorations complexes sur poisson zèbre et approches transcriptomiques s'implémentent mal dans la routine diagnostique.

1.4.6. Données *in silico*

La difficulté d'obtenir des données fonctionnelles, transférables en routine diagnostique, incite les équipes à inclure des annotations bio-informatiques combinées, pour l'interprétation des données moléculaires et phénotypiques, dans le but d'affiner la classification des variations candidates identifiées. Cette approche algorithmique est notamment appliquée en présence de variations candidates faux-sens, et est intégrée dans les guides de bonnes pratiques, pour la classification des variations, proposés par les groupes ACMG/AMP (*The American College of Medical Genetics and American College of Pathologists*) (156,157).

Cette approche algorithmique est notamment basée sur les conservations inter-espèces des séquences nucléotidiques et protéiques (158), la localisation dans une structure secondaire, une région ou un domaine protéique connu (159), la comparaison des propriétés physico-chimiques des acides aminés impliqués dans la substitution (160) ou la fréquence dans les bases de données de populations contrôles (161). Ainsi, des scores de prédictions ont pu être établis par exemple par les outils algorithmiques PolyPhen2 et SIFT afin de classer les variations candidates en « bénignes » ou « délétères » selon leurs propriétés (162,163).

La nécessité d'optimiser cette approche algorithmique a entrainé une évolution des outils de prédictions avec le développement de méta-prédicteurs, combinant parfois plus d'une centaine de paramètres. Ces outils, tels que CADD, REVEL ou BayesDel, définissent un score continu, dont le seuil est à définir par chaque utilisateur (160,164,165).

MED13L présente peu de variations dans la population générale pour une protéine très conservée entre les espèces (29,166). Aussi, de très larges blocs protéiques dans MED13L

sont prédits comme intolérants aux substititions (Figure 15). Dans ce contexte, les scores de prédictions

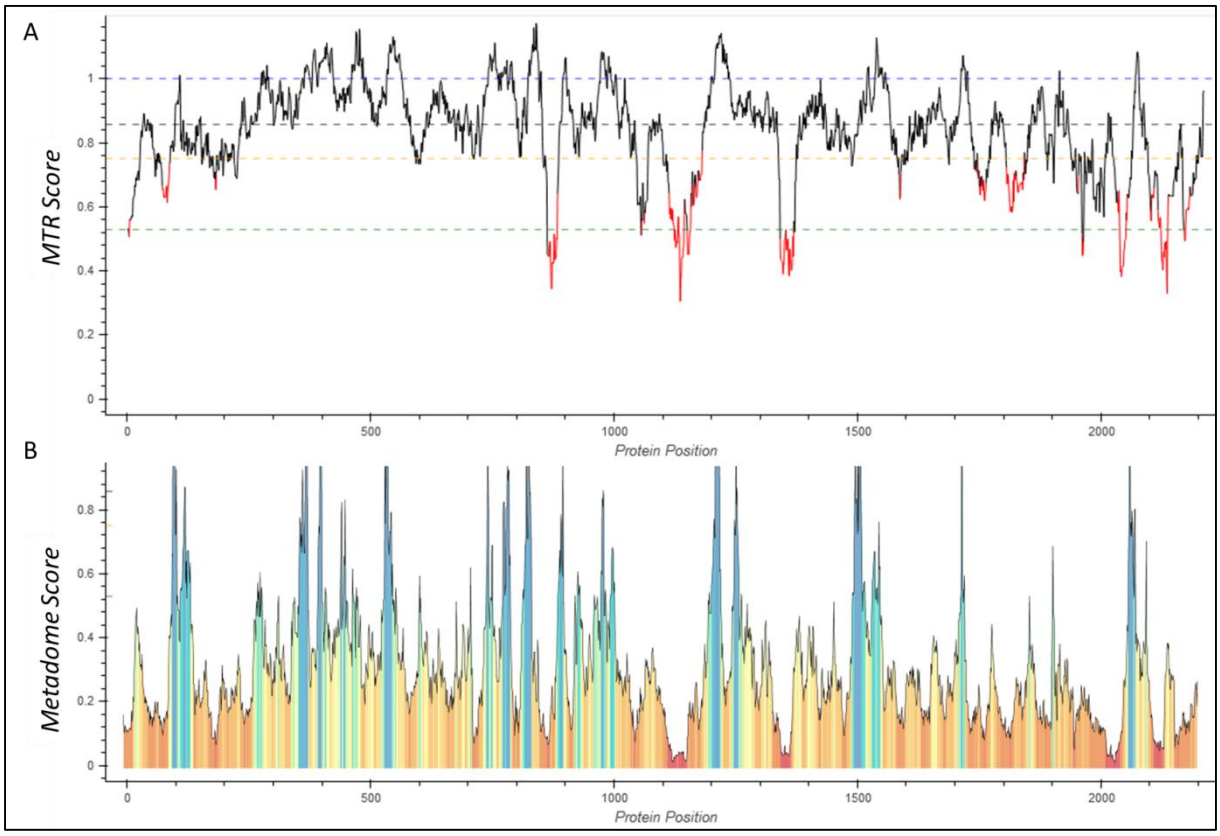


Figure 15. ***Prédictions des régions critiques pour les substitutions dans la protéine MED13L*** *(A) Base MTR, les régions protéiques identifiées en rouge sur le graphique présentent un éloignement significatif par rapport aux variations neutres. Graphique adapté de (167) ; (B) Base Metadome, les régions identifiées en orange/rouge, sont associées à un score élevé d'intolérance à la substitution. Graphique adapté de (168).*

2. OBJECTIFS

Les descriptions nombreuses de variations faux-sens candidates dans le gène *MED13L* nous ont conduit à nous interroger sur plusieurs points :

- **Existe-il des spécificités phénotypiques associées aux variations faux-sens de *MED13L* ?**

Initialement, très peu de variations candidates faux-sens candidates ont été rapportées dans *MED13L*. Les descriptions de patients concernaient très majoritairement des variations hétérozygotes non-sens ou par décalage de cadre de lecture en lien avec une possible haplo-insuffisance. L'application des stratégies de séquençage non ciblées dans les troubles du neuro-développement a conduit de nombreuses équipes à identifier des variations faux-sens dans des spectres phénotypiques, parfois atypiques par rapport aux connaissances.

Nous avons conduit une étude de cohorte sur 36 patients porteurs de variations pathogènes dans *MED13L* dont 7 variations faux-sens *de novo*. Ce travail publié dans la revue *NEUROGENETICS* en mai 2018 a permis de doubler le nombre de descriptions de variations pathogènes décrites en littérature et d'évoquer les premières spécificités en lien avec les variations faux-sens.

- **Quel est le profil *in silico* des variations faux-sens pathogènes en lien avec le Syndrome MED13L ?**

En parallèle du développement d'une approche fonctionnelle pour l'étude des variations faux-sens, nous avons constitué une base de données autour des annotations associées aux variations pathogènes décrites dans la littérature, les bases de données (ClinVar, LOVD, DECIPHER) et les demandes d'avis. Les informations ont été croisées pour définir un profil lié à la conservation, la co-localisation, les propriétés physico-chimiques des acides aminés impliqués.

Cette comparaison entre variations pathogènes candidates et variations faux-sens de la littérature est soumis et en révision favorable dans la revue *EUROPEAN JOURNAL OF MEDICAL GENETICS* en novembre 2021. A terme, une publication et un maintien de la base

seront proposés en ligne sur un modèle proche des observatoires des variations de *GRIN2A-GRIN2B* et *MECP2-CDKL5*.

- **Quelles évaluations fonctionnelles pourraient être proposées pour valider la pathogénicité de variations faux-sens candidates identifiées en routine diagnostique ?**

Les difficultés d'interprétation des variations faux-sens de *MED13L* nous ont conduits à nous interroger sur le développement d'explorations fonctionnelles. A ce jour, très peu d'études de validations fonctionnelles ont été proposées pour interpréter les variations de *MED13L*. Nous proposons ici une approche basée sur une des voies de régulation de la protéine MED13L via la voie de l'ubiquitine protéasome.

Cette évaluation fonctionnelle conduite sur 5 variations faux-sens représentatives de la diversité des variations faux-sens pathogènes rapportées dans le syndrome MED13L a été soumise pour publication et est actuellement en cours de révision.

3. RESULTATS

Phénotype assocé au syndrome *MED13L* – Description d'une série de 36 patients

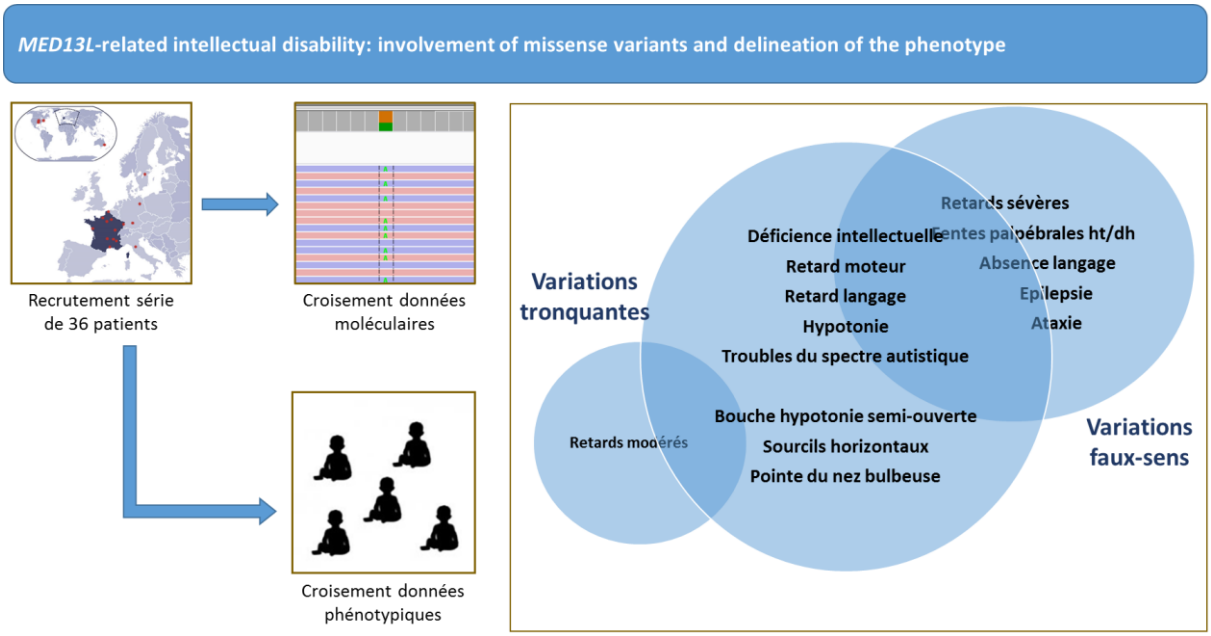
Contexte de l'article :

Les premières données phénotypiques présentant le syndrome MED13L proviennent d'une étude de séquençage ciblé sur *MED13L*, et réalisée dans une série de patients, porteurs d'une cardiopathie cono-troncale. Cette étude de 2003 posait l'hypothèse d'un lien entre une altération de *MED13L* et la transposition des gros vaisseaux (35). Le phénotype a été redéfini 10 ans plus tard par Asadollahi et collaborateurs, avec la description de 2 patients porteurs de microdélétions intragéniques de MED13L et d'un patient porteur d'une microduplication. Ces déséquilibres avaient été identifiés dans le cadre d'exploration de troubles du neurodéveloppement non syndromique (61). Depuis, plusieurs petites séries ont été publiées sur *MED13L* : 2 patients par van Haelst et collaborateurs, 8 patients par Adegbola et collaborateurs, et 3 autres patients par Cafiero et collaborateurs (91–93). Ces cas rapportés concernaient exclusivement des variations tronquantes, et ont permis de préciser le phénotype de MED13L sur le plan neurodéveloppemental.

Avec l'application du séquençage haut-débit dans le laboratoire de routine diagnostique du CHU de Lille, j'ai pu identifier 6 nouvelles variations hétérozygotes *de novo* pathogènes de *MED13L* dont une variation faux-sens p.(Ser2163Leu). Afin de proposer une description du syndrome MED13L sur une large série, j'ai organisé de Juin 2016 à Janvier 2018 un appel à collaboration internationale *via* la plateforme d'échange « GeneMatcher » (80), et *via* le groupe de travail des laboratoires français impliqués dans les diagnostics des déficiences intellectuelles.

Cette étude collaborative a permis de recenser les données moléculaires et phénotypiques de 36 patients dont 7 variations faux-sens. Ce travail a fait l'objet d'une publication dans **Neurogenetics en mai 2018** (62).

J'ai présenté les résultats de cette série en communication orale aux assises de génétique de Nantes en 2018, et en communication affichée à l'ESHG de 2017.


MED13L-related intellectual disability: involvement of missense variants and delineation of the phenotype
Recrutement série de 36 patients
Croisement données moléculaires
Croisement données phénotypiques
Variations tronquantes
Variations faux-sens
Retards sévères
Déficience intellectuelle
Fentes palpébrales ht/dh
Retard moteur
Absence langage
Retard langage
Epilepsie
Hypotonie
Ataxie
Troubles du spectre autistique
Retards modérés
Bouche hypotonie semi-ouverte
Sourcils horizontaux
Pointe du nez bulbeuse

ARTICLE 1 : PHÉNOTYPE ASSOCIÉ AU SYNDROME *MED13L* – DESCRIPTION D'UNE SERIE DE 36 PATIENTS

MED13L-related intellectual disability: involvement of missense variants and delineation of the phenotype

Smol T[1,2], Petit F[2,3], Piton A[4], Keren B[5], Sanlaville D[6], Afenjar A[7], Baker S[8], Bedoukian EC[10], Bhoj EJ[8], Bonneau D[9], Boudry-Labis E[1], Bouquillon S[1], Boute-Benejean O[2,3], Caumes R[3], Chatron N[6], Colson C[2,3],Coubes C[11], Coutton C[12], Devillard F[12], Dieux-Coeslier A[2,3], Doco-Fenzy M[13], Ewans LJ[14], Faivre L[15,16], Fassi E[17], Field M[18], Fournier C[4], Francannet C[19], Genevieve D[11], Giurgea I[20], Goldenberg A[21], Green AK[22], Guerrot AM[21], Heron D[5], Isidor B[23], Kenna BA[24] , Krock BL[8], Kuentz P[16], Lapi E[25], Le Meur N[21], Lesca G[6], Li D[8], Marey I[5], Mignot C[5], Nava C[5], Nesbitt A[8], Nicolas G[21], Roche-Lestienne C[1], Roscioli T[14], Satre V[12], Santani A[8], Stefanova M[22], Steinwall Larsen S[22], Saugier-Veber P[21], Picker-Minh S[26], Thuillier C[1], Verloes A[27], Vieville G[12], Wenzel M[24], Willems M[11], Whalen S[5], Zarate YA[29], Ziegler A[9], Manouvrier-Hanu S[2,3], Kalscheuer VM[30], Gerard B[4] and Ghoumid J[2,3]

[1] Institut de Génétique Médicale, Hôpital Jeanne de Flandre, CHU Lille, Lille, France ; [2] University of Lille, EA 7364-RADEME, Lille, France ; [3] Service de Génétique Clinique, Hôpital Jeanne de Flandre, CHU Lille, Lille, France ; [4] Laboratoire de diagnostic génétique, Institut de Génétique Médicale d'Alsace, Hôpitaux Universitaires de Strasbourg, Strasbourg, France ; [5] AP-HP, Groupe Hospitalier Pitié-Salpêtrière, Département de Génétique, Paris, France ; [6] Service de Génétique, Hospices Civils de Lyon, Lyon, France ; [7] AP-HP, Service de Génétique, Hôpital d'Enfants Armand-Trousseau, Paris, France ; [8] Department of Pathology Laboratory Medicine, Children's Hospital of Philadelphia, USA ; [9] Service de Génétique, CHU d'Angers, Angers, France ; [10] Roberts Individualized Medical Genetics Center, Children's Hospital of Philadelphia, USA ;[11] Département de Génétique Médicale, CHU Montpellier, Montpellier, France ;[12] Laboratoire de Génétique Chromosomique, CHU Grenoble Alpes, Grenoble, France ;[13] Service de Génétique, EA3801, SFR-CAP Santé, CHU de Reims, Reims, France ; [14] St Vincent's Clinical School, University of New South Wales, Darlinghurst, New South Wales, Australia ;[15] Centre de Génétique et Centre de Référence Maladies Rares 'Anomalies du Développement, CHU Dijon, Dijon, France ;[16] Equipe GAD, UMR INSERM 1231, Université de Bourgogne, Dijon, France ; [17] Division of Genetics and Genomic Medicine, Department of Pediatrics, Washington University School of Medicine, St. Louis, Missouri, USA ;[18] The Genetics of Learning Disability Service, Waratah, New South Wales, Australia ;[19] Service de Génétique Médicale, CHU de Clermont-Ferrand, Clermont-Ferrand, France ;[20] AP-HP, Service de Génétique, Hôpital Trousseau, Paris, France. ;[21] Service de Génétique et Inserm U1079, Centre Normand de Génomique Médicale et Médecine Personnalisée, CHU de Rouen, Inserm et Université de Rouen, Rouen, France ;[22] Department of Clinical Genetics, University Hospital Linköping, Linköping, Sweden ;[23] Service de Génétique Médicale, Unité de Génétique Clinique, CHU de Nantes, Nantes, France ;[24] Clinical Genetics, Children's Hospital of Philadelphia, USA ; [25] Medical Genetics Unit, Anna Meyer Children's University Hospital, Florence, Italy ;[26] Department of Pediatric Neurology, Charité-Universitätsmedizin Berlin, Berlin, Germany ; [27] AP-HP, Unité Fonctionnelle de Génétique Clinique, Hôpital Robert Debré, Paris, France ;[28] Genetikum, Neu-Ulm, Germany ;[29] Section of Genetics and Metabolism, Department of Pediatrics, University of Arkansas for Medical Sciences, Little Rock, Arkansas, USA ;[30] Research Group Development and Disease, Max Planck Institute for Molecular Genetics, Berlin, Germany

Abstract

Molecular anomalies in *MED13L*, leading to haploinsufficiency, have been reported in patients with moderate to severe intellectual disability (ID) and distinct facial features, with or without congenital heart defects. Phenotype of the patients was referred to "*MED13L* haploinsufficiency syndrome." Missense variants in *MED13L* were already previously described to cause the *MED13L*-related syndrome, but only in a limited number of patients. Here we report 36 patients with *MED13L* molecular anomaly, recruited through an international collaboration between centers of expertise for developmental anomalies. All patients presented with intellectual disability and severe language impairment. Hypotonia, ataxia, and recognizable facial gestalt were frequent findings, but not congenital heart defects. We identified seven de novo missense variations, in addition to protein-truncating variants and intragenic deletions. Missense variants clustered in two mutation hot-spots, i.e., exons 15–17 and 25–31. We found that patients carrying missense mutations had more frequently epilepsy and showed a more severe phenotype. This study ascertains missense variations in *MED13L* as a cause for *MED13L*-related intellectual disability and improves the clinical delineation of the condition.

Keywords: *MED13L*; intellectual disability; Mediator complex; cardiopathy

Introduction

Mediator is a large coregulator complex conserved from yeast to humans. The complex has emerged as a master coordinator of cell lineage determination, integrating signaling from various transcription factors, epigenetic regulators and non-coding RNAs [1]. In response to various stimuli, mediator undergoes conformational changes and creates a DNA loop between activated enhancer elements and promoter, notably through interactions with cohesins [1]. Mediator physically bridges transcription factors bound at enhancer elements with the RNA polymerase II transcription machinery at core promoter regions [1]. Mediator is organized into four modules, i.e., the tail-, the middle-, the head-, and the CDK8-kinase module [2]. In vertebrates, the latter module is composed of CCNT1 and three additional proteins: CDK8, MED12, and MED13; or their respective paralogs: CDK19, MED12L, and MED13L [3]. Disease-causing variations have been identified in genes encoding the CDK8-module proteins. *MED12* variants cause syndromic intellectual disabilities (ID), namely Opitz-Kaveggia syndrome (MIM #305450), Lujan-Fryns syndrome (MIM #309520), and Ohdo syndrome (MIM #300895) [4]. *CDK19* interruption by a translocation breakpoint has been found to cause moderate ID with microcephaly and retinal folds in a patient [5]. Recently, *MED13L* haploinsufficiency have been identified in patients with moderate to severe ID, hypotonia, and distinctive facial gestalt (OMIM #616789) [6-10]. The recognizable syndrome was delineated by Asadollahi et al. [7] and broadened by further reports [6, 8, 10-15]. The gene is located on chromosome 12q24.21 and encodes MED13L (*alias* TRAP240L), expressed in heart and brain tissues [9]. Originally, the interruption of *MED13L* by a translocation breakpoint was identified in a patient with dextro-loop transposition of the great arteries (dTGA- MIM #608808) and intellectual disability (ID). Given that association, a cohort of 97 individuals with isolated dTGA was screened for *MED13L* sequence variations. Rare heterozygous missense variants were identified in four patients [9]. Familial segregation was not available for three variants and showed that the remaining variant was inherited from a healthy parent [9]. Updated annotations of the four variations showed that the variant c.2056A>C was reported 472 times in GnomAD database (http://gnomad.broadinstitute.org/). Variants c.752A>G and c.6068A>G were reported in GnomAD database once and the variant c.5615G>A was reported once in 1000G database (http://www.internationalgenome.org/). Therefore, clinical relevance of these variants remains unclear. It was hypothesized that missense variants were associated with congenital heart defects (CHDs), particularly dTGA, without intellectual disability [10].

To date, 33 additional patients with a *MED13L* variants or intragenic deletion were reported [6-8, 10, 12, 13, 15-22]. The DDD studies identified at least 19 patients with a *MED13L* variant, highlighting *MED13L* as one of the most common ID-causing gene [14, 23]. Variants were either identified by targeted sequencing, indicating that the condition could be suspected prior to the molecular analysis, or by exome sequencing [12]. Strikingly, no further dTGA was found and all patients presented with ID, characteristic facial gestalt, and less commonly aspecific CHD in 6/25 cases (patent foramen ovale, Fallot tetralogy, pulmonary atresia). To our knowledge, seven missense variants were identified in 11 patients, but the lack of precise clinical data in most of them precluded clarification of their clinical relevance or possible genotype-phenotype correlation [8, 13, 14, 17]. Here, we report on 36 patients with *MED13L* variations affecting its function, including seven missense variants in nine patients. We aim to better delineate the phenotype and discuss possible genotype-phenotype correlation.

Subjects and methods

Patients

Thirty-six patients from 35 families were recruited through an international collaboration between centers of expertise for developmental anomalies. All patients were clinically examined by a clinical geneticist. Two patients have been published previously, but more detailed information was reported here: P7 (ref. [19]) and P8 (ref. [18]). Informed consents were obtained for genetic tests, data sharing, and publication of patients' photographs.

Genetic analyses

Molecular investigations were performed in different diagnostic laboratories according to their routine procedure regarding testing in patients affected with ID. *MED13L* intragenic deletions were identified by array-CGH. *MED13L* sequence variants were identified by either next generation sequencing of custom gene panels designed for ID (P5, P6, P7, P8, P9, P10, P11, P23, P25, P26, P27, and P31) [19, 24] or by whole exome sequencing for the other patients. Missense variations were evaluated using the Alamut interface (Interactive Biosoftware, Rouen, France). Pathogenicity scores were predicted in silico with SIFT (http://sift.jcvi.org), PolyPhen2 (http://genetics.bwh.harvard.edu/pph2/), and MutationTaster (http://www.mutationtaster.org/) softwares. All coordinates are provided for NM_015335.4

transcript in hg19 (genome build: GRCh37) and NP_056150.1 protein. Variant data have been submitted to ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/).

Results

MED13L molecular anomalies

MED13L intragenic deletions, ranging in size from 47 to 200 kb, were identified in five patients (P1, P2, P3, P4, and P24) (Fig. 1 and Table 1). When available, parental segregation showed that these occurred de novo. We identified a *MED13L* sequence variants in 31 patients from 30 families. In one family, recurrence in two sibs (P12 and P13) was observed, presumptively due to parental germinal mosaicism. Protein-truncating variants were identified in 27 patients and were distributed all over the gene (Fig. 1a). One protein-truncating variant, c.1708_1709del, was identified in two unrelated patients (P7-P15). Four variants were predicted to affect splicing (c.5588 + 1G>A - c.1009 + 1G>C and c.2345-3C>G and c.6225 + 1G>A). Variant c.1009 + 1G>C was identified in two unrelated patients (P25-P30). Seven likely pathogenic heterozygous missense variants were identified in nine patients (P14, P20, P21, P22, P23 P28, P32, P33, P35): c.2597C>T p.(Pro866Leu); c.2605C>T p.(Pro869Ser); c.2930C>T p.(Ala977Val); c.6005C>T p.(Ser2002Leu) c.6485C>T p.(Thr2162Met); c.6488C>T p.(Ser2163Leu) c.6530C>A p.(Ser2177Tyr) and were absent from GnomAD and ExAC database in well-covered regions. Both missense variants c.2605C>T p.(Pro869Ser) and c.6488C>T p.(Ser2163Leu) were identified in two unrelated patients, respectively in P28-P35 and P21-P23. Patient P20-P14 and P33 carried respectively the previously reported variants c.2597C>T p.(Pro866Leu), c.6005C>T p.(Ser2002Leu), and c.6485C>T p.(Tr2162Met)[14]. All missense variations were predicted to be “probably damaging” for PolyPhen2 (score > 0.98) and deleterious” for SIFT (score < 0.03). The seven missense variants were predicted to induce substitutions involving Pro866, Pro869, Ala977, Ser2002, Thr2162, Ser2163, and Ser2177 residues, which are highly conserved across vertebrates (PhyloP score 5.69 to 6.18—Fig. 1a). Except for Ala977, which presented a PhyloP score of 6.18, residues involved in these substitutions are also conserved in MED13, the MED13L paralog (Fig. 1).

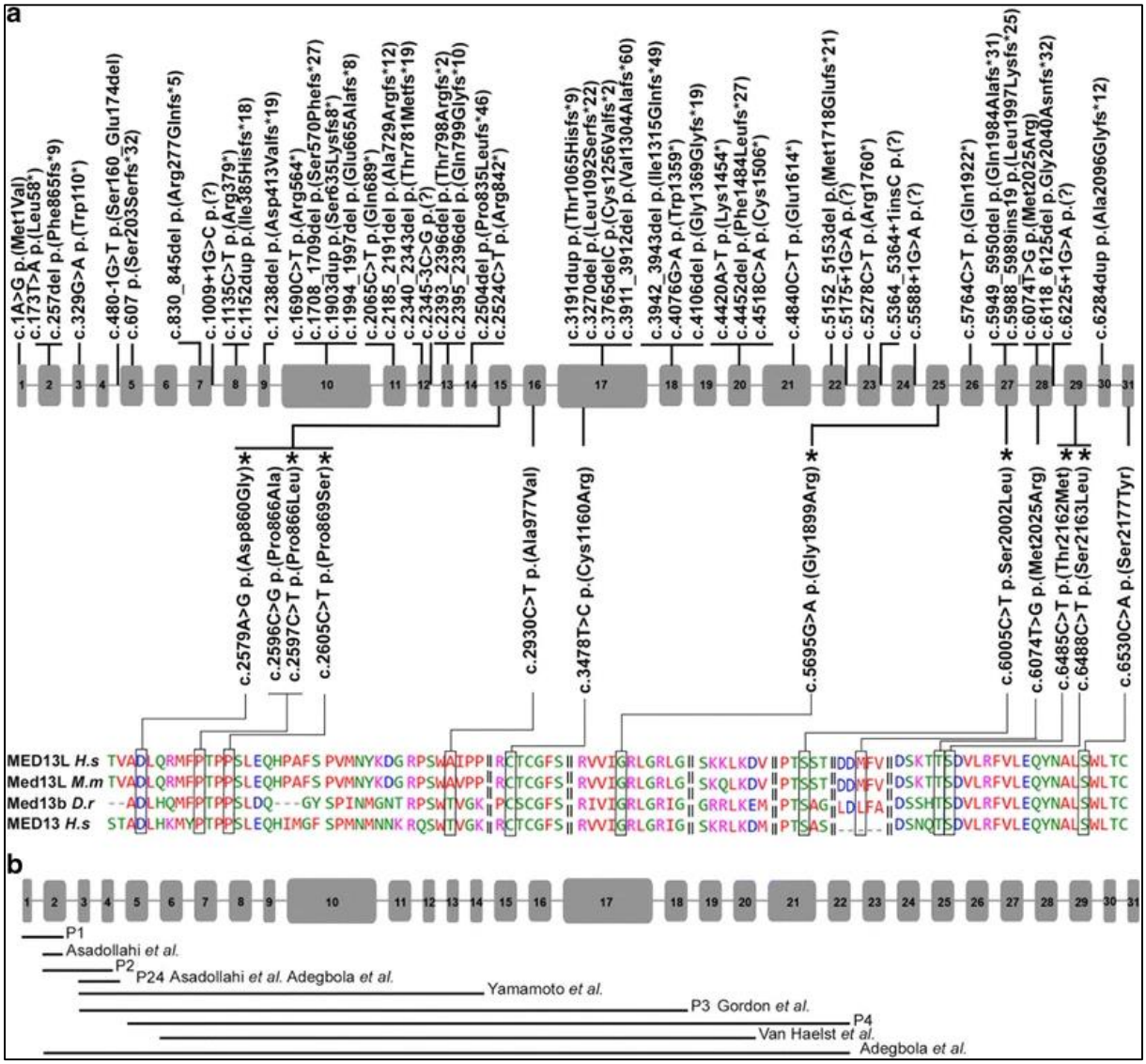


***Figure 1**: A. Summary of MED13L variants reported in the literature and in our cohort. Truncating variants are represented above the gene and missense variants under the gene. Asterisks indicate recurrent missense variants. Inter-species conservation of MED13L missense variations identified in the literature and in our series is shown. H.s., Homo sapiens; M.m, Mus musculus; D.r., Danio rerio. B. Schematic representation of MED13L and location of intragenic deletions reported in our cohort (P1-P2-P3-P4 and P24) and in the literature, indicated by horizontal bars.*

Phenotypic findings

Patient with protein-truncating mutation

Protein-truncating variants were identified in 27 patients. No remarkable prenatal history was reported and birth parameters were normal for all individuals. Motor skills were delayed,

median age for independent walking being 25 months (range from 17 to 41 months). One patient did not achieve walking, but he was only 32-month-old. Speech was also severely impaired in most individuals, composed of few words (16/21–52%) or even absent (5/21–24%). All patients showed moderate to severe ID. Global hypotonia was observed in 20/25 (80%) patients. Ataxia was noticed in 9/25 patients (36%), consisting mainly in dynamic ataxia and dysarthria in four patients. We did not retrieve age at onset of the cerebellar signs, but mean age of the patients with ataxia was 12 years (ranging from 2 to 39 years). One patient presented with seizures (1/26–4%). Autistic features were noticed in 5/21 (24%) cases and behavioral troubles in 10/26 (39%), consisting in aggressive behavior when specified. Brain magnetic resonance (MRI) imaging showed various non-specific anomalies comprising ventriculomegaly, myelination defect, corpus callosum anomaly, or focal cortical dysplasia (Table 1). Majority of the patients shared common facial features with wide open mouth, protruding tongue (without macroglossia), full cheeks, bulbous nasal tip, and horizontal eyebrows. Some patients showed thin vermilion, deep philtrum, and cupid-bow upper lip (Fig. 2 and Table 1). Echocardiography revealed patent foramen ovale in two patients (P1) and pulmonary valvular stenosis in one patient (P25) (Supplemental Table 1).

Patient with missense variation

Missense variants were identified in nine patients. Intra-Uterine Growth Retardation (IUGR) was observed only in P33. Median age for independent walking was also 25 months (range from 18 to 30 months), but 4/9 patients (44%) were not able to walk at the age of examination (P20-P28-P33-P35). The latter patients either did not achieve independent walking (P28-P33) or achieved independent walking and then lost ability to walk because of worsening of epilepsy (P20-P35). Speech was absent in 5/9 patients (56%) and composed of few words in 4/9 patients (44%). Global hypotonia was observed in 6/9 patients (67%). Ataxia was noticed in 3/7 patients (43%). Six patients presented with seizures 6/9 patients (67%), consisting in febrile seizure (P32), late onset infantile spasms (P20) and Lennox-Gastaut syndrome (P35). Abnormal brain magnetic resonance imaging (MRI) showed various non-specific anomalies in 4/7 patients (57%), comparable with these found in patients with protein-truncating variations (supplemental Table 1). We observed atypical facial features in three patients (P20-P28-P35). They showed long down-slanting palpebral fissures, with everted lower eyelids (Fig. 2d). Echocardiography revealed a patent foramen ovale in one patient (P33) (Table 1).

***Figure 2**: Morphological features of a selection of patients. A- Patients with protein-truncating variants. Core facial features comprises depressed nasal bridge, horizontal eyebrows, full cheeks and large open mouth. Majority of patients show also cupid-bow upper lip, thin vermilion border and deep philtrum. B- Patients P20-P28 and P35 (at different age) show atypical facial gestalt with long down-slanting palpebral fissures and everted lower eyelids. C-Some patients, notably in infancy have broad, stubby and tapering fingers. Feet showed long halluces and sandal-gap deformity in some patients. D-Photo enlargement of the palpebral features of patients P20-P28 and P35.*

Discussion

Here we report on a cohort of 36 patients carrying *MED13L* anomalies, including two previously published cases, allowing a better delineation of the associated phenotype. All individuals had motor delay, speech delay, and moderate to severe ID. In most patients, language was limited to few words or was even absent. Patients showed various degrees of

cerebellar dysfunction. Ataxia was observed in 11/32 patients, and was reported especially in the older cases (P9 –P10-P13-P15-P17-P20-P23-P32–P32-P36). Since this feature was under-reported in younger patients, we hypothesize that cerebellar involvement probably worsened with age. No patient showed cerebellar anomaly on brain MRI, but repeated imaging may be needed to explore possible progressive atrophy. MRI identified aspecific features comprising myelination defects, corpus callosum abnormalities, white matter anomalies, ventriculomegaly and focal cortical dysplasia. This study confirms that most patients show a recognizable facial gestalt, which could phenotypically overlap with deletion 1p36 microdeletion syndrome (OMIM# 607872) in some patients [12]. Core facial features comprise depressed nasal bridge, horizontal eyebrows, full cheeks, and large open mouth [7, 12]. More subtle features like cupid-bow upper lip, thin vermilion border, and deep philtrum can be observed. However, in a few patients, these core facial features were absent (P22-P27). We also noticed that non-recurrent facial features can be associated, especially in patients with a missense variant located in the exon 15 (P20-P28-P35) (Fig. 2b). In our cohort, patients were not suspected with the condition prior to genetic testing, but secondarily facial comparison allowed the description of common features. These data highlight the valuable role of clinical geneticists in the precision of the phenotype, a critical step to determine the pathogenicity of *MED13L*-variants.

We found only three patients with CHD, consisting in patent foramen ovale or pulmonary valvular stenosis but no complex CHD (Table 1). Patients carried respectively *MED13L* intragenic deletion, splice-site variant, and missense variant (P1-P25-P33), confirming that frequency of complex CHD is less than initially expected and is not correlated with *MED13L*-missense variants. Concerning patients reported by Muncke et al., showing dTGA and probably no developmental delay [9], it is unlikely that their variants affect the protein function. *MED13L* is located next to *TBX3* and *TBX5* genes. These genes respectively cause ulnar-mammary syndrome (OMIM#181450) and Holt-Oram syndrome (OMIM#142900), both conditions comprising CHD. Heart-specific enhancers have been identified within regulatory domains of both genes; however, they do not overlap with regions contacting MED13L [25]. Therefore, it is unlikely that *MED13L* variants could affect cis-regulatory elements controlling *TBX3* or *TBX5* expression during heart development. Involvement of *MED13L* in the dTGA of these patients remains to be explained.

We observed clinical variability, notably in patients who carried recurrent variations (P12-P13; P21-P23; P7-P15; P25-P30; P28-P35) (Table 1). As suggested by Asadollahi et al., we found that patients with missense variants were more prompt to develop epilepsy, compared to patients with protein-truncating variants (4/9 versus 1/26) [8]. Severe neurodevelopmental phenotype (absent speech in 5/9 versus 5/21—non ambulatory 4/9 versus 1/21) and malformations are also more frequent (Supplemental Table 1). More precisely, patients P20 and P35 lost the ability to walk consecutively to worsening of epilepsy. They needed gastrostomy tube feeding and P35 had hearing impairment and severe myopia. P33 showed a severe phenotype associating IUGR, absent speech, microcephaly, colobomatous microphthalmia, and never achieved sitting. Moreover, P20-P28 and P35 showed atypical facial features with long palpebral fissures and even everted lower eyelid in P20 and P35 (Fig. 2b). Based on these particular palpebral features, P20 was initially diagnosed with Kabuki syndrome (OMIM #147920). All these data are supported by clinical features from the 11 patients reported in the literature, carrying missense variants [8, 13, 14, 17]. Clinical details of the patients reported by the DDD study were retrieved from Decipher database (https://decipher.sanger.ac.uk/) and corresponded to patients' ID: #272205, #260542, #262717, #258131, #268019, #262545, #265953, and #323183. Among them, four patients had epilepsy, and two patients had IUGR and atypical features were also noted (craniosynostosis, microcephaly, major feeding difficulties, limb malformations). Since patients with missense mutation seem to have a more severe phenotype, we could hypothesize that they induce a dominant negative effect, contrarily to protein-truncating variants. We did not consider the patient reported by Mullegama et al., who suffered from speech delay, ASD and Mediterranean fever. He carried *MED13L*, *DEAF1*, and *MEFV* variants [26]. Inheritance of the *MED13L* missense variant c.5282C>T p.(Prol1761Leu) could not be determined, since the patient was adopted. Polyphen2 and SIFT software predicted in silico that the variant was respectively benign and tolerated. There is no experimental evidence of the deleterious effect on the function of the protein. Thus, there was not enough evidence to consider the variant as the cause of the neurodevelopmental disorder of the patient.

The seven missense variants identified in this study, as well as the *MED13L*-missense variants previously reported in the literature, cluster in exons 15–17 and 25–31 (Fig. 1) [8, 13, 14, 17]. Both localizations constitute hot-spots of mutations. As expected, majority of missense variations are recurrent [8]. Previously reported variants c.2597C>T p.(Pro866Leu), c.6005C>T p.(Ser2002Leu), c.6485C>T p.(Thr2162Met) were identified in three patients [8,

13, 14, 17]. We identified four novel missense variants in six patients: variant c.2605C>T p.(Pro869Ser) was identified in two patients and as well as variant c.6488C>T p.(Ser2163Leu). Variants c.2930C>T p.(Ala977Val) and c.6530C>A p.(Ser2177Tyr) were not recurrent. One of the MED13/MED13L functions is to physically link the CDK8-module to the core Mediator complex, mainly by interacting with MED19 and CDK8 [27]. Dissociation of the CDK8-module components from the core Mediator is mediated by Mediator-bound MED13/MED13L ubiquitylation and degradation [28]. Both subunits can also relay information from temporal/spatial signals or transcription factors to the RNA polymerase II machinery, thus controlling the expression of specific genes, notably genes involved in Wnt, FGF, and Rb/E2F pathways [8]. Since all the residues involved in the substitutions are located in highly conserved across vertebrate regions and even conserved in MED13, we can assume that these residues are probably implicated in such mechanisms. Further studies are needed to unravel the deleterious mechanism induced by these molecular changes.

Conclusion

In this cohort of 36 patients with *MED13L*-related intellectual disability, we confirmed recognizable facial gestalt and intellectual involvement. We highlighted possible arising of progressive cerebellar signs. We did not confirm congenital heart defects as a major feature of the condition. Patients with missense variant are significantly more at risk to develop epilepsy and seemed to have a more severe phenotype, suggesting possible dominant negative effect of the missense variants. We observed clustering of missense variants in specific domains of the protein. Substitution involving the highly conserved across species residues Asp860, Pro866, Pro869, Gly1899, Ser2002, Thr2162, and Ser2163 were identified in at least two patients. Precise roles of these domains and specific residues remain to be determined to better understand molecular mechanisms underlying *MED13L*-related intellectual disability.

Table 1: Clinical and molecular data of the 36 patients harbouring likely pathogenic *MED13L* variants and comparison with the literature. (**patients 1 to 7**)

Patient	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Gender	F	M	M	M	M	F	M
Age at first examination (years)	2	3.8	9	9	11	4	4
DD/ID	+	+	+	+	+ (moderate)	+ (severe)	+ (moderate)
Speech delay	+	+	+	+	+	+	+
Speech features	abs.	few words	NA	few words	few words	abs.	abs.
Motor delay	+	+	+	+	+	+	+
Walk (months)	23	29	25	36	23	17	24
Hypotonia	+	–	+	+	+	+	–
Ataxia	–	–	–	+ (dysarthria)	–	–	–
Seizures	–	–	–	–	–	–	–
Autistic features	–	+	NA	+	–	+	–
Behavioral troubles	–	–	–	+	–	+	–
MRI	NA	NA	NA	Moderate VM	NA	WM hypersignal	Normal
Upslanting palpebral fissures	+	–	–	–	NA	–	–
Bulbous nasal tip	+	+	+	+	NA	+	+
Cupid-bow upper lip	–	+	+	+	NA	–	+
Hypotonic open-mouth	+	+	+	+	NA	+	+
Thin vermillon	–	+	+	+	NA	–	–

border							
Deep philtrum	+	NA	+	+	NA	–	+
CHD	+ (patent foramen ovale)	NA	–	NA	NA	–	–
Miscellaneous	Duplicated thumb	II-III toes syndactyly, cryptorchidism		Micropenis		Hirsutism	Adducted thumb
MED13L	exon 1–2 deletion	exon 2 to 4 deletion	exon 3 to 18 deletion	exon 5 to 22 deletion	c.173 T > A	c.5278C > T	c.1708_1709del
Parental segregation	NA	de novo	de novo	de novo	de novo	de novo	de novo

Table 1: patients 8 to 14

Patient	8	9	10	11	12	13	14
Gender	M	M	M	F	M	F	F
Age at first examination (years)	10	13	4.75	2.75	9	19	4.25
DD/ID	+ (moderate)	+ (moderate)	+ (severe)	+ (severe)	+	+	+
Speech delay	+	+	+	+	+	+	+
Speech features	few words	NA	few words	abs. (disyllabic)	few words	NA	few words
Motor delay	+	+	+	+	+	+	+
Walk (months)	22	30	25	30	30	30	26
Hypotonia	+	+	+	+	+	–	–
Ataxia	+ (dysarthria)	+	+	–	+ (dysarthria)	+	–
Seizures	–	–	–	–	–	–	–
Autistic features	+	–	–	NA	NA	NA	–
Behavioral troubles	+	+	+	–	NA	+	–
MRI	Normal	Normal	Normal	WM hypersignal	Labyrinthic malformations	Pituitaty adenoma	Normal
Upslanting palpebral fissures	–	+	–	+	–	–	–
Bulbous nasal tip	+	+	+	+	+	+	–
Cupid-bow upper lip	+	+	+	+	–	–	–
Hypotonic open-mouth	+	–	+	+	+	+	+
Thin vermillon border	–	–	+	–	–	–	+
Deep philtrum	–	+	+	+	–	–	–

CHD	–	–	–	–	–	–	–
Miscellaneous	Syndactyly, coloboma	Large incisors	Club feet	CBH syndrome	Obsesity, cryptorchidism, pyramidal syndrome	Obsesity, pyramidal syndrome	Kidney cysts
MED13L	c.6118_6125del	c.1152dup p.(Ile385fs*18)	c.4452del	c.5152_5153del	c.5588 + 1G > A	c.5588 + 1G > A	c.6485C > T
Parental segregation	de novo	NA	de novo	de novo	de novo	de novo	de novo

Table 1: patients 15 to 21

Patient	15	16	17	18	19	20	21
Gender	M	M	F	M	M	M	F
Age at first examination (years)	7	8.5	34	31	10	15	9
DD/ID	+ (moderate)	+ (severe)	+ (moderate)	+ (moderate)	+ (moderate)	+ (severe)	+ (severe)
Speech delay	+	+	+	+	+	+	+
Speech features	NA	abs.	few words	few words	few sentences	abs.	abs.
Motor delay	+	+	+	+	+	+	+
Walk (months)	NA	24	20	NA	41	abs.	30
Hypotonia	+	+	+	NA	+	+	+
Ataxia	+	–	+	–	–	+	–
Seizures	NR	+	–	–	–	+	+
Autistic features	NA	+	–	–	–	+	+
Behavioral troubles	–	+	–	–	–	–	–
MRI	Subarachnoid spaces enlargement	Normal	NA	CC agenesis	Focal polymicrogyria	Normal	Normal
Upslanting palpebral fissures	–	–	–	+	–	NA	–
Bulbous nasal tip	+	–	+	+	+	+	+
Cupid-bow upper lip	NA	–	–	NA	+	–	–
Hypotonic open-mouth	NA	–	+	+	–	+	+
Thin vermillon	+	+	–	–	–	–	–

border							
Deep philtrum	NA	+	–	NA	+	–	–
CHD	NA	–	–	–	–	–	–
Miscellaneous	Posterior cleft palate	Strabismus, cryptorchidism	Hypopallesthesia	Hirsutism, strabismus	Robin sequence, 3–4-5 Toe clinodactyly, severe scoliosis	Nystagmus, craniosynostosis	
MED13L	c.1708_1709del	c.6284dup	c.830_845del	c.2065C > T	c.3942_3943del	c.2597C > T	c.6488C > T
Parental segregation	de novo	de novo	de novo	NA	de novo	de novo	NA

Table 1: patients 22 to 28

Patient	22	23	24	25	26	27	28
Gender	F	F	F	F	M	M	F
Age at first examination (years)	8	12	1.6	0.7	2.5	2.7	12
DD/ID	+	+ (severe)	+	+	+	+	+
Speech delay	+	+	+	NA	+	+	+
Speech features	few words	few words	few words	NA	few words	NA	abs.
Motor delay	+	+	+	+	+	+	+
Walk (months)	25	26	Not acquired yet	NA	23	abs.	abs.
Hypotonia	–	+	+	+	+	+	–
Ataxia	–	+	–	NA	+ (dysarthria)	NA	–
Seizures	–	–	–	–	–	–	+
Autistic features	+	–	–	NA	–	–	NA
Behavioral troubles	+	–	–	–	+	+	NA
MRI	Focal cortical dysplasia	NA	NA	NA	Normal	NA	CCH-Hypomyelination
Upslanting palpebral fissures	–	+	–	–	+	–	+
Bulbous nasal tip	–	+	+	–	–	–	+
Cupid-bow upper lip	–	+	+	+	+	–	+
Hypotonic open-mouth	–	+	+	+	+	–	+
Thin vermillon border	–	+	–	+	+	+	+
Deep philtrum	–	+	–	+	+	–	+

CHD	–	–	–	Pulmonary vavular stenosis	–	–	–
Miscellaneous		Vertebral artery occlusion	Bilateral talipes, umbilical hernia			Bilateral talipes	IUGR clynodactyly-double ureter-CoAo
MED13L	c.2930C > T	c.6488C > T	exon3to4deletion	c.1009 + 1G > C	c.2340_2343del	c.3911_3912delp.(Val1304Alafs*60)	c.2605C > T
Parental segregation	de novo	de novo	de novo	de novo	de novo	NA	de novo

Table 1: patients 29 to 36

Patient	29	30	31	32	33	34	35	36
Gender	M	F	F	M	F	M	M	F
Age at first examination (years)	2	39	4	5	9	17	24	3
DD/ID	+	+	+	+	+	+	+	+
Speech delay	+	+	+	+	+	+	+	+
Speech features	few words	few sentences	few words	few words	abs.	few words	abs.	few words
Motor delay	+	+	+	+	+	+	+	+
Walk (months)	NA	20	NA	18	abs.	30	abs.	19
Hypotonia	+	–	NA	+	+	+	+	–
Ataxia	–	+	–	+	NA	–	NA	+
Seizures	–	–	–	+	–	–	+	–
Autistic features	–	–	–	+		–	–	–
Behavioral troubles	–	+	–	+	–	–	–	–
MRI	NA	Normal	Normal	NA	CCH-VM	Fine splenium	Diffuse cortical atrophy	
Upslanting palpebral fissures	+	+	NA	NA	+	+	–	–
Bulbous nasal tip	+	+	NA	–	+	+	+	+
Cupid-bow upper lip	+	+	NA	NA	–	+	+	–
Hypotonic open-mouth	+	+	NA	NA	–	+	+	–

Thin vermillon border	+	+	NA	NA	+	–	+	+
Deep philtrum	+	+	NA	NA	–	–	+	+
CHD	–	–	NA	–	+ (patent foramen ovale)	–	–	NA
Miscellaneous	5 Toe clinodactyly		NA	Bilateral club foot	IUGR,colobomatous microphtalmia, bilateral talipes	Assymetric face - assimetric microphtalmia	Feeding difficulties - hearing impairment	Feeding difficulties in infancy
MED13L	c.1903dupA	c.1009 + 1G > C	c.2345-3C > G	c.6530C > A	c.6005C > T	c.6225 + 1G > A	c.2605C > T	c.5764C > T
Parental segregation	de novo	de novo	de novo	de novo	de novo	de novo	de novo	de novo

Table 1:

Patient	This study (*n* = 36)	Previous studies (*n* = 30)	Total (*n* = 66)
DD/ID	100% (36/36)	100% (30/30)	100% (66/66)
Speech delay	100% (36/36)	100% (30/30)	100% (66/66)
Speech features			
Motor delay	100% (36/36)	100% (19/19)	100% (55/55)
Walk (months)			
Hypotonia	77% (26/34)	62% (18/29)	68% (44/63)
Ataxia	34% (11/32)	32% (9/28)	33% (20/60)
Seizures	17% (6/35)	14% (4/28)	16% (10/63)
Autistic features	34% (10/29)	21% (6/29)	28% (16/58)
Behavioral troubles	31% (10/32)	32% (6/19)	31% (16/51)
MRI			
Upslanting palpebral fissures	34% (11/32)	62% (15/23)	47% (26/55)

Bulbous nasal tip	82% (27/33)	74% (17/23)	79% (44/56)
Cupid-bow upper lip	58% (18/31)		
Hypotonic open-mouth	78% (25/32)	78% (14/18)	78% (39/50)
Thin vermillon border	52% (17/33)		
Deep philtrum	58% (18/31)		
CHD	11% (3/27)	24% (6/25)	17% (9/52)

References

1.Allen BL, Taatjes DJ (2015) The mediator complex: a central integrator of transcription. Nat Rev Mol Cell Biol 16(3):155–166

2.Yin J-w, Wang G (2014) The mediator complex: a master coordinator of transcription and cell lineage development. Development 141(5):977–987

3.Daniels DL (2013) Mutual exclusivity of MED12/MED12L, MED13/13L, and CDK8/19 paralogs revealed within the CDK-mediator kinase module. J Proteomics Bioinformat 01(S2)

4.Graham JM, Schwartz CE (2013) MED12 related disorders. Am J Med Genet A 161A(11):2734–2740

5.Mukhopadhyay A, Kramer JM, Merkx G, Lugtenberg D, Smeets DF, Oortveld MAW, Blokland EAW, Agrawal J, Schenck A, van Bokhoven H, Huys E, Schoenmakers EF, van Kessel AG, van Nouhuys CE, Cremers FPM (2010) CDK19 is disrupted in a female patient with bilateral congenital retinal folds, microcephaly and mild mental retardation. Hum Genet 128(3):281–291

6.Adegbola A, Musante L, Callewaert B, Maciel P, Hu H, Isidor B, Picker-Minh S, le Caignec C, Delle Chiaie B, Vanakker O, Menten B, Dheedene A, Bockaert N, Roelens F, Decaestecker K, Silva J, Soares G, Lopes F, Najmabadi H, Kahrizi K, Cox GF, Angus SP, Staropoli JF, Fischer U, Suckow V, Bartsch O, Chess A, Ropers HH, Wienker TF, Hübner C, Kaindl AM, Kalscheuer VM (2015) Redefining the MED13L syndrome. Eur J Hum Genet 23(10):1308–1317

7.Asadollahi R, Oneda B, Sheth F, Azzarello-Burri S, Baldinger R, Joset P, Latal B, Knirsch W, Desai S, Baumer A, Houge G, Andrieux J, Rauch A (2013) Dosage changes of MED13L further delineate its role in congenital heart defects and intellectual disability. Eur J Hum Genet 21(10):1100–1104

8.Asadollahi R et al. Genotype-phenotype evaluation of MED13L defects in the light of a novel truncating and a recurrent missense mutation. Eur J Med Genet 60(9):451-464

9.Muncke N, Jung C, Rüdiger H, Ulmer H, Roeth R, Hubert A, Goldmuntz E, Driscoll D, Goodship J, Schön K, Rappold G (2003) Missense mutations and gene interruption in PROSIT240, a novel TRAP240-like gene, in patients with congenital heart defect (transposition of the great arteries). Circulation 108(23):2843–2850

10.van Haelst MM, Monroe GR, Duran K, van Binsbergen E, Breur JM, Giltay JC, van Haaften G (2015) Further confirmation of the MED13L haploinsufficiency syndrome. Eur J Hum Genet 23(1):135–138

12.Cafiero C, Marangi G, Orteschi D, Ali M, Asaro A, Ponzi E, Moncada A, Ricciardi S, Murdolo M, Mancano G, Contaldo I, Leuzzi V, Battaglia D, Mercuri E, Slavotinek AM, Zollino M (2015) Novel de novo heterozygous

loss-of-function variants in MED13L and further delineation of the MED13L haploinsufficiency syndrome. Eur J Hum Genet 23(11):1499–1504

13.Caro-Llopis A, Rosello M, Orellana C, Oltra S, Monfort S, Mayo S, Martinez F (2016) De novo mutations in genes of mediator complex causing syndromic intellectual disability: mediatorpathy or transcriptomopathy? Pediatr Res 80:809–815

14.Deciphering Developmental Disorders, S (2017) Prevalence and architecture of de novo mutations in developmental disorders. Nature 542(7642):433–438

15.Gordon CT, Chopra M, Oufadem M, Alibeu O, Bras M, Boddaert N, Bole-Feysot C, Nitschké P, Abadie V, Lyonnet S, Amiel J (2018) MED13L loss-of-function variants in two patients with syndromic Pierre Robin sequence. Am J Med Genet A 176(1):181–186

16.Iossifov I, Ronemus M, Levy D, Wang Z, Hakker I, Rosenbaum J, Yamrom B, Lee YH, Narzisi G, Leotta A, Kendall J, Grabowska E, Ma B, Marks S, Rodgers L, Stepansky A, Troge J, Andrews P, Bekritsky M, Pradhan K, Ghiban E, Kramer M, Parla J, Demeter R, Fulton LL, Fulton RS, Magrini VJ, Ye K, Darnell JC, Darnell RB, Mardis ER, Wilson RK, Schatz MC, McCombie WR, Wigler M (2012) De novo gene disruptions in children on the autistic spectrum. Neuron 74(2):285–299

17.Gilissen C, Hehir-Kwa JY, Thung DT, van de Vorst M, van Bon BWM, Willemsen MH, Kwint M, Janssen IM, Hoischen A, Schenck A, Leach R, Klein R, Tearle R, Bo T, Pfundt R, Yntema HG, de Vries BBA, Kleefstra T, Brunner HG, Vissers LELM, Veltman JA (2014) Genome sequencing identifies major causes of severe intellectual disability. Nature 511(7509):344–347

18.Hamdan FF, Srour M, Capo-Chichi JM, Daoud H, Nassif C, Patry L, Massicotte C, Ambalavanan A, Spiegelman D, Diallo O, Henrion E, Dionne-Laporte A, Fougerat A, Pshezhetsky AV, Venkateswaran S, Rouleau GA, Michaud JL (2014) De novo mutations in moderate or severe intellectual disability. PLoS Genet 10(10):e1004772

19.Redin C, Gérard B, Lauer J, Herenger Y, Muller J, Quartier A, Masurel-Paulet A, Willems M, Lesca G, el-Chehadeh S, le Gras S, Vicaire S, Philipps M, Dumas M, Geoffroy V, Feger C, Haumesser N, Alembik Y, Barth M, Bonneau D, Colin E, Dollfus H, Doray B, Delrue MA, Drouin-Garraud V, Flori E, Fradin M, Francannet C, Goldenberg A, Lumbroso S, Mathieu-Dramard M, Martin-Coignard D, Lacombe D, Morin G, Polge A, Sukno S, Thauvin-Robinet C, Thevenon J, Doco-Fenzy M, Genevieve D, Sarda P, Edery P, Isidor B, Jost B, Olivier-Faivre L, Mandel JL, Piton A (2014) Efficient strategy for the molecular diagnosis of intellectual disability using targeted high-throughput sequencing. J Med Genet 51(11):724–736

20.Utami KH, Winata CL, Hillmer AM, Aksoy I, Long HT, Liany H, Chew EG, Mathavan S, Tay SK, Korzh V, Sarda P, Davila S, Cacheux V (2014) Impaired development of neural-crest cell-derived organs and intellectual disability caused by MED13L haploinsufficiency. Hum Mutat 35(11):1311–1320

21.Wang T, Guo H, Xiong B, Stessman HAF, Wu H, Coe BP, Turner TN, Liu Y, Zhao W, Hoekzema K, Vives L, Xia L, Tang M, Ou J, Chen B, Shen Y, Xun G, Long M, Lin J, Kronenberg ZN, Peng Y, Bai T, Li H, Ke X, Hu Z, Zhao J, Zou X, Xia K, Eichler EE (2016) De novo genic mutations among a Chinese autism spectrum disorder cohort. Nat Commun 7:13316

22.Popp B, Ekici AB, Thiel CT, Hoyer J, Wiesener A, Kraus C, Reis A, Zweier C (2017) Exome Pool-Seq in neurodevelopmental disorders. Eur J Hum Genet 25(12):1364–1376

23.Deciphering Developmental Disorders, S (2015) Large-scale discovery of novel genetic causes of developmental disorders. Nature 519(7542):223–228

24.Lehalle D, Mosca-Boidron AL, Begtrup A, Boute-Benejean O, Charles P, Cho MT, Clarkson A, Devinsky O, Duffourd Y, Duplomb-Jego L, Gérard B, Jacquette A, Kuentz P, Masurel-Paulet A, McDougall C, Moutton S, Olivié H, Park SM, Rauch A, Revencu N, Rivière JB, Rubin K, Simonic I, Shears DJ, Smol T, Taylor Tavares AL, Terhal P, Thevenon J, van Gassen K, Vincent-Delorme C, Willemsen MH, Wilson GN, Zackai E, Zweier C, Callier P, Thauvin-Robinet C, Faivre L (2017) STAG1 mutations cause a novel cohesinopathy characterised by unspecific syndromic intellectual disability. J Med Genet 54:479–488

25.van Weerd JH, Badi I, van den Boogaard M, Stefanovic S, van de Werken HJG, Gomez-Velazquez M, Badia-Careaga C, Manzanares M, de Laat W, Barnett P, Christoffels VM (2014) A large permissive regulatory domain exclusively controls Tbx3 expression in the cardiac conduction system. Circ Res 115(4):432–441

26.Mullegama SV, Jensik P, Li C, Dorrani N, UCLA Clinical Genomics Center, Kantarci S, Blumberg B, Grody WW, Strom SP (2017) Coupling clinical exome sequencing with functional characterization studies to diagnose a patient with familial Mediterranean fever and MED13L haploinsufficiency syndromes. Clin Case Rep 5(6):833–840

27.Tsai KL, Sato S, Tomomori-Sato C, Conaway RC, Conaway JW, Asturias FJ (2013) A conserved mediator-CDK8 kinase module association regulates mediator-RNA polymerase II interaction. Nat Struct Mol Biol 20(5):611–619

28.Davis MA, Larimore EA, Fissel BM, Swanger J, Taatjes DJ, Clurman BE (2013) The SCF-Fbw7 ubiquitin ligase degrades MED13 and MED13L and regulates CDK8 module association with mediator. Genes Dev 27(2):151–156

Supplemental Table 1 : Comparison of clinical data of patients harbouring *MED13L* protein-truncating variants and missense variants.

Patient	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	This Study (n=31)	Previous Studies (n=24)	Total (n=55)
Gender	F	M	M	M	M	F	M	M	M	M	F	M	F	F	M	M	F	M	M	M	F	F	F	F	F	M	M	F	M	F	F	M			
Age at first examination (y)	2	3,8	9	9	11	4	4	10	13	4,75	2,75	9	19	4,25	7	8,5	34	31	10	15	9	8	12	1,6	0,7	2,5	2,7	12	2	39	NA	5			
DD/ID	+	+	+	+	+ (moderate)	+ (severe)	+ (moderate)	+ (moderate)	+ (moderate)	+ (severe)	+ (severe)	+	+	+	+ (moderate)	+ (severe)	+ (moderate)	+ (moderate)	+ (moderate)	+ (severe)	+ (severe)	+	+ (severe)	+	+	+	+	+	+	+	NA	+	100% (30/30)	100% (23/23)	100% (52/52)
Speech Delay	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	NA	+	+	+	+	+	NA	+	100% (30/30)	100% (17/17)	100% (46/46)
Speech features	abs.	few words	NA	orofacial apraxia	few words	abs.	abs.	dysarthria	NA	few words (<10)	dysyllabic	dysarthria	NA	few words	NA	abs.	few words	few words	few sentences	abs.	abs.	single word	< 10 words	echolalia	NA	dysarthria	NA	abs.	single words	few sentences	NA	few words			
Motor Delay	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	NA	+	100% (30/30)	100% (15/15)	100% (44/44)
Walk (months)	23	29	25	36	23	17	24	22	30	25	30	30	30	26	NA	24	20	NA	41	absent	30	25	26	absent	NA	23	absent	absent	NA	20	NA	18			
Hypotonia	+	-	+	+	+	+	-	+	+	+	+	+	-	-	+	+	+	+	+	+	+	-	+	+	+	+	+	-	+	-	NA	+	77% (23/30)	70% (12/17)	76% (35/46)
Ataxia	-	-	-	-	-	-	-	-	+	+	-	+	+	-	+	-	+	-	-	+	-	-	+	-	NA	-	NA	-	-	+	NA	+	32% (9/28)	29% (4/14)	29% (12/41)
Seizures	-	-	-	-	-	-	-	-	-	-	-	-	-	-	NR	+	-	-	-	+	+	-	-	-	-	-	-	+	-	-	NA	+	14% (4/29)	21% (4/19)	17% (8/47)
Autistic features	-	+	NA	+	-	+	-	+	-	-	NA	NA	NA	-	NA	+	-	-	-	+	+	+	-	-	NA	-	-	NA	-	-	NA	+	36% (8/23)	35% (7/20)	36% (15/42)
Behavioral troubles	-	-	-	+	-	+	-	+	+	+	-	NA	+	-	-	+	-	-	-	-	-	+	-	-	-	+	+	NA	-	+	NA	+	36% (10/28)	40% (6/15)	36% (15/42)
MRI	NA	NA	NA	moderate VM	NA	WM hypersignal	normal	normal	normal	normal	WM hypersignal	labyrinthic malformations	pituitary adenoma	normal	subarachnoid spaces enlargement	normal	NA	CC agenesis	focal polymicrogyria	normal	normal	focal cortical dysplasia	NA	NA	NA	normal	NA	CCH- Hypomyelination	NA	normal	NA	NA			
Upslanting palpebral fissures	+	-	-	-	NA	-	-	-	+	-	+	-	-	-	-	-	-	+	-	NA	-	-	+	-	-	+	-	+	+	+	NA	NA	29% (9/28)	59% (10/17)	41% (18/44)
Bulbous nasal tip	+	+	+	+	NA	+	+	+	+	+	+	+	+	-	+	-	+	+	+	+	+	-	+	+	-	-	-	+	+	+	NA	NA	79% (23/29)	87% (13/15)	81% (35/43)
Cupid-bow upper lip	-	+	+	+	NA	-	+	+	+	+	+	-	-	-	NA	-	-	NA	+	-	-	-	+	+	+	+	-	+	+	+	NA	NA	59% (16/27)		
Hypotonic open-mouth	+	+	+	+	NA	+	+	+	-	+	+	+	+	+	NA	-	+	+	-	+	+	-	+	+	+	+	-	+	+	+	NA	NA	82% (23/28)	71% (10/14)	78% (32/41)
Thin vermillion border	-	+	+	+	NA	-	-	-	-	+	-	-	-	+	+	+	-	-	-	-	-	-	+	-	+	+	+	+	+	+	NA	NA	48% (14/29)		
Deep philtrum	+	NA	+	+	NA	-	+	-	+	+	+	-	-	-	NA	+	-	NA	+	-	-	-	+	-	+	+	-	+	+	+	NA	NA	59% (16/27)		
CHD	+ (patent foramen ovale)	NA	-	NA	NA	-	-	-	-	-	-	-	-	-	NA	-	-	-	-	-	-	-	-	-	Pulmonary vavular stenosis	-	-	-	-	-	NA	-	8% (2/24)	29% (6/21)	18% (8/44)
Miscellaneous	duplicated thumb	II-III toes syndactyly, cryptorchidism		micropenis		hirsutism	adducted thumb	syndactyly, coloboma	large incisors	club feet	CBH syndrome	obsesity, cryptorchidism, pyramidal syndrome	obsesity, pyramidal syndrome	kidney cysts	posterior cleft palate	strabismus, cryptorchidism	hypopallesthesia	hirsutism, strabismus	Robin sequence 3-4-5 Toe clinodactyly severe scoliosis	nystagmus, craniosynostosis			vertebral artery occlusion	bilateral talipes, umbilical hernia			bilateral talipes	clynodactyly- double ureter-CoAo	5 Toe clinodactyly		NA	bilateral club foot			
***MED13L* molecular anomalies**	exon 1-2 deletion	exon 2 to 4 deletion	exon 3 to 18 deletion	exon 5 to 22 deletion	c.173T>A p.(Leu58*)	c.5278C>T p.(Arg1760*)	c.1708_1709del p.(Ser570Profs*27)	c.6118_6125del p.(Glu2040Asnfs*32)	c.1152dup p.(Ile385fs*18)	c.4452del p.(Pro1484Leufs*27)	c.5152_5153del p.(Met1718Ilefs*21)	c.5589+1G>A p.(?)	c.5589+1G>A p.(?)	c.6485C>T p.(Thr2162Met)	c.1708_1709del p.(Ser570Profs*27)	c.6284dup p.(Ala2096Glyfs*12)	c.838_845del p.(Arg2772ar*5)	c.2065C>T p.(Gln689*)	c.3942_3943del p.(Ile1315Glnfs*49)	c.2597C>T p.(Pro866Leu)	c.6488C>T p.(Ser2163Leu)	c.2930C>T p.(Ala977Val)	c.6488C>T p.(Ser2163Leu)	exon 3 to 4 deletion	c.1009+1G>C p.(?)	c.2348_2343del p.(Thr781Metfs*19)	c.3911_3912del p.(Val1304Ala*40)	c.2805C>T p.(Pro935ser)	c.1903dupA (p.Ser635fs)	c.1009+1G>C p.(?)	c.2345-3C>G p.(?)	c.6530C>A p.(Ser2177Tyr)			
Parental segregation	NA	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*	NA	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*	NA	*de novo*	*de novo*	NA	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*	NA	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*			

Phénotype assocé au syndrome *MED13L* – Description d'une série de 36 patients

Conclusions de l'article

Cet article reste à ce jour la plus grande série publiée de patients porteurs d'un syndrome MED13L. A partir d'une série de 36 patients, nous avons pu identifier 9 variations faux-sens, dont plusieurs variations récurrentes impliquant des acides aminés très conservés. Le phénotype du syndrome MED13L a pu être précisé, sur le plan du neurodéveloppement avec une déficience intellectuelle modérée à sévère, des particularités du visage et la présence de signes cérébelleux, parfois progressifs.

Le phénotype associé aux variations faux-sens semble présenter un degré de sévérité plus élevé, avec notamment la description plus fréquente d'épilepsie et de troubles moteurs. Ces variations faux-sens sont par ailleurs regroupés autour des exons 15 à 17 et 25 à 31, contrairement aux variations dites tronquantes, observées sur l'ensemble du gène.

Suite à la publication de cet article, nous avons mis à jour la fiche ORPHANET du syndrome MED13L.

Profil *in silico* des variations faux-sens pathogènes du gène *MED13L*

Contexte de l'article :

La publication d'une série de 36 patients, suivie par la publication de plusieurs descriptions par d'autres équipes nous ont conduit à nous interroger sur la distribution des variations faux-sens, et le regroupement autour des exons 15 à 17 et 25 à 31 (120,122,169). En l'absence de données fonctionnelles, et de données structurales concernant la protéine MED13L, nous nous sommes penchés sur les annotations nucléotidiques et protéiques, dont nous disposions dans les bases de données afin d'identifier des paramètres discriminants entre variations pathogènes et neutres.

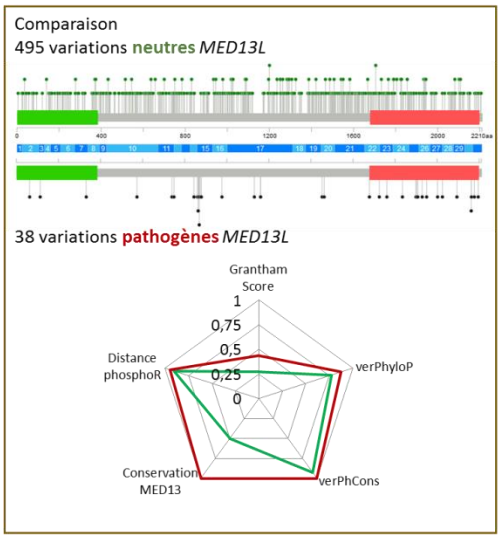
L'objectif de cette étude était d'aider à affiner la classification des variations faux-sens identifiées dans un contexte diagnostique. En effet, suite à la première publication, nous sommes sollicités pour proposer une expertise diagnostique, à la fois phénotypique et moléculaire, lors d'identification de variations faux-sens rares de *MED13L*. En l'absence de données fonctionnelles applicables en routine, nous nous sommes focalisés sur les données d'annotations dont nous disposions.

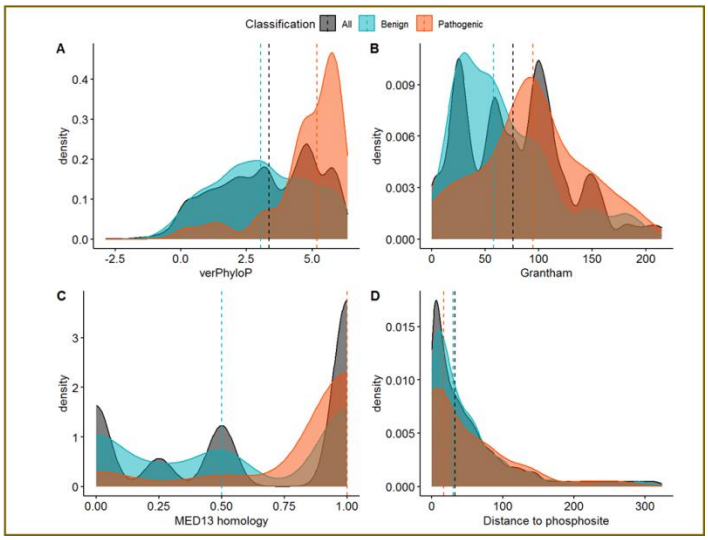
Une étude de la bibliographie, combinée à une analyse des bases de données ClinVar, DECIPHER, LOVD, denovo-db a permis d'identifier 38 variations faux-sens candidates dans le syndrome MED13L que nous avons comparées à 495 variations identifiées dans les populations contrôles (55,94,170).

Ce travail a fait l'objet d'une publication dans **European Journal of Medical Genetics en janvier 2022** (171).

J'ai présenté les résultats de cette série en communication affichée aux Journées Ouvertes en Biologie, Informatique et Mathématiques en 2017.

Performance of meta-predictors for the classification of *MED13L* missense variations, implication of raw parameters





ARTICLE 2 : PROFIL *IN SILICO* DES VARIATIONS FAUX-SENS PATHOGENES DU GÈNE *MED13L*

Performance of meta-predictors for the classification of *MED13L* missense variations, implication of raw parameters

Thomas Smol[1,2], Frédéric Frénois[1,3], Sylvie Manouvrier-Hanu[1,3], Florence Petit[1,3], Jamal Ghoumid[1,3]

[1] Université de Lille, EA7364 RADEME, F-59000, Lille, France ; [2] CHU Lille, Institut de Génétique Médicale, F-59000, Lille, France ; [3] CHU Lille, Clinique de Génétique « Guy Fontaine », F-59000, Lille, France

Abstract

MED13L syndrome is a rare congenital disorder comprising moderate intellectual disability, hypotonia and facial dysmorphism. Whole exome or genome sequencing in patients with non-specific neurodevelopmental disorders leads to identification of an increasing number of *MED13L* missense variations of unknown signification. The aim of our study was to identify relevant annotation parameters enhancing discrimination between candidate pathogenic or neutral missense variations, and to assess the performance of seven meta-predictor algorithms: BayesDel, CADD, DANN, FATHMM-XF, M-CAP, MISTIC and REVEL for the classification of *MED13L* missense variants. Significant differences were identified for five parameters: global conservation through verPhyloP and verPhCons scores; physico-chemical difference between amino acids estimated by Grantham scores; conservation of residues between MED13L and MED13 protein; proximity to phosphorylation sites for pathogenic variations. Among the seven *in-silico* tools, BayesDel, REVEL, and MISTIC provided the most interesting performances to discriminate pathogenic from neutral missense variations.

Individual gene parameter studies with *MED13L* have provided expertise on elements of annotation improving meta-predictor choices. The *in-silico* approach allows us to make useful hypotheses to predict the involvement of these amino acids in *MED13L* pathogenic missense variations.

Key-words

MED13L; missense; conservation; *in-silico* algorithm; variant interpretation

Introduction

In-silico predictive algorithms are now widely used for annotation of punctual variation in gene sequencing analysis. The American College of Medical Genetics and Genomics edited guidelines for proper use of these *in-silico* predictors [1]. Initially, first prediction strategies were based on either sequence/evolutionary conservation, or protein sequence, or supervised learning methods [2]. More complex approaches are now developed by combining multiple single tools [3]. These combinations, which define meta-predictors algorithms, often offer a continuous scoring rather than a 2-side classification. The vast majority of the studies compare the precision, accuracy and efficiency of *in-silico* tools on large multigene variation datasets such as the ClinVar database [4], and rarely focus on a specific gene for the classification of missenses variants.

The *MED13L* syndrome is an autosomal dominant syndrome due to heterozygous pathogenic variations within *MED13L*, comprising moderate intellectual disability, hypotonia, and facial dysmorphism [5–10]. *MED13L* encode the subunit 13-like protein, member of the CDK8 module kinase, part of the mediator complex [11]. The first description of pathogenic variations involved either *de novo* heterozygous nonsense, or frameshift, or intragenic microdeletion assumed to lead to haploinsufficiency [5,7,8]. Interpretations of protein-truncating variations or recurrent missense variations identified in patients with typical phenotype will not cause any particular difficulties. However, whole exome or genome sequencing in patients with non-specific neurodevelopmental disorders leads to identification of an increasing number of *MED13L* missense variations of unknown signification [9,10,12–22].

The aim of our study was to identify relevant annotation parameters enhancing discrimination between candidate pathogenic or neutral missense variations, and to assess the performance of seven meta-predictor algorithms (BayesDel [23], CADD [3], DANN [24], FATHMM-XF [25], M-CAP [26], MISTIC [27] and REVEL [28]) for the classification of known *MED13L* missense pathogenic variants.

Methods

Variation data and annotation

All possible combinations of *MED13L* missense variations were generated, corresponding to 14,545 different nucleotide substitutions. Missense variations were annotated with Ensembl Variant Effect Predictor (VEP), using genome assembly GRCh37, and with Combined Annotation Dependent Depletion (CADD), using version 1.3 [3,29]. We considered data parameters extracted from VEP and CADD annotations. Molecular anomalies with reported evidence of a splice defect, substitutions in the first or last three bases of each exon, were excluded from the analysis to prevent the pathogenicity classification associated with splice effect rather than pathogenicity due to an amino-acid substitution (n=398, 2.73%).

A comparison of conserved residues between MED13L and MED13 proteins was performed using Clustal Omega alignment with canonical transcript sequence of MED13L (Q71F56) and MED13 (Q9UHV7) obtained from Uniprot. Clustering was performed in R (v4.0.1). For each aligned amino acid, a score ranging from 0 to 1 was applied, according to the conservation of residues between MED13L and MED13. Fully conserved residues were annotated with the higher score “1”. Substitutions between two residues belonging to a group with strongly similar properties were annotated “0.5” and “0.25” respectively for groups with weakly similar properties. An unconserved substitution in the Gonnet PAM 250 matrix between MED13L and MED13 was annoted “0”. Distance to phosphorylation sites of each neutral or pathogenic missense on MED13L was estimated from PhosphoSite database on the protein model THRAP2, the previous name of MED13L [30].

Variations were classified according to ACMG 2015 criteria [1]. Frequencies of known reported alleles were considered using gnomAD, BRAVO, FREXAC and HGVD databases. Likely pathogenic and pathogenic variations described in the MED13L syndrome were identified in ClinVar, DECIPHER, HGMD and LOVD databases. Likely pathogenic and pathogenic variations previously reported were merged in the pathogenic missense group. Benign and likely benign missenses reported in ClinVar and LOVD, and missense variations above or equal to one allele in gnomAD control subpopulation, were merged in the control/neutral group.

Algorithms

Seven *in-silico* combined tools were used to analyze all the *MED13L* missense variants. All algorithms and methodology associated with the prediction tools are described in the reference articles. BayesDel is a deleteriousness measure, combining multiple annotation scores including PolyPhen2, SIFT, FATHMM, LRT, Mutation Taster, Mutation Assessor, PhyloP, GERP++, Siphy, and the minor allele frequency across populations [23]. BayesDel gave an important place to conservation measured and population frequency including ExAC [31]. CADD, Combined Annotation-Dependent Depletion, establishes a continuous Phred-like score from conservation matrix (GERP, PhastCons, PhyloP), functional annotations and protein-level scores (Grantham, SIFT, PolyPhen2) [3]. DANN, for deleterious annotation of genetic variants using neutral networks, is based on the same training data as CADD, but used a deep neural network rather a linear kernel support vector machine to score the variations [24]. FATHMM-XF scores are obtained from a supervised machine learning model using features from 27 data sets [25]. M-CAP, standing for Mendelien Clinically Applicable Pathogenicity Score, is notably based on pathogenicity scores as SIFT, PolyPhen2 and CADD [26]. MISTIC, for MISsense deleTeriousness predICTor, combines a soft voting system based on Random Forest and Logistic Regression [27]. REVEL, for Rare Exome Variant Ensemble Learning, combines 9 individual tools: MutPred, FATHMM, VEST, PolyPhen2, SIFT, PROVEAN, Mutation Assessor, Mutation Taster and LRT, and 4 conservation scores: GERP, SiPhy, PhyloP and PhastCons [28]. To identify thresholds for *MED13L* gene variations for all tools, a logistic regression has been performed with OptimalCutpoints R package, with default parameters [32].

Performance comparison

For all algorithms, we calculated the overall accuracy (ACC), specificity, sensitivity and area under curve (AUC) from Receiver operating characteristics (ROC) curves. The positive predictive values (PPV) and the negative predictive values (NPV) were determined for each algorithm. The Matthews correlation coefficient (MCC) was used to provide a balanced comparison between *in-silico* tools. Comparison of pathogenic and neutral classifications was performed with the following R-packages: rstatix, mltools and pROC.

Results

Among the 14,545 missense variations, 38 were classified as pathogenic or likely pathogenic (0.26%) and 495 as benign or likely benign (3.42%). The descriptions of missenses are listed in supplemental table 1.

Comparison of annotation parameters

To identify relevant parameters, a comparison of 19 scores provided by VEP, and based on conservation, functional and protein-level annotations, was performed. Comparisons of features between the 495 control variations and the 38 pathogenic variations with statistically significant differences were described Table 1. Statistically significant differences were identified for five parameters. Mean degree of physicochemical difference between pairs of amino acids estimated by Grantham scores was estimated at 58.00 [0.00 – 215.00] and at 94.50 [0.00 – 194.00] respectively for neutral and pathogenic variations (*p=5.31e-4, Wilcoxon test*). Significant differences in side chain atomic composition, polarity and size between the two amino acids were associated with missense pathogenic variations.

Analysis of conservation across the entire length of the MED13L protein was then performed. *MED13L* encodes a conserved protein associated with a global verPhyloP score of 3.30 [-2.87 – 6.53] over the entire protein sequence. Higher level of amino-acid and nucleotide conservations were observed for pathogenic variations compared to neutral variations with mean vertebrate verPhyloP scores respectively of 5.19 [0.17 – 6.17] and 3.04 [-0.84 – 6.32] (*p=1.14e-8*), as well as for verPhCons with mean scores of 1.00 [1.00 – 1.00] and 0.93 [0.00 – 1.00] (*p=5.17e-3, Wilcoxon test*). Therefore, MED13L appeared to be a conserved protein in which the more highly conserved amino acids could be involved in pathogenic variations. Then, we compared the conservation between both MED13 and MED13L paralogs. Sequence alignments using Clustal Omega analysis allowed to identify a homology rate of 54.80% between the paralogs, and a high degree of conservation of amino-acids involved in pathogenic variations *versus* neutral variations: median values of 1.00 *vs* 0.50 (*p=2.95e-5, Wilcoxon test*). Only 4 out of the 38 pathogenic missenses involved “unconserved” amino acids in MED13 protein: p.(Pro573Leu), p.(Leu844Ile), p.(Ser878Phe) and p.(Pro879Leu). In

MED13 protein sequence, the four amino acids corresponded respectively to Thr551, Ser810, Phe847, and Ser848 residues.

Then, we analyzed the proximity with a phosphorylation site for each missense variant. Indeed, phosphodegron motifs were previously identified in MED13L and MED13, and were found to be critical for their degradation and dissociation from the CDK8-module [33,34]. Pathogenic missense variations were significantly closer to phosphorylation sites than neutral missenses: 17.50 [0.00 – 263.00] *vs* 30.00 [0.00 – 321.00] (*p=2.95e-3, Wilcoxon test*). Twenty out of the 38 (52.6%) pathogenic missenses were identified within ±20 amino acids of a phosphorylation site.

Evaluation of in-silico classification tool performance

As a first step, optimal cut-offs were determined for the seven *in-silico* tools which provide a global raw value. Using OptimalCutpoints R package, the thresholds were defined for BayesDel, CADD, DANN, FATHMM-XF, M-CAP, MISTIC and REVEL (Table 2). The performances of the *in-silico* evaluations were also described in Table 2. The ROC curves of the individual *in-silico* combined tools for *MED13L* showed that MISTIC presented the highest AUC with a 0.902 value (Figure 1A). The ACC scores ranked MISTIC, BayesDel and REVEL as the best tools with respectively values of 0.912, 0.902 and 0.886. Based on the MCC scores, MISTIC, BayesDel and REVEL provided the best performances with respectively values of 0.533, 0.492 and 0.484. Among the seven *in-silico* tools, MISTIC, REVEL, and BayesDel provided the most interesting performances to discriminate pathogenic from neutral missense variations.

Concordance analysis

Considering all reported pathogenic missense variations, we have assessed the level of correlation of all predictors. Only 18 out of the 38 pathogenic missense variations were classified as deleterious by all tools (47.4%). Twenty-five out of the 38 variations were considered as deleterious for at least 6 tools (65.8%). The correlations of the seven *in-silico* tools were then compared to predict the pathogenic and the neutral missense classifications.

For pathogenic classifications, the strongest correlation was observed between BayesDel and REVEL (R^2 = 0.872) (Figure 1D). For the neutral group, the strongest correlation was also observed between BayesDel and REVEL (R^2 = 0.682) (Figure 1C). Indeed, combination of BayesDel, REVEL and MISTIC showed the highest rate of correct prediction targeting pathogenic variations.

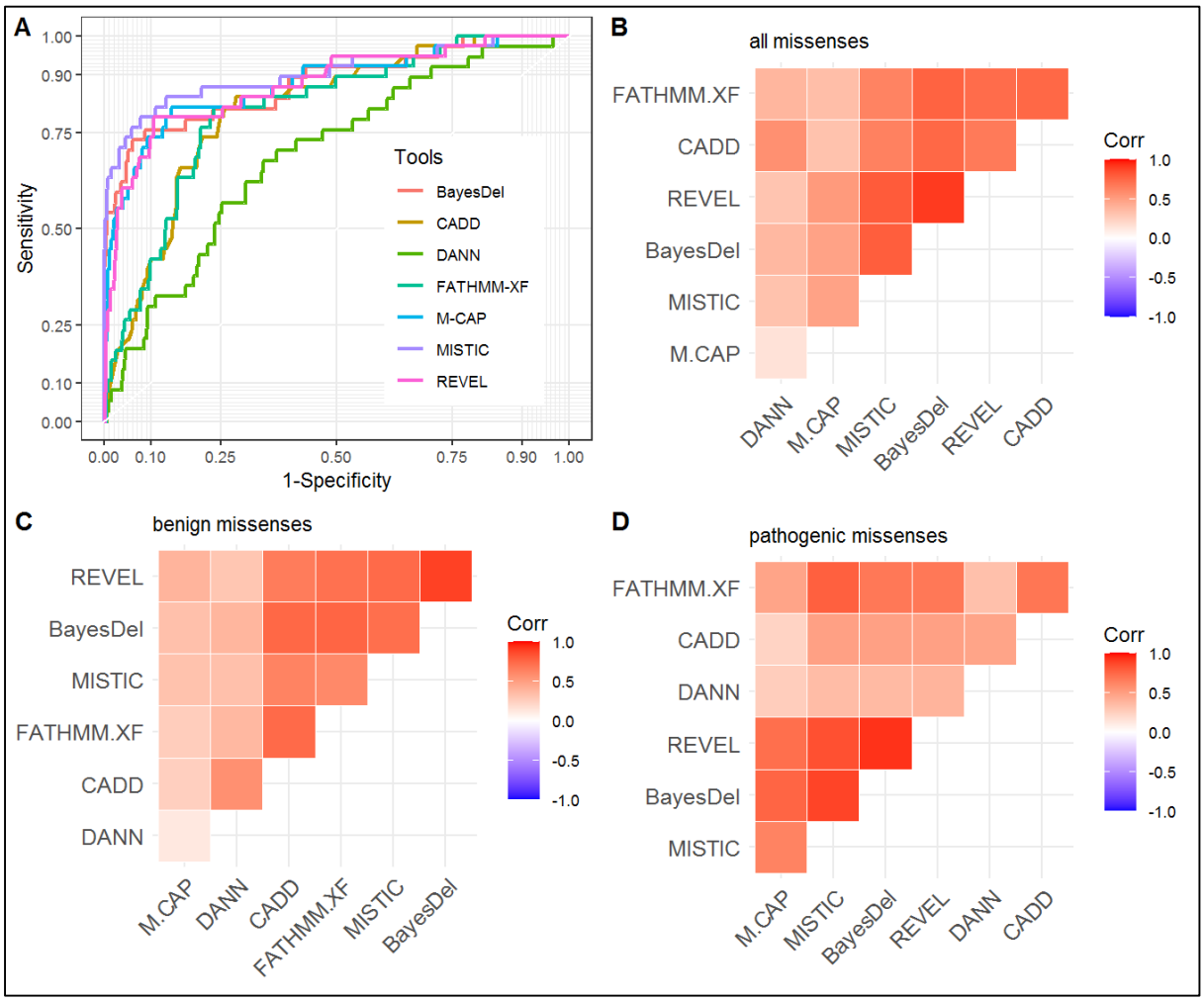


Figure 1: Comparison and correlations of meta-predictor algorithms. (A) *ROC curve performances;* ***(B)*** *Correlation regarding all missense classifications;* ***(C)*** *Correlation regarding neutral missense classifications;* ***(D)*** *Correlation regarding pathogenic missense classifications.*

Discussion

Our study targeted the most discriminant parameters and has evaluated the performances of seven *in-silico* tools in order to evaluate *MED13L* missense classification. The ACMG/AMP guideline included the concordance of computational *in-silico* predictive programs in the

criteria of variation classification, either with multiple lines of computational evidence for a deleterious effect on the gene, as PP3 criteria, or with multiple lines of computational evidence suggesting no impact on gene, as BP4 criteria [1]. When first reported *MED13L* pathogenic variations corresponded to *de novo* truncating variations as well as nonsense, frameshift, or intragenic deletions [5,7,8], the most recent now involve an important number of missense variations [9,10,35]. As the description of pathogenicity mechanisms in MED13L syndrome is still ongoing, the knowledge of *in-silico* missense annotations could improve performance in data interpretations, since computational parameters were routinely used in sequencing data analysis.

Based on all parameters evaluated, the verPhyloP scores have been considered as the most discriminant parameters with scores of 5.19 for pathogenic and 3.04 for neutral missenses (*p=1.14e-8*). The verPhyloP scores measured the evolutionary conservation at individual alignment sites in vertebrae [36]. MED13L is considered as a highly conserved protein across species with a global verPhyloP conservation score of 3.30 [11], and is involved in the conserved CDK8 kinase module. However, some regions show higher conservation scores and are associated with clustering of *bona fide* pathogenic missense variations [9,35]. The importance of the conservation score was emphasized by the comparison between the two paralogs MED13 and MED13L [37]. We undercovered a strong association between amino acids involved in pathogenic missense variations and conserved residues, within MED13 and MED13L. These results would suggest a critical role played by these residues, whose function remains to be determined. The reciprocal is also observed within MED13. The five missense variations previously described in MED13-related neurodevelopmental disorder were highly conserved: p.(Thr326Ile), p.(Pro327Gln), p.(Pro327Ser), p.(Pro540Thr) and p.(Ala2064Val) (Supplemental Table 1) [34].

The physicochemical distance between two residues, evaluated by the Grantham score, is also considered as a robust parameter to discriminate pathogenic from neutral missense variations. The most frequently substituted amino-acids in pathogenic missenses were proline, serine, and threonine residues (43%, Supplemental Table 1). The recurrence of these three amino-acids in pathogenic substitutions could suggest an implication of MED13L phosphorylation

sites [38]. One hypothesis could be an alteration of residues phosphorylated by proline-directed protein kinase targeting serine or threonine residues located next to a proline. Interestingly, CDK proteins, as well as CDK8, are members of the proline-directed serine/threonine-protein kinase family [39]. These residues could represent targets of cyclin C-CDK8 phosphorylation of MED13L altered in pathogenic missense variations [33].

Prediction of deleterious effect of *MED13L* missense variations using optimized thresholds showed relatively correct outcomes for the identification of potential neutral variations with negative predictive values ranging between 0.964 and 0.984 (Table 2). Conversely, as expected, the positive predictive values were associated with the lowest scores, ranging from 0.132 for DANN to 0.441 for MISTIC (Table 2). Among the different *in-silico* tools, MISTIC, BayesDel and REVEL have outperformed other algorithms. The approach for MISTIC, with two different classes of machine-learning algorithms, including information of multiple conservation scores and functional measures, was associated with the highest scores for specificity, PPV, MCC, ACC and AUC (Table 2). Gene level evaluation using REVEL and BayesDel tools has previously considered one of the most discriminant *in-silico* tools for clinical variation classification considering actionable genes [40].

Our data highlighted the difficulty of the application of *in-silico* parameters in routine diagnosis when these global statistical approaches were applied to one gene targeting. Individual gene parameter studies, such as *MED13L*, provide expertise on elements of annotation improving meta-predictor choices. Moreover, the *in-silico* approach allows us to make useful hypotheses to predict the involvement of these amino acids in *MED13L* pathogenic missense variations.

Acknowledgements

We gratefully acknowledge the following: all members of the EA7364 RADEME team of Lille University for their comments; European Reference Network ITHACA for support.

References

1 Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med Off J Am Coll Med Genet. 2015 May;17(5):405–24.

2 Tavtigian SV, Greenblatt MS, Lesueur F, Byrnes GB, IARC Unclassified Genetic Variants Working Group. In silico analysis of missense substitutions using sequence-alignment based methods. Hum Mutat. 2008 Nov;29(11):1327–36.

3 Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014 Mar;46(3):310–5.

4 Ghosh R, Oak N, Plon SE. Evaluation of in silico algorithms for use with ACMG/AMP clinical variant interpretation guidelines. Genome Biol. 2017 Nov;18(1):225.

5 Asadollahi R, Oneda B, Sheth F, Azzarello-Burri S, Baldinger R, Joset P, et al. Dosage changes of MED13L further delineate its role in congenital heart defects and intellectual disability. Eur J Hum Genet EJHG. 2013 Oct;21(10):1100–4.

6 Cafiero C, Marangi G, Orteschi D, Ali M, Asaro A, Ponzi E, et al. Novel de novo heterozygous loss-of-function variants in MED13L and further delineation of the MED13L haploinsufficiency syndrome. Eur J Hum Genet EJHG. 2015 Nov;23(11):1499–504.

7 van Haelst MM, Monroe GR, Duran K, van Binsbergen E, Breur JM, Giltay JC, et al. Further confirmation of the MED13L haploinsufficiency syndrome. Eur J Hum Genet EJHG. 2015 Jan;23(1):135–8.

8 Adegbola A, Musante L, Callewaert B, Maciel P, Hu H, Isidor B, et al. Redefining the MED13L syndrome. Eur J Hum Genet EJHG. 2015 Oct;23(10):1308–17.

9 Smol T, Petit F, Piton A, Keren B, Sanlaville D, Afenjar A, et al. MED13L-related intellectual disability: involvement of missense variants and delineation of the phenotype. Neurogenetics. 2018 Mar DOI: 10.1007/s10048-018-0541-0

10 Tørring PM, Larsen MJ, Brasch-Andersen C, Krogh LN, Kibæk M, Laulund L, et al. Is MED13L-related intellectual disability a recognizable syndrome? Eur J Med Genet. 2019 Feb;62(2):129–36.

11 Tsai K-L, Sato S, Tomomori-Sato C, Conaway RC, Conaway JW, Asturias FJ. A conserved Mediator-CDK8 kinase module association regulates Mediator-RNA polymerase II interaction. Nat Struct Mol Biol. 2013 May;20(5):611–9.

12 Bowling KM, Thompson ML, Amaral MD, Finnila CR, Hiatt SM, Engel KL, et al. Genomic diagnosis for children with intellectual disability and/or developmental delay. Genome Med. 2017 May;9(1):43.

13 Aoi H, Mizuguchi T, Ceroni JR, Kim VEH, Furquim I, Honjo RS, et al. Comprehensive genetic analysis of 57 families with clinically suspected Cornelia de Lange syndrome. J Hum Genet. 2019 Oct;64(10):967–78.

14 Ziats MN, Ahmad A, Bernat JA, Fisher R, Glassford M, Hannibal MC, et al. Genotype-phenotype analysis of 523 patients by genetics evaluation and clinical exome sequencing. Pediatr Res. 2020 Mar;87(4):735–9.

15 Gieldon L, Mackenroth L, Kahlert A-K, Lemke JR, Porrmann J, Schallner J, et al. Diagnostic value of partial exome sequencing in developmental disorders. PloS One. 2018;13(8):e0201041.

16 Mullegama SV, Jensik P, Li C, Dorrani N, UCLA Clinical Genomics Center, Kantarci S, et al. Coupling clinical exome sequencing with functional characterization studies to diagnose a patient with familial Mediterranean fever and MED13L haploinsufficiency syndromes. Clin Case Rep. 2017 Jun;5(6):833–40.

17 Thiffault I, Farrow E, Zellmer L, Berrios C, Miller N, Gibson M, et al. Clinical genome sequencing in an unbiased pediatric cohort. Genet Med Off J Am Coll Med Genet. 2019 Feb;21(2):303–10.

18 Gilissen C, Hehir-Kwa JY, Thung DT, van de Vorst M, van Bon BWM, Willemsen MH, et al. Genome sequencing identifies major causes of severe intellectual disability. Nature. 2014 Jul;511(7509):344–7.

19 Yamamoto T, Imaizumi T, Yamamoto-Shimojima K, Lu Y, Yanagishita T, Shimada S, et al. Genomic backgrounds of Japanese patients with undiagnosed neurodevelopmental disorders. Brain Dev. 2019 Oct;41(9):776–82.

20 Hamdan FF, Myers CT, Cossette P, Lemay P, Spiegelman D, Laporte AD, et al. High Rate of Recurrent De Novo Mutations in Developmental and Epileptic Encephalopathies. Am J Hum Genet. 2017 Nov;101(5):664–85.

21 Jiménez-Romero S, Carrasco-Salas P, Benítez-Burraco A. Language and Cognitive Impairment Associated with a Novel p.Cys63Arg Change in the MED13L Transcriptional Regulator. Mol Syndromol. 2018 Feb;9(2):83–91.

22 Martínez F, Caro-Llopis A, Roselló M, Oltra S, Mayo S, Monfort S, et al. High diagnostic yield of syndromic intellectual disability by targeted next-generation sequencing. J Med Genet. 2017 Feb;54(2):87–92.

23 Feng B-J. PERCH: A Unified Framework for Disease Gene Prioritization. Hum Mutat. 2017 Mar;38(3):243–51.

24 Quang D, Chen Y, Xie X. DANN: a deep learning approach for annotating the pathogenicity of genetic variants. Bioinformatics. 2015 Mar;31(5):761–3.

25 Rogers MF, Shihab HA, Mort M, Cooper DN, Gaunt TR, Campbell C. FATHMM-XF: accurate prediction of pathogenic point mutations via extended features. Bioinforma Oxf Engl. 2018 Feb;34(3):511–3.

26 Jagadeesh KA, Wenger AM, Berger MJ, Guturu H, Stenson PD, Cooper DN, et al. M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. Nat Genet. 2016 Dec;48(12):1581–6.

27 Chennen K, Weber T, Lornage X, Kress A, Böhm J, Thompson J, et al. MISTIC: A prediction tool to reveal disease-relevant deleterious missense variants. PloS One. 2020;15(7):e0236962.

28 Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am J Hum Genet. 2016 Oct;99(4):877–85.

29 McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, et al. The Ensembl Variant Effect Predictor. Genome Biol. 2016 Jun;17:122.

30 Hornbeck PV, Chabra I, Kornhauser JM, Skrzypek E, Zhang B. PhosphoSite: A bioinformatics resource dedicated to physiological protein phosphorylation. Proteomics. 2004 Jun;4(6):1551–61.

31 Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016 Aug;536(7616):285–91.

32 López-Ratón M, Rodríguez-Álvarez MX, Cadarso-Suárez C, Gude-Sampedro F. OptimalCutpoints: An R Package for Selecting Optimal Cutpoints in Diagnostic Tests. J Stat Softw. 2014 Nov;61(1):1–36.

33 Stieg DC, Willis SD, Ganesan V, Ong KL, Scuorzo J, Song M, et al. A complex molecular switch directs stress-induced cyclin C nuclear release through SCFGrr1-mediated degradation of Med13. Mol Biol Cell. 2018 Feb;29(3):363–75.

34 Snijders Blok L, Hiatt SM, Bowling KM, Prokop JW, Engel KL, Cochran JN, et al. De novo mutations in MED13, a component of the Mediator complex, are associated with a novel neurodevelopmental disorder. Hum Genet. 2018 May;137(5):375–88.

35 Asadollahi R, Zweier M, Gogoll L, Schiffmann R, Sticht H, Steindl K, et al. Genotype-phenotype evaluation of MED13L defects in the light of a novel truncating and a recurrent missense mutation. Eur J Med Genet. 2017 Jun DOI: 10.1016/j.ejmg.2017.06.004

36 Hubisz MJ, Pollard KS, Siepel A. PHAST and RPHAST: phylogenetic analysis with space/time models. Brief Bioinform. 2011 Jan;12(1):41–51.

37 L. Daniels D. Mutual Exclusivity of MED12/MED12L, MED13/13L, and CDK8/19 Paralogs Revealed within the CDK-Mediator Kinase Module. J Proteomics Bioinform. 2013;01(S2). DOI: 10.4172/jpb.S2-004

38 Holt LJ. Regulatory modules: Coupling protein stability to phopshoregulation during cell division. FEBS Lett. 2012 Aug;586(17):2773–7.

39 Malumbres M. Cyclin-dependent kinases. Genome Biol. 2014;15(6):122.

40 Tian Y, Pesaran T, Chamberlin A, Fenwick RB, Li S, Gau C-L, et al. REVEL and BayesDel outperform other in silico meta-predictors for clinical variant classification. Sci Rep. 2019 Sep;9(1):12752.

Table 1: Evaluation of features for pathogenic and neutral missense variations

	Pathogenic (n = 38)	Neutral (n = 495)	significance
Grantham Score	94.50 [0.00 - 194.00]	58.00 [0.00 - 215.00]	*p=5.31e-4*
verPhyloP	5.19 [0.17 - 6.17]	3.04 [-0.84 - 6.32]	*p=1.14e-8*
verPhCons	1.00 [1.00 - 1.00]	0.93 [0.00 - 1.00]	*p=5.17e-3*
MED13 residue conservations	1.00 [0.00 - 1.00]	0.50 [0.00 - 1.00]	*p=2.95e-5*
Phosphorylation site distance	17.50 [0.00 - 263.00]	30.00 [0.00 - 321.00]	*p=2.95e-3*

Table 2: Performances of *in-silico* tools

	Thresholds	Se	Sp	PPV	NPV	MCC	ACC	AUC
BayesDel	0.234	0.757	0.913	0.400	0.980	0.492	0.902	0.877
CADD	25.10	0.842	0.717	0.186	0.983	0.308	0.726	0.810
DANN	0.995	0.676	0.658	0.132	0.964	0.174	0.660	0.691
FATHMM-XF	0.768	0.816	0.764	0.209	0.982	0.333	0.767	0.807
M-CAP	0.056	0.816	0.856	0.304	0.984	0.438	0.853	0.874
MISTIC	0.663	0.789	0.923	0.441	0.983	0.533	0.912	0.902
REVEL	0.634	0.789	0.893	0.361	0.982	0.484	0.886	0.873

ACC = Accuracy ; AUC = Area Under Curve ; MCC = Matthews Correlation Coefficient ; NPV = Negative Predictive Value ; PPV = Predictive Positive Value ; Se = Sensiblity ; Sp = Specificity

Supplemental Table 1: Pathogenic and neutral evaluated missense variation

Profil *in silico* des variations faux-sens pathogènes du gène *MED13L*

Conclusions de l'article

A partir de cette étude, nous avons pu à la fois identifier des annotations associées aux variations faux-sens pathogènes, et proposer des seuils de bascule de classification pour les algorithmes de méta-predictions utilisés en routine dont CADD et REVEL.

En l'absence de données structurales précices concernant MED13L, permettant de comprendre le regroupement des variations faux-sens pathogènes, les données obtenues nous ont permis de définir des caractéristiques en lien avec ces substitutions. En effet, les variations pathogènes concernent des acides aminés fortement conservés dans l'évolution, mais également au sein du paralogue MED13 de MED13L. Leur localisation peut ainsi être en faveur de sites critiques au sein de la protéine MED13L. Nous avons également observé que la fonction de ces résidus était probablement altérée, puisque les substitutions pathogènes étaient associées à des scores de Grantham significativement élevés, témoignant d'un changement significatif des propriétés physico-chimiques entre l'acide aminé de référence et l'acide aminé de substitution. Par ailleurs, nous avons identifié des motifs de phosphorylation à proximité des substitutions pathogènes, notamment au sein du large domaine intrinséquement désorganisé. La constatation d'une co-localisation entre les substitutions pathogènes et les sites de phosphorylation a permis d'orienter les expérimentations proposées dans l'article 3 de la thèse.

Devant le degré élevé de conservation de MED13L, ainsi que les scores d'intolérance aux variations faux-sens déduits par les bases de données gnomAD et GeVIR (166,172), nous avons observé des scores élevés des seuils associés aux prédicteurs BayesDel, CADD, DANN, FATHMM, M-CAP, MISTIC et REVEL (171). Ceci témoigne de la difficulté d'analyser la pertinence de ces scores pour des variations faux-sens rares identifiées dans des gènes très conservés comme *MED13L*. Les performances observées pour REVEL et BayesDel sont concordantes avec d'autres études « gènes spécifiques » (173).

Analyse fonctionnelle des variations faux-sens de *MED13L*

Contexte de l'article :

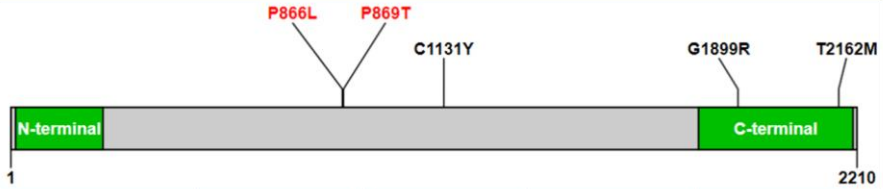
Les deux premiers travaux de la thèse ont mis en évidence la présence de particularités phénotypiques chez les patients porteurs de variations faux-sens pathogènes dans le gène *MED13L*, ainsi qu'une localisation des variations faux-sens à proximité de sites de phosphorylation et impliquant des acides aminés très conservés (62,171).

Pour comprendre les mécanismes fonctionnels pouvant être dérégulés par ces substitutions nous avons exploré plusieurs aspects des propriétés de MED13L. Nous sommes partis d'une part des données publiées concernant MED13L dans la littérature, et d'autre part des constatations des deux précédents articles. Nous avons choisi 5 variations faux-sens : p.(Pro866Leu) et p.(Pro869Leu) localisées dans l'exon 15 associées à un phénotype sévère ; p.(Cys1131Tyr) identifié également chez un patient avec une présentation sévère ; p.(Gly1899Arg), récurrente dans les publications, et p.(Thr2162Met) observées chez des patients avec une présentation phénotypique plus classique.

Ce travail est actuellement en révision.

Une partie des résultats a été présentée à l'ASHG de 2018.

***MED13L* missense variations cause multifaceted functional consequences underlying severe phenotypes**



	P866L	P869T	C1131Y	G1899R	T2162M
Phénotype	Sévère	Sévère	Sévère	Modérée	Modérée
Localisation cellulaire	Nucléaire	Nucléaire	Nucléaire	Nucléo-cytoplasmique	Nucléaire
Expression protéique	Réduite	Augmentée	Augmentée	Réduite	Augmentée
Dégradation protéasome	Non-altérée	Non-altérée	Non-altérée	Altérée	Non-altérée
Intégration CKM/Médiateur	Normale	Normale	Normale	Réduite	Normale
Etat de phosphorylation	Augmenté	Augmenté	-	-	-
Prédiction structure	-	-	Hélice α	Feuillet β	Liaisons hydrogènes
Hypothèse mécanisme	Dominant négatif	Dominant négatif	?	Perte de fonction	Perte de fonction

ARTICLE 3 : ETUDES FONCTIONNELLES DE VARIATIONS FAUX-SENS PATHOGENES DU GENE *MED13L*

MED13L missense variations cause multifaceted functional consequences underlying severe phenotypes

Thomas Smol[1,2], Frédéric Frénois[1,3], Morgane Billotte[1], Odile Boute[1,3], Roseline Caumes[1,3], Margarita Stefanova[4], Damien Sanlaville[5], Leonie Menke[6], Amara Nassar Sheikh Rashid[7], Caroline Thuillier[2], Samra Holuigue[1], Didier Monté[8], Catherine Roche-Lestienne[2], Sylvie Manouvrier-Hanu[1,3], Florence Petit[1,3], Alexis Verger[8], Jamal Ghoumid[1,3]

[1] Université de Lille, EA7364 RADEME, F-59000, Lille, France ; [2] CHU Lille, Institut de Génétique Médicale, F-59000, Lille, France ; [3] CHU Lille, Clinique de Génétique « Guy Fontaine », F-59000, Lille, France ; [4] Department of Clinical Genetics, Linköping University Hospital, Linköping, Sweden ; [5] Service de Génétique, Centre de Biologie et de Pathologie Est, Hospices Civils de Lyon, Lyon, France; Lyon Neuroscience Research Centre, GENDEV Team, CNRS UMR 5292, INSERM U1028, France; Université Claude Bernard Lyon 1, France ; [6] Department of Pediatrics, Emma Children's Hospital, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands ; [7] Department of pediatric immunology, rheumatology and infectious diseases, Emma Children's Hospital, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands ; [8] CNRS ERL 9002 Integrative Structural Biology, Inserm U 1167 – RID-AGE, Univ. Lille, CHU Lille, Institut Pasteur de Lille, F-59000 Lille, France

Abstract

Purpose: Pathogenic *MED13L* variants cause a neurodevelopmental disorder characterized by a moderate-to-severe intellectual disability. The majority of the patients carry a protein-truncating variation leading to haploinsufficiency and a typical phenotype. Few patients with a missense variant were reported with a broader and more severe phenotype. The exact biological consequences underlying atypical phenotypes remain unknown.

Method: We selected five *MED13L* reported pathogenic variants located in exon 15, 17, and 30, causing either a severe phenotype (p.Pro866Leu, p.Pro869Ser and p.Cys1131Tyr) or a typical phenotype (p.Gly1899Arg and p.Thr2162Met). To unravel pathogenic mechanisms, we analyzed MED13L wild type and mutant subcellular localization, integration into CDK8-module (CKM) and core Mediator complex, proteasome-mediated degradation, and phosphorylation.

Result: The p.Pro866Leu and p.Pro869Ser variants were likely to induce a hyper-phosphorylated status of the MED13L$^{823\text{-}930}$ domain while residues Cys1131 and Thr2162 are predicted to be critical for MED13L proper folding. Cytoplasmic relocalization and reduced integration into CKM and core Mediator complex were observed only with the p.Gly1899Arg.

Conclusion: *MED13L* missense variants are likely to be responsible for different pathogenic mechanisms. *MED13L* variants associated with typical phenotypes (p.Gly1899Arg and p.Thr2162Met) probably induce a loss of function, while variants associated with severe phenotypes probably cause a dominant-negative effect. This study highlights potential novel MED13L functions that should now be properly deciphered.

Introduction

Pathogenic variations of the Mediator complex subunit 13 like gene (*MED13L* - #MIM 608771) leading to haploinsufficiency were reported as the cause of *MED13L*-related intellectual disability *alias MED13L* syndrome.[1–4] First reported series allowed *MED13L* syndrome to be delineated in patients harboring moderate to severe global developmental delay, distinctive facial features with or without cardiac defects.[1–7] Heterozygous nonsense or frame-shift variations leading to a loss-of-function as well as intragenic or entire-gene deletions were associated with the condition.[1–3] Subsequently, several pathogenic missense variants, clustering in exons 15-17, 25, and 27-31 were reported in patients with a broader clinical presentation, including epilepsy and severe gross motor delay.[4,6,7]

MED13L functions remain largely unraveled and thus precise functional consequences of missense pathogenic variations remain unknown[3,8]. The encoded protein is the paralog of MED13, another subunit of the Mediator complex. Both subunits are members of the CDK8 kinase module (CKM).[9] The CKM is highly conserved across species and includes 4 subunits: MED12 or MED12L; MED13 or MED13L; CDK8 or CDK19, and cyclin-C. Except for cyclin-C, variants in all genes encoding the CKM components cause neurodevelopmental disorders: *MED13* is associated with an autosomal dominant intellectual disability (MIM#618009); *MED12* with Lujan-Fryns syndrome (MIM#309520), Ohdo syndrome (MIM#300895), and Opitz-Kaveggia syndrome (MIM#305450); *MED12L* with Nizon-Isidor syndrome (MIM#618872); *CDK8* with a recently described neurodevelopmental disorder (MIM#618748), and *CDK19* with a syndromic intellectual disability and epileptic encephalopathy (MIM# 618916).[10–14] Besides the CKM, the core Mediator complex is organized into three modules, *i.e.* Tail, Middle, and the Head. The Mediator plays a crucial role in assembling the Pre-Initiation Complex (PIC), taking part in the recruitment of the

RNA Polymerase II (RNA pol II) and other PIC factors to the promoter.[15] It has emerged as a master coordinator of development and cell lineage determination through interactions with various transcription factors, long intergenic noncoding RNAs (lincRNAs), and epigenetic regulators. The CKM is assumed to be involved in the repression of gene expression by preventing interactions between the core Mediator complex and the RNA polymerase II.[15] MED13 and MED13L mediate the embedding of the CKM to the core Mediator complex by interacting with at least MED19.[15] Within the CKM, MED13L can replace MED13 and conversely, in an exclusive manner depending on the cellular context.[16] Disassembly of the CKM from the core Mediator complex is mediated by MED13 or MED13L degradation, through a ubiquitin-proteasome mediated mechanism.[9] Then, the association between the core Mediator complex and the RNA polymerase II can be achieved for gene transcription.[17]

Based on these data, we aimed to unravel the biological mechanisms associated with pathogenic *MED13L* missense variations. We selected 5 variations located in exon 15, 17, and 30, causing either a severe or a typical *MED13L* syndrome phenotype (p.Pro866Leu – p.Pro869Ser – p.Cys1131Tyr – p.Gly1899Arg and p.Thr2162Met). For each substitution, we studied the subcellular localization of the encoded protein, the MED13L proteasome-mediated degradation, and the ability of the mutated protein to be embedded into the CKM and the core Mediator complex. Finally, we investigated the potential modification of the phosphorylation status induced by the pathogenic variations. We observed that different mechanisms were involved in the pathogenicity of missense variations, probably leading to either haploinsufficiency or potential gain-of-function and loss-of-function, correlating with the severity of the disease.

Materials and Methods

Plasmid construction

Expression plasmids containing either the full-length cDNA for human *MED13L* gene with an N-terminus triple 3xFLAG tag or the amino acid sequence from position 823 to 930 (*MED13L*$^{823\text{-}930}$ domain) were custom designed and obtained from VectorBuilder (Cyagen Biosciences Inc, Santa Clara, US-CA) for transfection assays. The Q5 Site-Directed Mutagenesis kit from NEB (New England Biolabs, Ipswich, US-MA) was used to introduce single nucleotide variations in the wild-type *MED13L* plasmid expression vector and the plasmid containing the *MED13L*$^{823\text{-}930}$ sequence according to the manufacturer's protocol. Mutagenic primers were designed using the NEBaseChanger tool available online. All mutagenesis products were verified by Sanger sequencing. Mutagenesis and sequencing primers are detailed in Supplementary Table 1.

Cell culture and transfection

HeLa and HEK293 cell lines were cultured in DMEM with 10% fetal bovine serum, 2mM L-glutamine, 100 UI/ml penicillin, and 100 µg/ml streptomycin at 37°C with 5% CO_2. Plasmids were transfected using Lipofectamine 3000 (Thermo Fisher, Waltham, US-MA) following the manufacturer's instructions. The day before transfection, $0.5x10^6$ cells were plated in 6-well plates. Then, cells were transfected with 2 µg of plasmid DNA of the full-length *MED13L* plasmid or 1 µg of the *MED13L*$^{823\text{-}930}$ domain-containing sequence plasmids. Experiments were performed in triplicates.

Western Blot analysis

Cells were harvested 60h after transfection. HeLa cells were then lysed in ice-cold lysis RIPA buffer complemented with Complete Protease Inhibitor Cocktail (Roche Life Science, Penzberg, Germany). Lysates were centrifuged at 15000 rpm for 5 minutes at 4°C. The

supernatant was separated from debris and protein concentration was determined by the BCA method. For Phos-tag gel assays, protein samples were dialyzed against a Tris/NaCl buffer containing Halt™ Phosphatase Inhibitor Cocktail 100X (ThermFisher Scientific, Courtaboeuf, France). For each condition, 20 µg of proteins were mixed with NuPAGE LDS Sample Buffer 4X (Thermo Fisher) and heated at 100°C for 5 minutes. Denatured proteins were separated on NuPAGE 3-8% tris-acetate protein gels (Thermo Fisher) and transferred to a nitrocellulose membrane. Blots were incubated overnight at 4°C with a Monoclonal anti-FLAG M2 Antibody (1:500; Sigma-Aldrich, Saint-Louis, US-MO) for 3xFLAG-MED13L detection and beta Tubulin Monoclonal antibody (1:5000; CliniSciences, Nanterre, France) as control. Membranes were then incubated with secondary antibody Goat anti-Mouse IgG (H+L) HRP conjugated (SouthernBiotech, Birmingham, US-AL) for 1h at RT and revealed with SuperSignal West Dura Extended Duration Substrate (Thermo Fisher). Besides the classical Western-blot approach, Phos-Tag gels were used to analyze the separation of phosphorylated protein fragments from their unphosphorylated counterparts on SDS-PAGE (Wako Pure Chemical Industries). Supercup Phos-tag gel at 12.5% was used according to the manufacturer's protocol. Quantification was performed using the ImageJ software (Rasband, Bethesda, US-MD). All experiments were conducted in triplicates.

Degradation assays

Protein stability was analyzed using the proteasome-specific inhibitor MG132 to the cell culture media. After 60h of transfection, cells were incubated with 10 µM of the proteasome inhibitor MG132 for 12h (Sigma-Aldrich). Transfections and incubations with MG132 were realized in triplicates. Quantification was performed using the ImageJ software (Rasband).

Co-immunoprecipitation

HEK-293 human cells were transfected with 6µg of pcDNA3 3xFlag wild-type and mutants *MED13L* plasmids using Lipofectamine based on the manufacturer's recommendations. Mock transfected cells were used as control. 24h later, cells were washed with PBS and lysed in 1 ml buffer containing 50 mM Tris/HCl pH 8, 250 mM NaCl, 1 mM EDTA, 1 mM TCEP, 10% glycerol, 3 mM $MgCl_2$ and 0.1% Igepal CA-630. After centrifugation, extracted proteins were incubated with 30µl of anti-Flag M2 affinity gel (Sigma A2220) on a rotation wheel overnight at 4°C. After extensive washing, the precipitated proteins were identified by Western-blot. The antibodies used to detect Mediator subunits were from Bethyl laboratories (anti-Med12 A300-774A (1:5000), anti-Med16 A303-668A (1:2500) and anti-CDK8 A302-500A (1:2500)). The antibody used to detect Flag hMED13L derivatives was from Sigma (monoclonal anti-Flag M2 peroxidase (HPR) A8592 (1:10000)). Western blot was performed using an ECL kit (Amersham GE Healthcare Life Sciences) based on the manufacturer's recommendations. HRP-conjugated secondary antibodies were used at 1:5000 and were purchased from Amersham GE Healthcare Life Sciences. All the membranes were scanned on an ImageQuant LAS 500 (GE Healthcare Life Sciences).

Immunofluorescence

For 3xFLAG-MED13L staining, HeLa cells were fixed in 2% paraformaldehyde for 40 minutes at RT, permeabilized 20 minutes at RT in permeabilization buffer (0.1% Triton X100), and then blocked 60 minutes at 4°C in blocking buffer (4% BSA, 0.05% Tween-20). Cells were incubated with anti-FLAG M2 Antibody (1:750) overnight at 4°C followed by AlexaFluor 488 anti-goat IgG (1:1000; Thermo Fisher) for 1 hour at RT. Cells were mounted in Vectashield mounting medium with DAPI containing 1.5 µg/ml DAPI. Protein localization was observed respectively in 100 cells, in 3 independent experiments (Figure 1AB).

3D structures predictions and sequence alignments

MED13L structure prediction was obtained from the AlphaFold protein structure database, based on UniProt Q71F56 protein.[18] MED13L domains were determined by homology from alignment with *S. cerevisiae* Med13 (P38931). Predictions of post-translational sites were carried out with the phosphosite database using THRAP2 model, the NetPhos 3.1 server, and the GPS 5.0 kinase-specific phosphorylation site prediction server using the Q71F56 sequence.[19–21]

Results

Phenotypic presentations linked to selected MED13L missense variations

Five previously reported pathogenic missense variations (p.Pro866Leu, p.Pro869Ser, p.Cys1131Tyr, p.Gly1899Arg, p.Thr2162Met) were studied (Table 1). Pathogenic missense variations could be divided into 2 groups. The first included the exon 15 missense variants (p.Pro866Leu and p.Pro869Ser) and p.Cys1131Tyr variant located in exon 17. Both selected variations in exon 15 led to an atypical severe phenotype combining epilepsy or abnormal EEG (4/5), severe motor phenotype with no ability to walk independently (6/6), and atypical facial features consisting of large palpebral fissures (5/5). The p.Cys1131Tyr variant led to a severe phenotype involving motor skills. At 6 years, the patient could only sit independently. The second group of variants included p.Gly1899Arg in exon 17 and p.Thr2162Met in exon 29. Both variants were identified in patients with a typical *MED13L* syndrome phenotype with moderate global developmental delay and facial particularities (Table 1).[4,7]

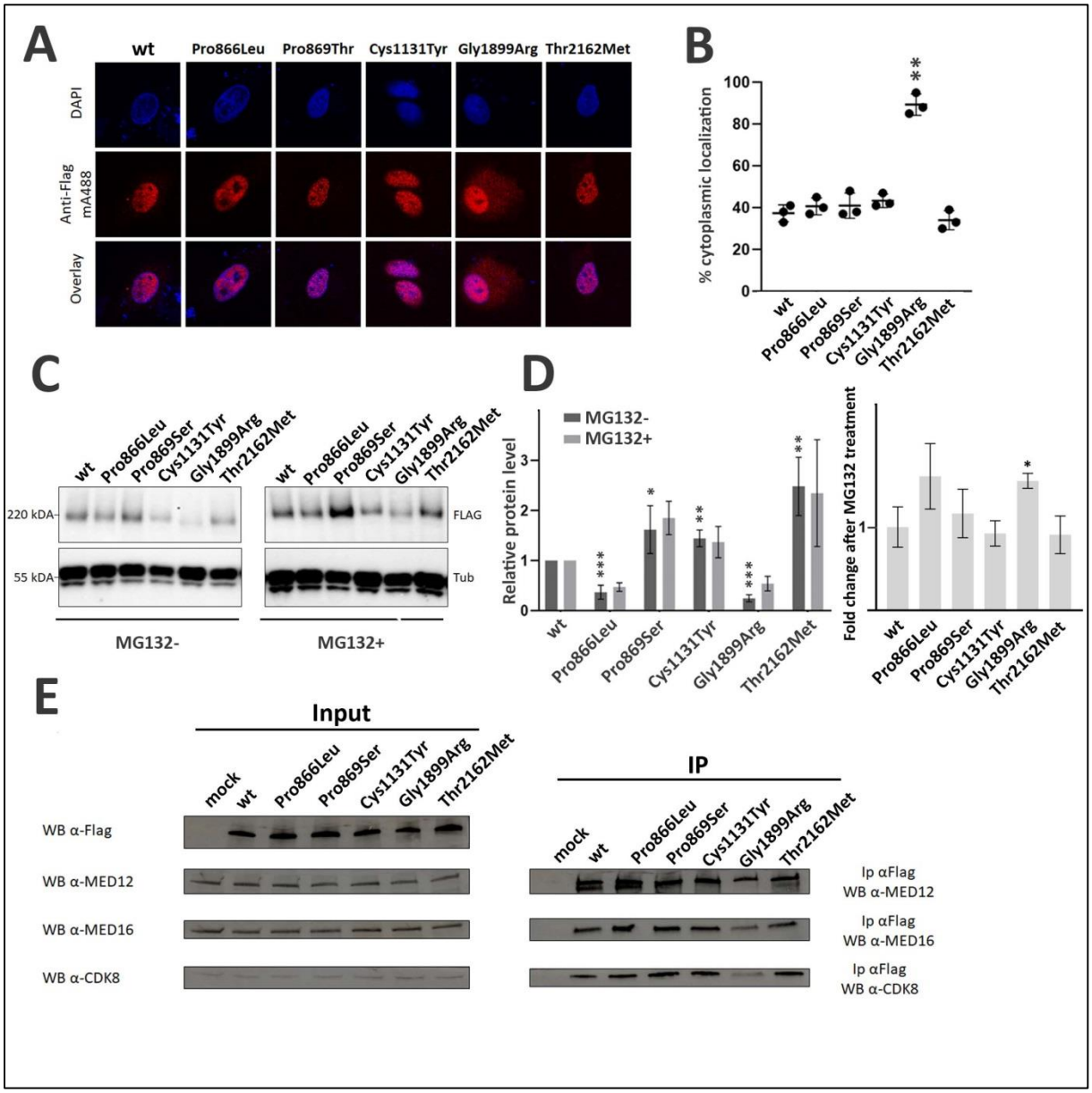


Figure 1. Localization, stability, and integration of MED13L missense pathogenic variations. (A) *Immunofluorescence studies revealed a nuclear localization for exon 15, 17, and 30 pathogenic missense variations. The p.Gly1899Arg is associated with nucleo-cytoplasmic relocalization.* ***(B)*** *Representation of the relative 3xFLAG-MED13L cytoplasmic localization. Cytoplasmic redistribution was quantified by counting the number of cells presenting with cytoplasmic staining over a total of 100 transfected cells in 3 independent experiments. Substitution p.Gly1899Arg showed an increased cytoplasmic localization compared to wild type.* ***(C)*** *Western blot showing protein levels in cells expressing wild-type or mutated MED13L, without and with MG132 proteasome specific inhibitor treatment.* ***(D)*** *Representation of the relative quantification of total MED13L without and with MG132*

*treatment, and representation of protein level fold change after MG132 treatment. Only the p.Gly1899Arg variant was associated with instability and a higher degradation level compared to other pathogenic variations. T-test results comparing MED13L wild-type and mutants are indicated thus, *P<0.05, **P<0.005, and ***P<0.0005.* ***(E)*** *Co-immunoprecipitation assays showed reduced MED13L p.Gly1899Arg variant interaction with MED12, CDK8, and MED16. 3xFLAG blot of the input samples shows an equal expression of the WT and mutant MED13L.*

Cellular localization of MED13L mutant proteins

As a subunit of the CKM, MED13L localizes into the nucleus. To identify abnormal localization associated with pathogenic missense variations, we performed immunofluorescence staining assays in the HeLa cell line. Nuclear localization was observed for MED13L wild-type, p.Pro866Leu, the p.Pro869Ser, the p.Cys1131Tyr, and p.Thr2162Met variants. A majority of nucleo-cytoplasmic diffused signals were noticed with the p.Gly1899Arg MED13L variant, with a significant difference with MED13L wild-type (89.3% versus 37.3% - $P<0.005$) (Figure 1AB). Conversely, the other pathogenic missense variations were not likely to affect the cellular localization of mutated proteins compared to MED13L wild-type.

In vitro analysis of transient expression of MED13L variations

After 60h transfection, Western-blot of whole-cell extracts from cells transiently expressing 3xFLAG-MED13L wild type and mutant revealed single bands at about 250kDa (Figure 1C). Compared to MED13L wild type level, we observed a decreased protein level in cells expressing variants p.Pro866Leu (0.37 fold $P<0.001$) and p.Gly1899Arg (0.25 fold – $P<0.0005$). Increased MED13L levels were observed in cells expressing p.Pro869Ser (1.6

fold – $P<0.05$), p.Cys1131Tyr (1.43 fold – $P<0.005$), and p.Thr2162Met (2.5 fold – $P<0.0005$) (Figure 1CD). To evaluate proteasome-mediated MED13L degradation, transfected cells were treated for 12h with 10 µmol/L of MG132, a proteasome inhibitor. Compared to MED13L wild type, only variant p.Gly1899Arg showed a significantly increased protein level after MG132 treatment (1.6 fold – $P<0.05$). This suggests that variant p.Gly1899Arg undergoes high specific proteasomal degradation (Figure 1CD).

Interaction between MED13L, CKM, and Mediator complex subunits

Direct association between MED13L and MED12 within the CKM was previously demonstrated for specific MED12 variants.[16] To test the ability of wild-type or mutated MED13L to interact with MED12 and CDK8, co-IP experiments were performed with an equal expression of the wild type and mutant MED13L in input samples. As shown in Figure 1E, using an anti-FLAG antibody for the immunoprecipitation, we detected a MED12 antibody signal with the same band intensity for the wild-type, p.Pro866Leu, p.Pro869Ser, p.Cys1131Tyr, and p.Thr2162Met variants. The intensity of the anti-MED12 signal was decreased only with the recurrent p.Gly1899Arg variant. The potential decreased interaction between the p.Gly1899Arg variant and the CKM was also noticed on the co-IP using a FLAG input and an anti-CDK8 revelation (Figure 1E). MED13L p.Gly1899Arg and Mediator complex loss of association was also observed on the co-IP using a FLAG input and an anti-MED16 revelation.

Post-translational studies

Multiple phosphorylation sites were predicted around Pro866 and Pro869 amino acid residues by PhosphoSitePlus database, NetPhos 3.1 server, and MusiteDeep server.[19,20,22] Thus, the four residues Thr867, Ser870, Ser878, and Ser889 were considered as phosphorylation sites

(data not shown). To evaluate the consequences of the p.Pro866Leu and p.Pro869Ser substitutions on phosphorylation status, cell lysates obtained from cells transiently expressing *MED13L*$^{823\text{-}930}$ were resolved on 12.5% Phos-tag gels and analyzed by Western blot (Figure 2AB). Since MED13L is a 240 kDa protein with a high number of phosphorylation sites, we use MED13L$^{823\text{-}930}$ rather than MED13L full length to enhance the capacity of the detection of differences in terms of phosphorylation.

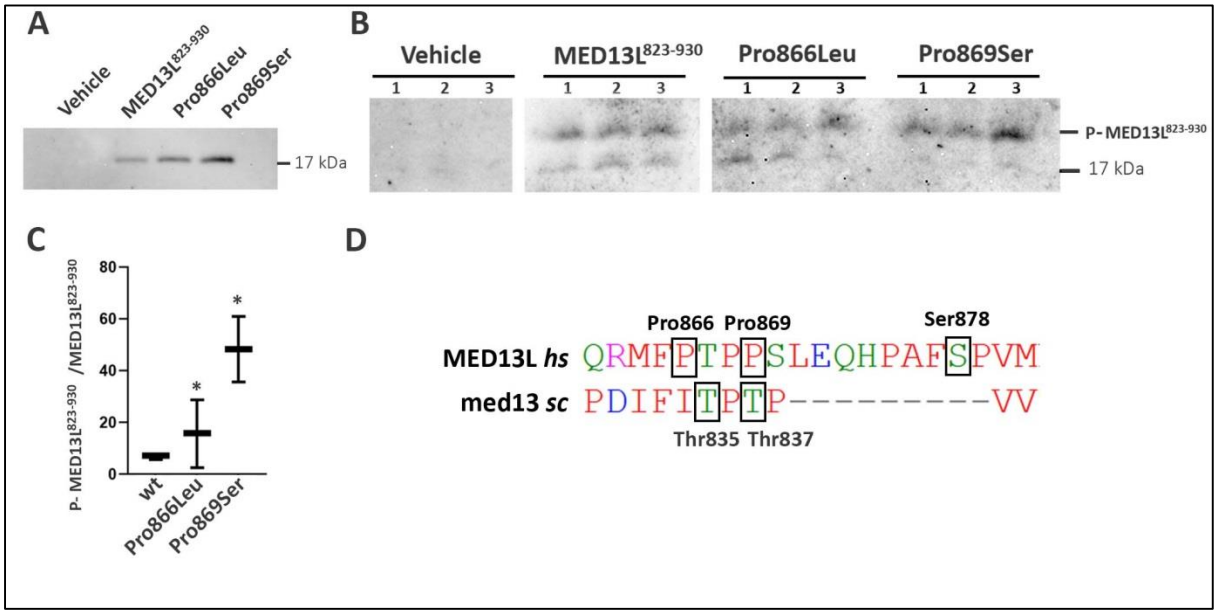


Figure 2. Detection of MED13L phosphorylation using phos-tag mobility shift assay. (A) *Expression of MED13L*$^{823\text{-}930}$ *wild type and mutated protein.* ***(B)*** *Whole-cell lysates of HeLa cells resolved on a Phos-tag gel and immunoblotted with anti-FLAG antibody, from 3 independent experiments (1, 2 and 3). P-MED13L*$^{823\text{-}930}$ *bands indicate the assumed MED13L*$^{823\text{-}930}$ *wild type and mutated phosphorylated forms, and the 17kDa band corresponds to the protein unphosphorylated forms.* ***(C)*** *Representation of the relative quantification of the MED13L*$^{823\text{-}930}$ *wild type and mutated phosphorylated relative to the unphosphorylated forms. Substitution Pro866Leu and Pro869Ser induce an increased phosphorylated status of the protein, compared to wild type. T-test results comparing MED13L wild-type and mutants are indicated thus, *P<0.05.* ***(D)*** *Alignment of the amino-acid sequence of the human MED13L*$^{862\text{-}881}$ *domain and the S. cerevisiae Med13*$^{830\text{-}840}$*. In human*

residue Ser878 is phosphorylated by CDK8. In yeast, Thr835 and Th837 stl2-mediated phosphorylation is triggered by Med13 cdk8-mediated phosphorylation.

The enhanced separation of phosphorylated proteins revealed a significantly higher level of phosphorylated status with p.Pro866Leu and p.Pro869Ser variants compared to MED13L$^{823-930}$ wild-type (respectively $P<0.05$ and $P<0.01$) (Figure 2C). The highest impact on phosphorylation status was noticed for the Pro869Ser variant. These results suggest an impact of missense located in exon 15 on phosphorylation pattern recognitions through the addition of phosphorylatable residues.

Discussion

The MED13L syndrome is characterized by the association of speech impairment, motor delay, intellectual disability, and distinctive facial features. The first reported molecular anomalies were consistent with a haploinsufficiency mechanism underlying the disease. Subsequently, description of patients harboring broader and more severe phenotypes, and carrying a missense variation have suggested that the pathogenic mechanisms could be different. Such discrepancies in variant biological consequences leading to variable phenotypes have already been reported in several genetic diseases, for instance in *SCN8A* and *GNAO1* related neurodevelopmental disorders.[23,24]

Here, we selected 5 missense variations identified in patients with the more severe spectrum of the disease, *i.e.* variants involving residues Pro866, Pro869, and Cys1131, and also 2 recurrent missense variants identified in patients with a typical MED13L-syndrome

phenotype, *i.e.* p.Gly1899Arg and p.Thr2162Met.[25–27] *MED13L* exon 15 has been reported as a hotspot of variations. Pathogenic variants cluster around 3 predicted phosphorylation sites: Thr867, Ser870, and Ser878 and are contiguous to proline residues Pro866, Pro868, Pro869, Pro875, and Pro879.[19,22] The domain between residues Asp860 and Pro879 is probably critical for MED13L function since heterozygous missense variations p.Asp860Gly, p.Met864Lys, p.Met864Leu, p.Pro866Ala, p.Pro866Leu, p.Thr867Ile, p.Pro869Ala, p.Pro869Ser, p.Pro869Thr, p.Ser878Phe, and p.Pro879Lys have been reported as pathogenic in publications or public databases, including LOVD and ClinVar.[4,8,28,29 30] The precise role of the domain remains unknown.[31] Among the exon 15 variants, p.Pro866Ala, p.Pro866Leu, p.Pro869Ser, and p.Pro869Leu were likely to induce the more severe phenotype of the disease. Patients have atypical facial features consisting of large palpebral fissures (4/4), epilepsy or abnormal EEG (5/5), inability to walk independently (6/6), and severe feeding difficulties (5/6). Clinical data of the novel patient P6 carrying the variant p.Pro869Leu support the specific phenotype associated with a variant involving the Pro869 residue (Table 1). We focused on the p.Pro866Leu variant and the recurrent p.Pro869Ser variation, that we previously identified.[4] We found that both substitutions do not modify the protein subcellular localization. When using MG132 as a specific proteasome-mediated degradation, we do not identify a higher increase in the protein level in cells expressing variants p.Pro866Leu and p.Pro869Ser, when compared to wild type (Figure 2CD). These data suggest that both variants do not show increased proteasome-mediated degradation. Since the residue Pro866 and Pro869 are close to multiple phosphorylation sites, we explored modification of the phosphorylation status due to p.Pro866Leu and p.Pro869Ser variants. Migration in Phos-tag gels followed by Western blotting revealed an increased phosphorylated status induced by the substitutions compared to $MED13L^{823-930}$ (respectively $P>0.05$ and $P>0.01$) (Figure 2ABC). The exact role of phosphorylation of the MED13L domain between residues Asp860 and

Pro879 is unknown. Both residues are located within intrinsic-disordered regions (IDR), defined by a continuous stretch of disordered promoting residues (Figure 3A). IDRs contribute to the formation of large interfaces, interacting with multiple partners. Thus, these domains have been involved in the key process of directing protein−protein interactions, including signaling and control pathways.[32]

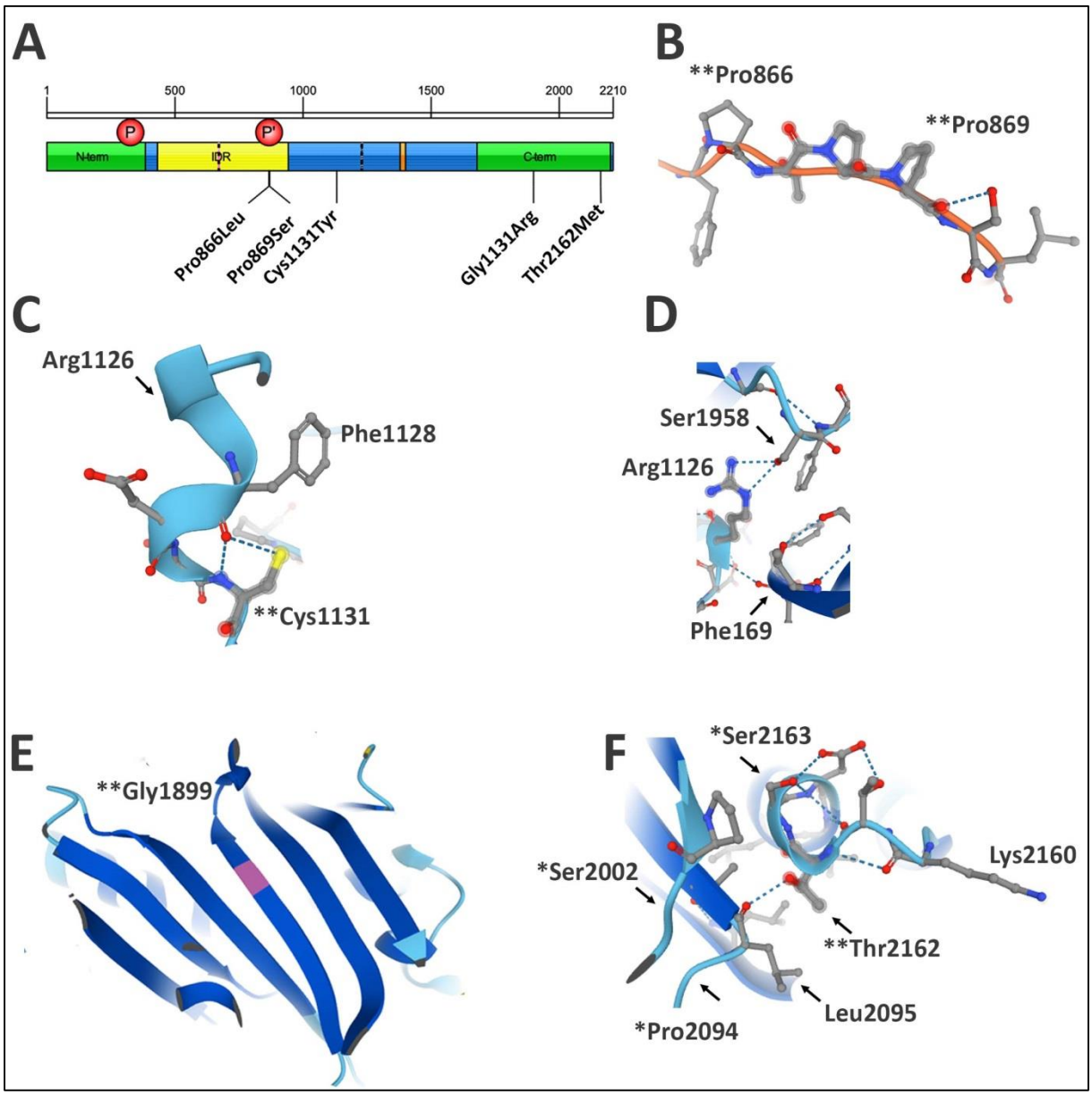


Figure 3. MED13L 3D structure prediction. (A) *Schematic representation of human MED13L with conserved N-terminal and C-terminal domains, potential intrinsic disorder region (IDR); P designates MED13L phosphodegron domain and P' designates MED13L domain located between residues*

Asp860 and Pro879, including predicted phosphorylation sites. ***(B-F)*** *AlphaFold prediction showing MED13L conformation of the domains containing the residues involved in the studied substitutions.* ***(B)*** *Residues Pro866 and Pro869 are located in an IDR, showing no specific structure.* ***(C)*** *Predicted hydrogen bound Cys1131-Phe1128 stabilizes an alpha-helix formed by residues Arg1126-Asn1127-Phe1128-Asp1129-Ser1130.* ***(D)*** *The side chain of residue Arg1126 is predicted to form two hydrogen bonds with the side chain of residue Ser1958, and the amino functional group of the residue Arg1126 is predicted to be involved in another hydrogen bond with residue Phe170 carboxylate functional group.* ***(E)*** *Residue Gly1899 is involved in a beta-strand, forming an 8-stranded antiparallel beta-sheet structure.* ***(F)*** *The residue Thr2162 side chain forms a hydrogen bond with the carboxylate functional group of the residue Leu2095. *designates residues involved in the studies substitutions. ** designates residues involved in the substitution causing MED13L-syndrome*

In humans, Ser878 residue is a high-confidence CDK8 substrate.[33] In yeast, Thr835 and Thr837 residues are involved in a three-step model of cyclin C release from Med13 following H2O2 stress (Figure 2D). The first step of this process is the priming modification of a so-called degron, located between residues 651 and 906, consisting of Cdk8-mediated phosphorylation. Once cyclin C is released, the third step of the process corresponds to slt2-mediated phosphorylation of Thr835 and Thr837 residues, leading to Med13 degradation.[34] Human MED13L and yeast Med13 protein alignments suggest that human MED13L Thr867 residue corresponds to yeast Thr835 residue (Figure 2D). It is thus conceivable that MED13L could undergo a two-kinase degron phosphorylation switch, leading to its degradation following Ser878 and Thr867 phosphorylation. This speculative mechanism would give cues to understand our results. CDK8 would be able to phosphorylate Ser878 residue, but the substitutions would impair the MED13L degradation process. This would explain the increased phosphorylated status associated with both variants and the absence of increased degradation. This mechanism is likely to be consistent with a dominant-negative effect of the variants and gives a putative explanation to the patients' severe phenotype.

The p.Cys1131Tyr variant is associated with a severe motor phenotype and no epilepsy in the reported patient.[35] Due to inter-individual variability, we cannot ascertain atypical facial gestalt in only one patient, but we can note that the patient shows short palpebral fissures.[7] We found that the protein subcellular localization was normal (Figure 1AB), and we did not identify an increased proteasome-mediated degradation (Figure 1CD). The substitution did not impair the assembly of the CKM, nor the CKM embedding into the core Mediator complex (Figure 1E). Residue Cys1131 is predicted to C-cap an alpha helix composed of residues Arg1126-Asn1127-Phe1128-Asp1129-Ser1130. Residue Cys1131 side chain is predicted to form a hydrogen bond with residue Phe1128 carboxylate functional group, stabilizing the alpha-helix (Figure 3C). The latter secondary structure is likely to be critical for MED13L proper folding, since the side chain of residue Arg1126 is predicted to form two hydrogen bonds with the side chain of residue Ser1958, and its amino functional group is predicted to be involved in another hydrogen bond with residue Phe169 carboxylate functional group (Figure 3D).[21] Substitution p.Cys1131Tyr probably disrupts the Phe1128-Cys1131 hydrogen bond, impairing the stabilization of the alpha-helix. This may partially explain the pathogenic mechanism.

Residue Gly1899 is involved in a beta-strand, forming an 8-stranded antiparallel beta-sheet structure (Figure 3E). The recurrent p.Gly1899Arg variation showed *ex vivo* a significant alteration of its stability, targeted by proteasome degradation (Figure 1ABCD). We observed that the substitution impaired MED13L interactions with CDK8 and MED12, probably leading to the reduction of the integration of the p.Gly1899Arg mutated protein into the CKM and then into the core Mediator complex (Figure 1E). This is consistent with the involvement of residue Gly1899 in a beta-sheet since these structures have been reported to facilitate protein-protein interactions.[36] The abnormal protein probably undergoes the accelerated proteasome-mediated degradation we observed (Figure 1CD). Altogether, the

pathogenic mechanism linked to the p.Gly1899Arg variant is consistent with a loss-of-function. Similar deleterious consequences were described in *MED12*-related diseases. In cancer, the p.Leu1224Phe variant was shown to affect the binding to MED13L. In FG and Lujan syndromes, p.Arg961Trp and p.Asn1007Ser variants were linked to anomalies in the recruitment of CDK8[37,38].

The p.Thr2162Met variant is also located in a hotspot of variations, including p.Thr2162Lys, p.Ser2163Leu, and p.Ser2177Tyr variants [4,39,40]. The residue Thr2162 side chain forms a hydrogen bond with the carboxylate functional group of the residue Leu2095, which is located at the N-terminus of a beta-strand involved in an antiparallel beta-sheet structure. The Leu2095-Thr2162 hydrogen bound brings close to each other residues Ser2002, Pro2094, Thr2162, and Ser2163, which are all involved in pathogenic missense variants. We can assume that the p.Thr2162Met disrupts the proper MED13L folding impairing a potential critical required proximity between these residues for MED13L function. Since residue Lys2160 is predicted to be a ubiquitination site, the folding is probably involved in the ubiquitination targeting process (Figure 3F). This could partially explain the absence of increased proteasome-mediated degradation observed and the relatively high level of the mutated protein when compared to wild type (2.5 fold – $P<0.0005$) (Figure 1CD). The typical phenotype observed in the patients would suggest a loss of function, rather than a dominant-negative effect of the substitution.

To conclude, functional analysis of *MED13L* pathogenic missense variation enabled the identification of different potential pathogenic mechanisms: dominant-negative variants associated with severe phenotypes (p.Pro866Leu, p.Pro869Ser, and p.Cys1131Tyr) and loss-of-function variants associated with a typical phenotype (p.Gly1899Arg and p.Thr2162Met). Our results allow us to partially unravel the pathogenic mechanism underlying the clinical

discrepancy in patients carrying a MED13L missense variation. Variants are likely to behave differently, impairing one of the multiple functions of MED13L. Finally, this study pointed to potential novel MED13L functions, i.e. possible involvement of exon 15 phosphorylation in MED13L degradation, involvement of Gly1899 residue in interaction allowing MED13L embedding into the CKM, and the critical role of the MED13L C-ter domain phosphorylation.

References

1. Asadollahi R, Oneda B, Sheth F, et al. Dosage changes of MED13L further delineate its role in congenital heart defects and intellectual disability. *Eur J Hum Genet*. 2013;21(10):1100-1104. doi:10.1038/ejhg.2013.17

2. van Haelst MM, Monroe GR, Duran K, et al. Further confirmation of the MED13L haploinsufficiency syndrome. *Eur J Hum Genet*. 2015;23(1):135-138. doi:10.1038/ejhg.2014.69

3. Adegbola A, Musante L, Callewaert B, et al. Redefining the MED13L syndrome. *Eur J Hum Genet*. 2015;23(10):1308-1317. doi:10.1038/ejhg.2015.26

4. Smol T, Petit F, Piton A, et al. MED13L-related intellectual disability: involvement of missense variants and delineation of the phenotype. *Neurogenetics*. 2018;19(2):93-103. doi:10.1007/s10048-018-0541-0

5. Cafiero C, Marangi G, Orteschi D, et al. Novel de novo heterozygous loss-of-function variants in MED13L and further delineation of the MED13L haploinsufficiency syndrome. *Eur J Hum Genet*. 2015;23(11):1499-1504. doi:10.1038/ejhg.2015.19

6. Asadollahi R, Zweier M, Gogoll L, et al. Genotype-phenotype evaluation of MED13L defects in the light of a novel truncating and a recurrent missense mutation. *Eur J Med Genet*. 2017;60(9):451-464. doi:10.1016/j.ejmg.2017.06.004

7. Tørring PM, Larsen MJ, Brasch-Andersen C, et al. Is MED13L-related intellectual disability a recognizable syndrome? *Eur J Med Genet*. 2019;62(2):129-136. doi:10.1016/j.ejmg.2018.06.014

8. Yi Z, Zhang Y, Song Z, et al. Report of a de novo c.2605C > T (p.Pro869Ser) change in the MED13L gene and review of the literature for MED13L-related intellectual disability. *Ital J Pediatr*. 2020;46(1):95. doi:10.1186/s13052-020-00847-y

9. Davis MA, Larimore EA, Fissel BM, Swanger J, Taatjes DJ, Clurman BE. The SCF–Fbw7 ubiquitin ligase degrades MED13 and MED13L and regulates CDK8 module association with Mediator. *Genes Dev*. 2013;27(2):151-156. doi:10.1101/gad.207720.112

10. Snijders Blok L, Hiatt SM, Bowling KM, et al. De novo mutations in MED13, a component of the Mediator complex, are associated with a novel neurodevelopmental disorder. *Hum Genet*. 2018;137(5):375-388. doi:10.1007/s00439-018-1887-y

11. Charzewska A, Maiwald R, Kahrizi K, et al. The power of the Mediator complex-Expanding the genetic architecture and phenotypic spectrum of MED12-related disorders. *Clin Genet*. 2018;94(5):450-456. doi:10.1111/cge.13412

12. Nizon M, Laugel V, Flanigan KM, et al. Variants in MED12L, encoding a subunit of the mediator kinase module, are responsible for intellectual disability associated with transcriptional defect. *Genet Med*. 2019;21(12):2713-2722. doi:10.1038/s41436-019-0557-3

13. Calpena E, Hervieu A, Kaserer T, et al. De Novo Missense Substitutions in the Gene Encoding CDK8, a Regulator of the Mediator Complex, Cause a Syndromic Developmental Disorder. *Am J Hum Genet*. 2019;104(4):709-720. doi:10.1016/j.ajhg.2019.02.006

14. Chung H-L, Mao X, Wang H, et al. De Novo Variants in CDK19 Are Associated with a Syndrome Involving Intellectual Disability and Epileptic Encephalopathy. *Am J Hum Genet*. 2020;106(5):717-725. doi:10.1016/j.ajhg.2020.04.001

15. Soutourina J. Transcription regulation by the Mediator complex. *Nat Rev Mol Cell Biol*. 2018;19(4):262-274. doi:10.1038/nrm.2017.115

16. L. Daniels D. Mutual Exclusivity of MED12/MED12L, MED13/13L, and CDK8/19 Paralogs Revealed within the CDK-Mediator Kinase Module. *Journal of Proteomics & Bioinformatics*. 2013;01(S2). doi:10.4172/jpb.S2-004

17. Tsai K-L, Sato S, Tomomori-Sato C, Conaway RC, Conaway JW, Asturias FJ. A conserved Mediator-CDK8 kinase module association regulates Mediator-RNA polymerase II interaction. *Nat Struct Mol Biol*. 2013;20(5):611-619. doi:10.1038/nsmb.2549

18. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583-589. doi:10.1038/s41586-021-03819-2

19. Hornbeck PV, Zhang B, Murray B, Kornhauser JM, Latham V, Skrzypek E. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res*. 2015;43(Database issue):D512-520. doi:10.1093/nar/gku1267

20. Blom N, Sicheritz-Pontén T, Gupta R, Gammeltoft S, Brunak S. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics*. 2004;4(6):1633-1649. doi:10.1002/pmic.200300771

21. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583-589. doi:10.1038/s41586-021-03819-2

22. Wang D, Liu D, Yuchi J, et al. MusiteDeep: a deep-learning based webserver for protein post-translational modification site prediction and visualization. *Nucleic Acids Res*. 2020;48(W1):W140-W146. doi:10.1093/nar/gkaa275

23. Talwar D, Hammer MF. SCN8A Epilepsy, Developmental Encephalopathy, and Related Disorders. *Pediatr Neurol*. 2021;122:76-83. doi:10.1016/j.pediatrneurol.2021.06.011

24. Feng H, Khalil S, Neubig RR, Sidiropoulos C. A mechanistic review on GNAO1-associated movement disorder. *Neurobiol Dis*. 2018;116:131-141. doi:10.1016/j.nbd.2018.05.005

25. Caro-Llopis A, Rosello M, Orellana C, et al. De novo mutations in genes of mediator complex causing syndromic intellectual disability: mediatorpathy or transcriptomopathy? *Pediatr Res*. 2016;80(6):809-815. doi:10.1038/pr.2016.162

26. Kosmicki JA, Samocha KE, Howrigan DP, et al. Refining the role of de novo protein-truncating variants in neurodevelopmental disorders by using population reference samples. *Nat Genet*. 2017;49(4):504-510. doi:10.1038/ng.3789

27. Deciphering Developmental Disorders Study, McRae JF, Clayton S, et al. Prevalence and architecture of *de novo* mutations in developmental disorders. *Nature*. 2017;542(7642):433-438. doi:10.1038/nature21062

28. Gilissen C, Hehir-Kwa JY, Thung DT, et al. Genome sequencing identifies major causes of severe intellectual disability. *Nature*. 2014;511(7509):344-347. doi:10.1038/nature13394

29. Thiffault I, Farrow E, Zellmer L, et al. Clinical genome sequencing in an unbiased pediatric cohort. *Genet Med*. 2019;21(2):303-310. doi:10.1038/s41436-018-0075-8

30. Mészáros B, Kumar M, Gibson TJ, Uyar B, Dosztányi Z. Degrons in cancer. *Sci Signal*. 2017;10(470). doi:10.1126/scisignal.aak9982

31. Welcker M, Clurman BE. FBW7 ubiquitin ligase: a tumour suppressor at the crossroads of cell division, growth and differentiation. *Nat Rev Cancer*. 2008;8(2):83-93. doi:10.1038/nrc2290

32. Fuxreiter M, Tóth-Petróczy Á, Kraut DA, et al. Disordered proteinaceous machines. *Chem Rev*. 2014;114(13):6806-6843. doi:10.1021/cr4007329

33. Poss ZC, Ebmeier CC, Odell AT, et al. Identification of Mediator Kinase Substrates in Human Cells using Cortistatin A and Quantitative Phosphoproteomics. *Cell Reports*. 2016;15(2):436-450. doi:10.1016/j.celrep.2016.03.030

34. Stieg DC, Willis SD, Ganesan V, et al. A complex molecular switch directs stress-induced cyclin C nuclear release through SCFGrr1-mediated degradation of Med13. *Mol Biol Cell*. 2018;29(3):363-375. doi:10.1091/mbc.E17-08-0493

35. Tørring PM, Larsen MJ, Brasch-Andersen C, et al. Is MED13L-related intellectual disability a recognizable syndrome? *Eur J Med Genet*. 2019;62(2):129-136. doi:10.1016/j.ejmg.2018.06.014

36. Remaut H, Waksman G. Protein-protein interaction through beta-strand addition. *Trends Biochem Sci*. 2006;31(8):436-444. doi:10.1016/j.tibs.2006.06.007

37. Kämpjärvi K, Kim NH, Keskitalo S, et al. Somatic MED12 mutations in prostate cancer and uterine leiomyomas promote tumorigenesis through distinct mechanisms. *Prostate*. 2016;76(1):22-31. doi:10.1002/pros.23092

38. Zhou H, Spaeth JM, Kim NH, et al. MED12 mutations link intellectual disability syndromes with dysregulated GLI3-dependent Sonic Hedgehog signaling. *Proc Natl Acad Sci U S A*. 2012;109(48):19763-19768. doi:10.1073/pnas.1121120109

39. Aoi H, Mizuguchi T, Ceroni JR, et al. Comprehensive genetic analysis of 57 families with clinically suspected Cornelia de Lange syndrome. *J Hum Genet*. 2019;64(10):967-978. doi:10.1038/s10038-019-0643-z

40. Bowling KM, Thompson ML, Amaral MD, et al. Genomic diagnosis for children with intellectual disability and/or developmental delay. *Genome Med*. 2017;9(1):43. doi:10.1186/s13073-017-0433-1

Table 1: phenotypic description of patients carrying analyzed MED13L missense variants

Patient	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
Reference	Smol eta l. (P20)	Decipher 268019	Smol et al (P28)	Smol et al (P35)	Yi et al.	Present study	Torring et al.	Caro-Llopis et al.	Decipher 262545	Decipher 323183	Smol eta l. (P14)	Bowling et al.	Decipher 272205
MED13L variation	p.Pro866Leu	p.Pro866Ala	p.Pro869Ser	p.Pro869Ser	p.Pro869Ser	p.Pro869Leu	p.Cys1131Tyr	p.Gly1899Arg	p.Gly1899Arg	p.Gly1899Arg	p.Thr2162Met	p.Thr2162Met	p.Thr2162Met
Age at examination	15 y.	NA	12 y.	24 y.	4y.5m.	3y. 3m.	6 y.	8y	NA	NA	4y.4m.	NA	NA
DD/ID	Severe	Severe	Severe	Severe	Severe	Severe	Severe	Severe	NA	ID	Moderate	Moderate	ID
Spech delay	Absent speech	NA	Absent speech	Absent speech	Absent speech	Absent speech	Absent	Absent	NA	NA	First words at 4y.	Speech delay	Absent
Motor skills	Ability to walk lost at 10y.	NA	No independent walking	Ability to walk lost at 10y.	Severe motor delay No independent walking	Walking not acquired	Head control, no independent sitting	Head control: 7 m., sitting: 17 m.	NA	NA	Walking at 26m.	NA	NA
Seizure	Late onset infantile spasms	Abnormal EEG and seizure		Lennox-Gastaut syndrom	No seizure	Abnormal EEG	No seizure	No epilepsy	NA	NA	No seizure	NA	NA

				e	Abnormal EEG (spike and slow wave rapid rhythm)								
Atypical Neurological findings	ASD	Primary microcephaly	ASD Microcrephaly	No			Major hypotonia spastic paraparesis, dystonia (extremities and tongue)	ASD with aggressive behavior	Microcephaly	Hypotonia Aggressive behavior ASD		Microcephaly	Sleep disturbance
Brain MRI	Normal	Agenisis of corpus callosum	Hypoplasia corpus callosum Delayed myelinisation				Enlarged ventricular system, partial agenesis of the corpus callosum, Dandy-Walker variant	Enlarged peri-cerebral space	NA	NA	Normal	Normal	NA
Feeding difficulties	Severe weight loss in spite of enteral	Gostrostomy tube feeding in infancy	Sond-feeded until 3 m. age	Severe weight loss in spite of	no	Severe weight loss in spite of	No	No	NA		no	no	

	nutrition			enteral nutrition		enteral nutrition – Cyclic vomiting						
Atypical facial feature	Large palpebral fissures	NA	Large palpebral fissures	Large palpebral fissures	NA	Large palpebral fissures	Blepharophimosis	No	Bilateral ptosis		no	Thick eyebrows
Miscellaneous	Nystagmus Craniosynostosis	Retinal dystrophy, severely reduced visual acuity, scoliosis		Severe myopia		Severe visual impairment - Scoliosis	Microphtalmia Bilateral corneal opacity	Renal cyst				Aplasia/hypoplasia involving bones of the hand

ASD: Autism Spectrum Disorder; EEG: Electroencephalography;NA: non-available.

Supplemental Table 1: Mutagenesis and sequencing primers

Variations	Primers	Mutagenesis primers in full length plasmid	Sequencing primers in full length plasmids	Mutagenesis primers in short length plasmids	Sequencing primers in short length plasmids
c.2597C>T ; p.(Pro866Leu)	Forward	AGGATGTTTCTCACTCCACCA	TGGACTACAAAGACCATGACGG	AGGATGTTTCTCACTCCACCA	TGGACTACAAAGACCATGACGG
	Reverse	TTGCAAGTCTGCAACTGTTG	AATTTCCTCGGGCTTGGGAC	TTGCAAGTCTGCAACTGTTG	AATTTCCTCGGGCTTGGGAC
c.2605C>A ; p.(Pro869Thr)	Forward	TCCCACTCCAACATCTTTGGA	TGGACTACAAAGACCATGACGG	TCCCACTCCAACATCTTTGGA	TGGACTACAAAGACCATGACGG
	Reverse	AACATCCTTTGCAAGTCTG	AATTTCCTCGGGCTTGGGAC	AACATCCTTTGCAAGTCTGC	AATTTCCTCGGGCTTGGGAC
c.2605C>T ; p.(Pro869Ser)	Forward	X	X	TCCCACTCCATCATCTTTGGA	TGGACTACAAAGACCATGACGG
	Reverse	X	X	AACATCCTTTGCAAGTCTG	AATTTCCTCGGGCTTGGGAC
c.2605C>G ; p.(Pro869Ala)	Forward	X	X	TCCCACTCCAGCATCTTTGGA	TGGACTACAAAGACCATGACGG
	Reverse	X	X	AACATCCTTTGCAAGTCTGC	AATTTCCTCGGGCTTGGGAC
c.3392G>A ; p.(Cys1131Tyr)	Forward	TTTGACAGCTATTGCATCTGTG	CCAGTGGTCAAGGGTCTGTT	TTTGACAGCTATTGCATCTGTG	ACAAGATGCAGCCAATTCCC
	Reverse	GTTTCTGTCTTTAAAGATATTCA	CATTAAGCGCCTCTCTGCAG	GTTTCTGTCTTTAAAGATATTCATC	AGAATCGGGGATGTAAAGCC

		TC			
c.5695G>C ; p.(Gly1899Arg)	Forward	AGTTGTAATCCGGCGACTTGG	GGGCTATTGTCTGTCTCACG	X	X
	Reverse	CTCCAGGGTAGAGATGTC	GAGCCCATTGTGACAGCATC	X	X
c.6485C>T ; p.(Thr2162Met)	Forward	TCCAAAACCATGTCGGATGTTT	CCTGGCCCTTGGGTATTTTG	TCCAAAACCACGTCGGATGTTTTAAG	TCAAAACCAGTGCCCTCTCT
	Reverse	GTCAAGAGGGTGTGGAAC	TGGCATTGTACAACTGAGTGAG	GTCAAGAGGGTGTGGAAC	ACGGGAAGGCAGGAAGTAC
All primers are described 5' to 3'.					

Analyse fonctionnelle des variations faux-sens de *MED13L*

Conclusions de l'article

Deux degrés de sévérité semblent émerger au sein du Syndrome MED13L. Les substitutions pathogènes de MED13L localisées dans l'exon 15, représentées ici par les variations faux-sens touchant les acides aminés proline 866 et 869, sont associées à un phénotype sévère du Syndrome MED13L. Celui-ci est notamment marqué par la fréquence plus importante d'épilepsies, d'un retard moteur très marqué et de la présence de fentes palpébrales élargies. Les variations faux-sens localisées en dehors de l'exon 15 sont identifiées chez des patients présentant un Syndrome MED13L plus « classique » avec une déficience intellectuelle modérée et un retard de langage et un retard moteur modéré, ainsi que des particularités du visage (62).

Fonctionnellement, la variation récurrente faux-sens p.(Gly1899Arg) possède des caractéristiques la classant parmi les variations en lien avec une haplo-insuffisance, avec une modification de la localisation cellulaire, une altération des interactions entre le CKM et le complexe médiateur, ainsi qu'une instabilité accrue de la protéine mutante. Un Syndrome MED13L « classique » est décrit pour cette variation récurrente (50,51,55,116). Cette présentation classique est également observée pour les variations p.(Cys1131Tyr) et p.(Thr2162) dont les modélisations 3D, ainsi que la localisation protéique des variations, laissent supposer une altération des structures secondaires, et une déstabilisation de l'environnement interactionnel. Les variations p.(Pro866Leu) et p.(Pro869Ser) sont, quant à elles, en lien avec une possible altération des marques de phosphorylation post-traductionnelles. La sévérité du phénotype associé pourrait être en lien avec un mécanisme physiopathologique, associé notamment à une modification des voies de signalisation

4. DISCUSSION ET CONCLUSION

L'élargissement des indications du séquençage non ciblé, chez des patients présentant des déficiences intellectuelles non syndromiques, s'est accompagné d'une détection de variations candidates basculant le concept d'approche phénotypique en première intention vers une approche génotypique. Ce changement de paradigme a mis en évidence la difficulté de se fonder sur des études de cohortes ciblées, rapportant le plus souvent des variations en lien avec une haplo-insuffisance, pour interpréter des variations faux-sens candidates, identifiées chez des patients présentant un spectre phénotypique moins classique. Ce constat est illustré par le Syndrome MED13L reposant initialement sur de nombreuses descriptions de variations dites par décalages de cadre de lecture ou non-sens associées à une probable haplo-insuffisance, et les publications récentes de variations faux-sens pathogènes (Figure 16) (62,94,105,120).

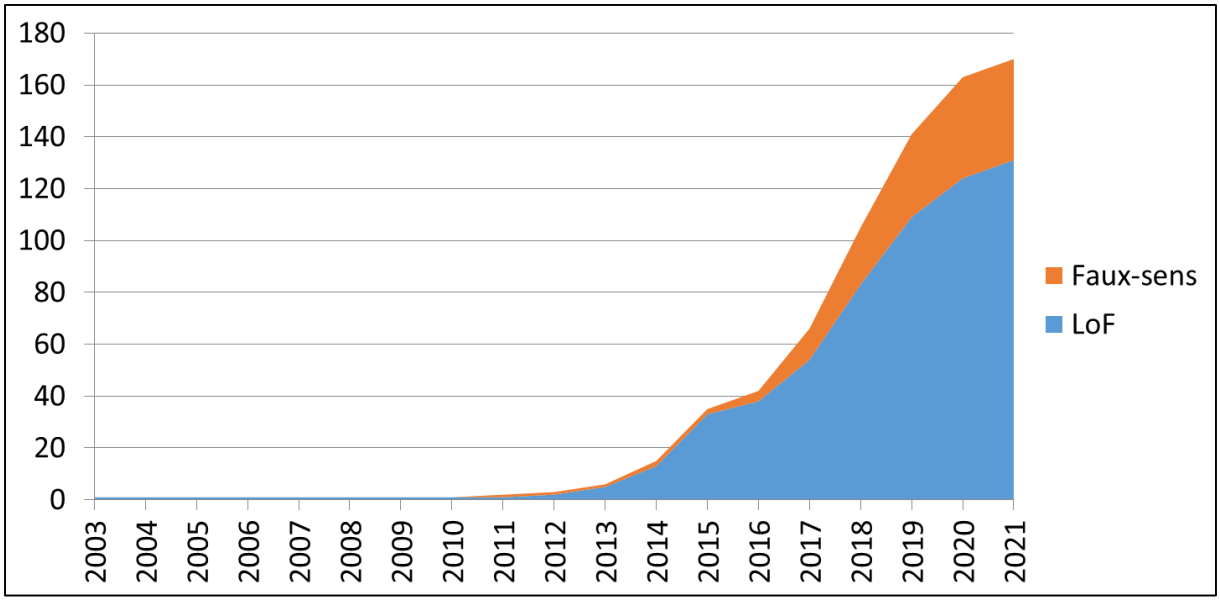


Figure 16. ***Nombre de variations candidates pathogènes dans MED13L rapportées en publication depuis 2003****. Les variations faux-sens sont en orange, les variations associées à une possible haplo-insuffisance sont en bleu (non-sens, décalage de cadre de lecture, microdélétions). LoF = variations « pertes de fonction ».*

Particularités phénotypiques des variations faux-sens

Notre publication d'une large cohorte de 36 patients présentant des variations pathogènes du gène *MED13L* nous a conduit à identifier plusieurs particularités phénotypiques et moléculaires associées aux variations faux-sens (**Article 1**) (62). Ainsi, les variations faux-sens présentent une distribution majoritairement groupée autour des exons 15 à 17 et 25 à 31

(62). Cette observation a été confirmée secondairement avec la publication des variations p.(Met864Lys) dans l'exon 15, p.(Cys1131Tyr) dans l'exon 17, p.(Leu1950Pro) dans l'exon 26 et p.(Pro2094Thr) dans l'exon 29 (105,110,121,132).

Associant des signes phénotypiques aspécifiques et retrouvés dans les déficiences intellectuelles non syndromiques, le Syndrome MED13L est le plus souvent défini cliniquement *a posteriori* des données moléculaires. Le phénotype associé aux variations faux-sens et aux variations haplo-insuffisantes présente de larges similitudes (Figure 17). Cependant la sévérité du phénotype est plus marqué chez les patients porteurs de variations faux-sens pathogènes, posant l'hypothèse de deux mécanismes conduisant à deux facettes du Syndrome MED13L (**Article 1** et **Article 3**) (62).

Par exemple, l'épilepsie est très peu rapportée chez les patients porteurs de variations pathogènes dites « tronquantes », dont p.(Ser203Phefs*) (épilepsie fébrile), p.(Lys1474*) ; p.(Pro835Leufs*) (épilepsie absence) et p.(Ala2096Glyfs*) (62,93,120). Inversement, la description d'épilepsies est fréquente (45%) chez les patients porteurs de variations faux-sens p.(Asp860Gly) (épilepsie réfractaire chez 2 patients), p.(Pro866Leu) (spasmes infantiles), p.(Pro869Ser) (n=3, dont 1 Lennox-Gastaut), p.(Pro1761Leu) (épilepsie absence), p.(Ser2163Leu), p.(2177Tyr) (épilepsie fébrile) (62,119,120,122,128).

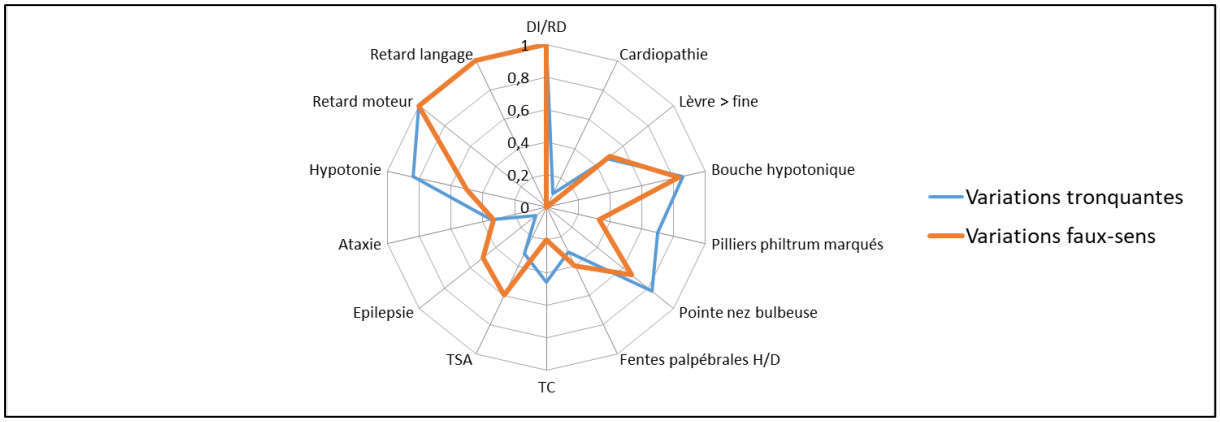


Figure 17. ***Comparaison des phénotypes associés aux variations faux-sens et pertes-de-fonction fondées sur une étude de 32 patients*** *(Communication personnelle Thomas SMOL – Assises de génétique 2018). DI/RD = Déficience intellectuelle – Retard de développement ; H/D = haut/dehors ; TC = troubles du comportement ; TSA = Troubles du spectre autistique.*

La sévérité du phénotype s'exprime également par un retard moteur très marqué chez les patients porteurs de variations faux-sens, notamment de variations de l'exon 15. Cette observation, que nous avions pu établir à partir de l'étude de cohorte, a été confirmée par

d'autres équipes (105,107,120,122). Ainsi les troubles moteurs avec une démarche ataxique associée à un retard moteur important, voire une perte de la marche, ont été observés chez plusieurs patients porteurs de variations faux-sens (62,105).

Nous avons pu noter la présence de fentes palpébrales élargies chez les patients porteurs de variations faux-sens de l'exon 15 appartenant à la petite cohorte décrite. Cette particularité semble moins marquée et moins décrite chez les patients porteurs de variations conduisant à une haplo-insuffisance (**Article 3**) (62,105,120,122).

Effet dominant négatif des variations faux-sens de l'exon 15 ?

En l'absence de caractérisation des domaines et structures protéiques de MED13L (Figure 6, Figure 7 & Figure 8), il est difficile de proposer un lien entre une possible altération d'une structure et les substitutions groupées dans l'exon 15 (Figure 18). Les données récentes publiées sur les déficiences intellectuelles en lien avec *MED13*, paralogue de *MED13L*, décrivent des variations faux-sens pathogènes groupées autour du motif phosphodegron de MED13, à proximité de l'acide aminé Thréonine 326, ce qui n'est pas le cas pour MED13L (63). Nous n'avons également pas identifié de variation faux-sens candidate dans le phosphodegron de MED13L lors de notre seconde étude recensant les variations faux-sens rapportées dans l'ensemble des bases de données disponibles en 2021 (**Article 2**) (Tableau 4 – Annexe).

Le regroupement des variations faux-sens est particulièrement remarquable entre les acides aminés 860 et 880, dans une région dont ni la conformation spatiale, ni le domaine, ni la fonction ne sont précisément connus (Figure 8B & Figure 18).

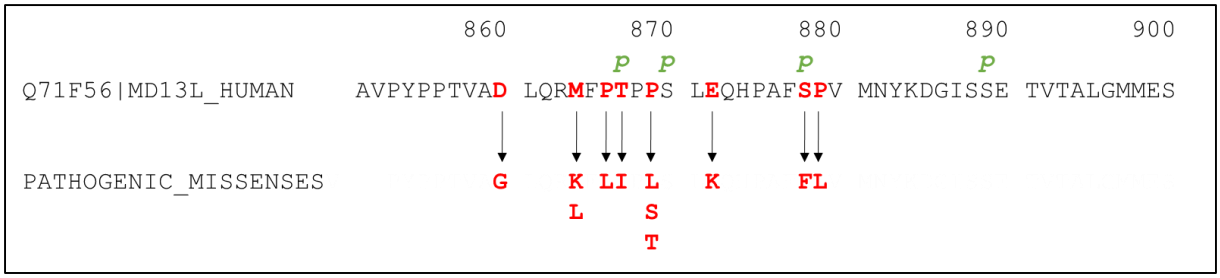


Figure 18. ***Localisation des variations faux-sens de l'exon 15****. Les sites connus de phosphorylation sont identifiés par la lettre p minuscule en vert, les acides aminés rapportés en pathologie sont en rouge et en gras, une flèche indique les substitutions décrites.*

A partir d'une étude croisée de la base de données PhosphoSite (174), nous avons identifié que les variants faux-sens de l'exon 15 impliquaient des acides-aminés localisés autour des sites de phosphorylations, c'est-à-dire la Thréonine 866, la Sérine 870, la Sérine 878 et la Sérine 889 (**Article 2**). Si la succession d'acides aminés entre les positions protéiques 860 et 880 peut mimer un phosphodegron « ΦXΦΦΦ(T/S)PPX(T/S/E/D) », l'alignement des séquences protéiques ne semble pas en faveur de cette hypothèse (175). Néanmoins, dans cette région, le résidu Sérine 878 est décrit comme une cible de phosphorylation de la cycline CDK8 (176). Chez la levure, en réponse au stress oxydatif, la phosphorylation de SSN2 (MED13) *Sce* par CDK8 *Sce* est suivie par la phosphorylation des résidus Thr835 et Thr837. L'alignement des séquences de MED13L *Hsa* et SSN2 (MED13) *Sce* semblerait indiquer que le résidu Thr837 de MED13 *Sce* correspondrait au résidu Thr867 de MED13L *Hsa*. Ainsi, nous pourrions émettre l'hypothèse que les variations de l'exon 15 p.Pro866Leu et p.Pro869Ser pourraient être impliqués dans le processus de dégradation de MED13L (Figure 10) (30,176). Ce processus, encore spéculatif pour MED13L *Hsa*, est à rapprocher de la dégradation faisant intervenir la phosphorylation multiple du motif phosphodegron en position protéique 326, suivie d'une reconnaissance par la protéine ubiquitine ligase E3 FBXW7, composant du complexe SCF (protéines SKP1, CUL1 et F-box) (30). La dégradation de MED13L ou de son paralogue MED13 est indispensable pour la libération du CKM du complexe médiateur, ce qui pourrait expliquer la sévérité du phénotype des patients. A noter qu'il existe des patients rapportés avec les variants p.(Ser878Phe) et p.(Pro879Leu). Malheureusement, il n'existe aucune précision phénotypique dans la base de données ClinVar (Tableau 4 – Annexe).

Les acides aminés impliqués dans les variations p.(Pro866Leu) et p.(Pro869Ser) sont localisés dans la région correspondante à l'IDR de MED13L *Hsa* (Figure 8 & Figure 14). Les IDR confèrent une flexibilité à la structure protéique et sont constitués d'un enchainement d'acides aminés favorisant un état désorganisé, tels que les résidus cis-proline (177). L'acide aminé Pro869 correspondrait à une cis-proline, la substitution p.Pro869Ser pourrait donc potentiellement induire un effet délétère plus large sur la structure de l'IDR. Or, ces domaines sont la cible de nombreuses modifications post-traductionnelles, grâce notamment à un enrichissement en sites de phosphorylations (178).

Importance du repliement de MED13L dans la fonction de la protéine

Le développement des prédictions de structure 3D des protéines par le logiciel d'intelligence artificielle AlphaFold est une véritable révolution, permettant de mieux appréhender les propriétés des protéines à partir de leur structure tertiaire (40). La modélisation de MED13L *Hsa* à partir de la séquence Q71F56 d'UniProt, nous a permis d'envisager que les mécanismes pathogènes associés aux variations p.Cys1131Tyr et p.Thr2162Met pouvaient être causés par une possible désorganisation de la protéine (**Article 3**). Ces deux substitutions n'entrainent pas de relocalisation de la protéine et ne sont ni associées à une augmentation de la dégradation de la protéine ni à une anomalie d'incorporation au CKM. Le résidu Cys1131 est supposé « coiffer » (C-cap) une hélice α composée des résidus Arg1126-Asn1127-Phe1128-Asp1129-Ser1130 (40). La chaîne latérale du résidu Cys1131 est prédite pour former une liaison hydrogène avec le groupe fonctionnel carboxylate du résidu Phe1128, stabilisant l'hélice α. Cette dernière structure secondaire pourrait être critique pour le repliement correct de MED13L, puisque la chaîne latérale du résidu Arg1126 est prédite pour former deux liaisons hydrogènes avec la chaîne latérale du résidu Ser1958, dont le groupement fonctionnel amine serait impliqué dans une autre liaison hydrogène avec le groupe fonctionnel carboxylate du résidu Phe169. La substitution p.Cys1131Tyr perturbe probablement la liaison hydrogène Phe1128-Cys1131, altérant la stabilisation de l'hélice α (**Article 3 - Figure 3**).

La variation p.Thr2162Met est située dans un *hotspot* de variations pathogènes, avec la description des variations pathogènes proches p.Thr2162Lys, p.Ser2163Leu, et p.Ser2177Tyr (**Article 2**) (62,134,135). La chaîne latérale du résidu Thr2162 forme une liaison hydrogène avec le groupement fonctionnel carboxylate du résidu Leu2095, situé à l'extrémité N-terminale d'un brin β impliqué dans une structure en feuillets β antiparallèles (40). La liaison hydrogène Leu2095-Thr2162 rapproche les résidus Pro2000, Ser2002, Pro2094, Thr2162, et Ser2163, impliqués dans des variations faux-sens pathogènes (51,60,62,110,135). Nous pouvons supposer que la substitution p.Thr2162Met perturbe également le repliement correct de MED13L en altérant une potentielle proximité critique requise entre tous ces résidus pour la fonction de MED13L. Puisque le résidu Lys2160 est prédit comme site d'ubiquitination, le repliement est probablement impliqué dans le processus de ciblage de l'ubiquitination (174). Cela pourrait aussi expliquer en partie l'absence d'augmentation de la dégradation médiée par le protéasome et le niveau relativement élevé de la protéine mutée par rapport au type sauvage (**Article 3**).

Importance de l'intégration de MED13L dans le CKM et le complexe médiateur

Nous avons identifié des anomalies d'incorporation dans le CKM et le complexe Médiateur uniquement pour la variation pathogène récurrente p.Gly1899Arg (50,51,55,116). Cette observation pourrait indiquer que les interactions de MED13L avec les autres sous-unités du complexe Médiateur ne seraient le fait que d'un domaine fonctionnel relativement restreint impliquant le résidu Gly1899. Ceci serait cohérent avec l'implication du résidu Gly1899 dans un feuillet β, puisque ces structures ont été rapportées comme facilitant les interactions protéine-protéine (179). D'autres expérimentations d'immuno-précipitations sont nécessaires afin de déterminer si d'autres domaines sont déterminants pour les interactions de MED13L et de ses partenaires.

Quelle approche fonctionnelle dans la réalité de la routine diagnostique ?

L'interprétation et la classification correcte de variations faux-sens, identifiées par une approche pangénomique non-ciblée, se heurtent régulièrement à l'absence de solution de validation fonctionnelle de ces substitutions. Parmi les approches les plus intéressantes, on peut noter l'utilisation de signatures épigéniques de méthylation spécifiques pour les protéines impliquées, notamment dans les processus de transferts de groupement méthyles. Une approche similaire a été conduite pour *MED12* sur une cohorte de 9 patients sans détection de profil spécifique de méthylation (180). Il ne semble pas y avoir d'études similaires conduites pour *MED13L* ou *MED13* à ce jour.

Parmi les approches que nous avons suivies pour caractériser les variations faux-sens, seule l'étude de l'état de phosphorylation a donné des résultats significatifs pour les variations p.(Pro866Leu) et p.(Pro869Ser). Cette analyse sur Phos-Tag gel présente plusieurs difficultés pour MED13L. La première concerne le nombre important de sites de phosphorylation candidats décrits sur la protéine MED13L. En effet, la base PhosphoSite dénombre 48 cibles potentielles de phosphorylation très majoritairement associées à l'IDR de MED13L (Figure 19). Partant de ce constat, il est difficile d'établir une analyse de l'état de phosphorylation sur l'ensemble de la protéine MED13L puisque le changement en lien avec une substitution sur

un petit motif de phosphorylation pourrait être masqué par le bruit de fond de l'état général de phosphorylation de la protéine.

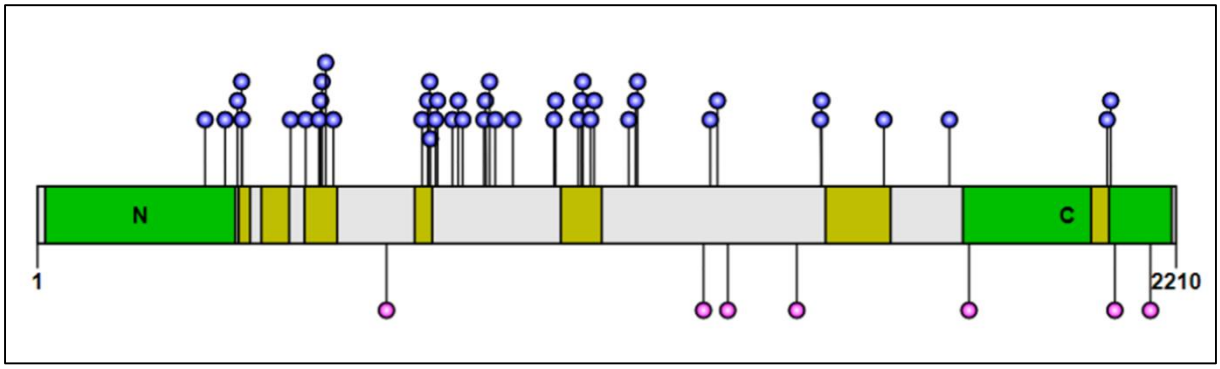


Figure 19. ***Localisation des sites potentiels de phosphorylation*** *(en bleu)* ***et d'ubiquitination*** *(en rose) sur MED13L selon la base PhosphoSite. Figure adaptée d'après (174).*

Pour obtenir des résultats exploitables, il nous a fallu passer par une étape de clonage d'une région d'intérêt, avant de pouvoir l'exprimer et l'étudier. Ces étapes intermédiaires augmentent le risque d'artificialisation du modèle avec la formation de motifs non stables ou le risque d'une non-reconnaissance de sites de phosphorylation par les protéines kinases. La mise au point technique sur plusieurs semaines de ce protocole très sensible ne rend pas cette approche facilement transposable en routine (181). Une alternative pourrait être l'étude de l'état de phosphorylation protéique par une approche par spectrométrie de masse, malheureusement associée à des coûts difficilement maîtrisables (182). Néanmoins, la proximité de sites de phosphorylations autour des variations faux-sens pathogènes souligne l'importance de l'exploration de l'état de phosphorylation comme aide à l'interprétation de certains faux-sens.

L'étude de la capacité d'association de MED13L avec ses partenaires du CKM et du complexe médiateur a permis d'associer la variation p.(Gly1899Arg) à un mécanisme probable d'haplo-insuffisance (**Article 3**). Cette variation récurrente est associée à un spectre phénotypique classique de Syndrome MED13L, sans association avec une épilepsie ou des difficultés motrices plus sévères (50,51,55,116,117). Nous n'avons pas noté de mécanisme similaire pour les autres variations faux-sens, dont les variations de l'exon 15. Cette étude par co-immunoprécipitation avait également été proposée pour étudier la variation faux-sens p.(Thr326Ala), par Davis et collaborateurs, afin de valider leur hypothèse d'une altération de l'intégration de MED13L dans le CKM consécutive à une modification du motif phosphodegron de ce dernier (30). Il est difficile de proposer cette approche *a priori* pour les

variations faux-sens de *MED13L* en dehors de ces deux cibles. En effet, il ne permettrait pas de distinguer une altération de cibles d'ubiquitination en lien avec une possible perte de fonction pour les variations proches de la lysine 2160 de *MED13L* avec un phénotype plus classique, ayant un possible effet dominant négatif suspecté pour les variations de l'exon 15 (**Article 3**). L'ensemble des variations faux-sens n'ayant pas été étudié, et les cibles d'interactions restant mal connues, cette approche pourrait être proposée en seconde intention pour des variations sans altération de la phosphorylation ou de l'ubiquitination.

Qu'attendre des données *in silico* pour la classification des variations faux-sens ?

A partir d'une comparaison des paramètres associés aux variations faux-sens pathogènes et neutres ou bénignes, nous avons pu établir un profil *in silico* des différents faux-sens (Figure 20). Une des difficultés concernant MED13L correspond à son caractère très conservé inter-espèce. Les scores de conservation, influençant de nombreux algorithmes de prédiction, sont généralement élevés, que la variation soit possiblement pathogène ou neutre/bénigne. Parmi les autres paramètres, nous avons pu identifier trois données associées à des résultats discriminants : le score de Grantham, la proximité de sites de phosphorylation et la conservation des séquences avec le paralogue MED13 (**Article 2**). Malgré tout, le nombre de variations impliquées ayant permis l'identification de ces paramètres, reste restreint, et ne permet pas d'établir un modèle spécifique à MED13L. En effet, cette approche nécessiterait de diviser les jeux de données en données tests et données d'apprentissage ce qui diminuerait encore les effectifs concernant les variations pathogènes.

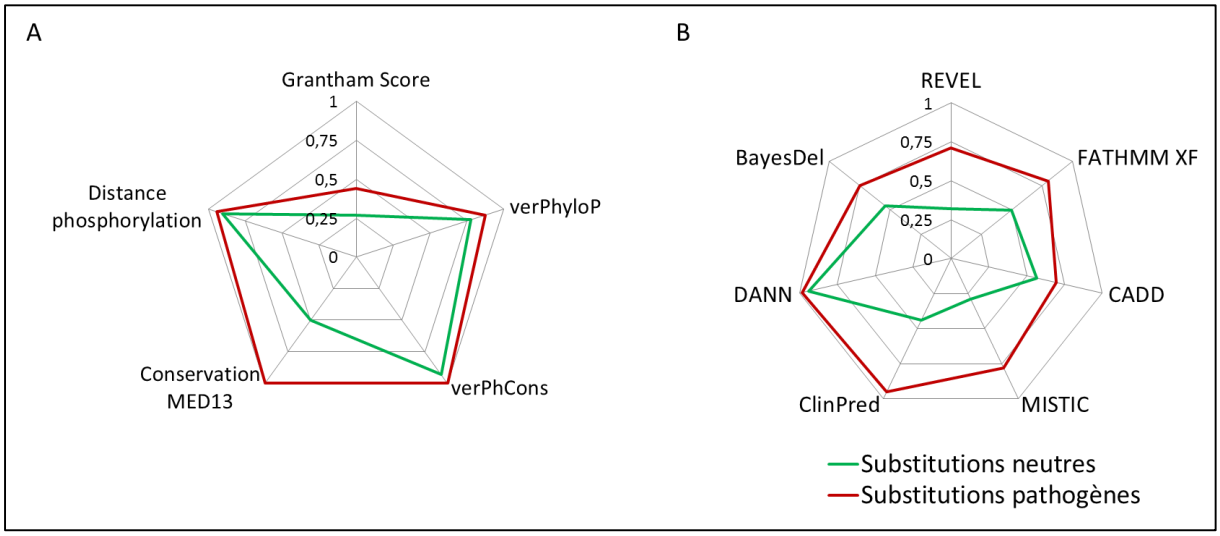


Figure 20. ***Comparaison des paramètres in silico (A) associés aux variations faux-sens et des valeurs médianes associées aux différents outils de prédictions in silico pour les substitutions pathogènes et neutres (B).***

Cette approche pourrait permettre d'organiser et de prioriser une analyse fonctionnelle d'une variation faux-sens en lien avec les différents mécanismes suspectés dans le travail, à savoir dominant négatif ou perte de fonction.

En conclusion, les approches multiples proposées dans ce travail associant données phénotypiques, fonctionnelles et bio-informatiques laissent entrevoir une complexité fonctionnelle du Syndrome MED13L. L'haplo-insuffisance, mise en avant par les premières études de variations pathogènes de *MED13L*, a d'abord servi de modèle pour affiner la description clinique du Syndrome MED13L. Ces descriptions ont permis dans un premier temps de préciser les contours du syndrome, en basculant d'une cause possible de cardiopathie cono-troncale à une déficience intellectuelle modérée à sévère (35,61,92). L'haplo-insuffisance explique également le mécanisme sous-jacent associé à certaines variations faux-sens identifiées chez des patients présentant un Syndrome MED13L dit classique. En revanche, l'hypothèse d'un second mécanisme, possiblement par effet dominant négatif semble émerger des observations faites sur les variations groupées de l'exon 15, se traduisant cliniquement par un phénotype plus sévère, et fonctionnellement par la survenue de modifications post-translationnelles probables. Cette hétérogénéité de mécanisme, suspectée dans ce syndrome rare, implique une réflexion globale lors de l'identification de variations

faux-sens candidates dépassant les prédictions *in silico* habituellement proposées. Les approches fonctionnelles ont permis d'établir un site critique dans l'association de MED13L avec le module kinase CDK8 autour de la position protéique Glycine 1899. La caractérisation des autres variations est plus complexe et pourrait faire intervenir des modifications des interactions avec des protéines partenaires, à ce jour mal connues, et/ou des changements de conformation de la protéine MED13L dans sa large région centrale intrinsèquement désorganisée.

REFERENCES

1. Allen BL, Taatjes DJ. The Mediator complex: a central integrator of transcription. Nat Rev Mol Cell Biol. 2015 Mar;16(3):155–66.

2. Tsai K-L, Sato S, Tomomori-Sato C, Conaway RC, Conaway JW, Asturias FJ. A conserved Mediator-CDK8 kinase module association regulates Mediator-RNA polymerase II interaction. Nat Struct Mol Biol. 2013 May;20(5):611–9.

3. Soutourina J. Transcription regulation by the Mediator complex. Nat Rev Mol Cell Biol. 2018 Apr;19(4):262–74.

4. Tsai K-L, Tomomori-Sato C, Sato S, Conaway RC, Conaway JW, Asturias FJ. Subunit architecture and functional modular rearrangements of the transcriptional mediator complex. Cell. 2014 Jun 5;157(6):1430–44.

5. Cevher MA, Shi Y, Li D, Chait BT, Malik S, Roeder RG. Reconstitution of active human core Mediator complex reveals a critical role of the MED14 subunit. Nat Struct Mol Biol. 2014 Dec;21(12):1028–34.

6. Tsai K-L, Yu X, Gopalan S, Chao T-C, Zhang Y, Florens L, et al. Mediator structure and rearrangements required for holoenzyme formation. Nature. 2017 Apr 13;544(7649):196–201.

7. L. Daniels D. Mutual Exclusivity of MED12/MED12L, MED13/13L, and CDK8/19 Paralogs Revealed within the CDK-Mediator Kinase Module. Journal of Proteomics & Bioinformatics [Internet]. 2013 [cited 2016 Jul 19];01(S2). Available from: http://www.omicsonline.org/mutual-exclusivity-of-med12-med12l-med13-13l-and-cdk8-19-paralogs-revealed-within-the-cdk-mediator-kinase-module-jpb.S2-004.php?aid=15770

8. Knuesel MT, Meyer KD, Bernecky C, Taatjes DJ. The human CDK8 subcomplex is a molecular switch that controls Mediator coactivator function. Genes Dev. 2009 Feb 15;23(4):439–51.

9. Amoasii L, Holland W, Sanchez-Ortiz E, Baskin KK, Pearson M, Burgess SC, et al. A MED13-dependent skeletal muscle gene program controls systemic glucose homeostasis and hepatic metabolism. Genes Dev. 2016 Feb 15;30(4):434–46.

10. Baskin KK, Grueter CE, Kusminski CM, Holland WL, Bookout AL, Satapati S, et al. MED13-dependent signaling from the heart confers leanness by enhancing metabolism in adipose tissue and liver. EMBO Mol Med. 2014 Dec;6(12):1610–21.

11. Musante L, Bartsch O, Ropers H-H, Kalscheuer VM. cDNA cloning and characterization of the human THRAP2 gene which maps to chromosome 12q24, and its mouse ortholog Thrap2. Gene. 2004 May 12;332:119–27.

12. Hamada N, Iwamoto I, Nishikawa M, Nagata K-I. Expression Analyses of Mediator Complex Subunit 13-Like: A Responsible Gene for Neurodevelopmental Disorders during Mouse Brain Development. Dev Neurosci. 2021;43(1):43–52.

13. Sierecki E. The Mediator complex and the role of protein-protein interactions in the gene regulation machinery. Seminars in Cell & Developmental Biology. 2020 Mar 1;99:20–30.

14. Soutourina J, Wydau S, Ambroise Y, Boschiero C, Werner M. Direct Interaction of RNA Polymerase II and Mediator Required for Transcription in Vivo. Science. 2011 Mar 18;331(6023):1451–4.

15. Imasaki T, Calero G, Cai G, Tsai K-L, Yamada K, Cardelli F, et al. Architecture of the Mediator head module. Nature. 2011 Jul 3;475(7355):240–3.

16. Grünberg S, Henikoff S, Hahn S, Zentner GE. Mediator binding to UASs is broadly uncoupled from transcription and cooperative with TFIID recruitment to promoters. EMBO J. 2016 Nov 15;35(22):2435–46.

17. Kim YJ, Björklund S, Li Y, Sayre MH, Kornberg RD. A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. Cell. 1994 May 20;77(4):599–608.

18. Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, et al. Mediator and cohesin connect gene expression and chromatin architecture. Nature. 2010 Sep 23;467(7314):430–5.

19. Benetatos L, Vartholomatos G. Enhancer DNA methylation in acute myeloid leukemia and myelodysplastic syndromes. Cell Mol Life Sci. 2018 Jun;75(11):1999–2009.

20. El Khattabi L, Zhao H, Kalchschmidt J, Young N, Jung S, Van Blerkom P, et al. A Pliable Mediator Acts as a Functional Rather Than an Architectural Bridge between Promoters and Enhancers. Cell. 2019 Aug 22;178(5):1145-1158.e20.

21. Dixon JR, Gorkin DU, Ren B. Chromatin Domains: The Unit of Chromosome Organization. Molecular Cell. 2016 Jun 2;62(5):668–80.

22. Chereji RV, Bharatula V, Elfving N, Blomberg J, Larsson M, Morozov AV, et al. Mediator binds to boundaries of chromosomal interaction domains and to proteins involved in DNA looping, RNA metabolism, chromatin remodeling, and actin assembly. Nucleic Acids Res. 2017 Sep 6;45(15):8806–21.

23. Georges A, Gopaul D, Denby Wilkes C, Giordanengo Aiach N, Novikova E, Barrault M-B, et al. Functional interplay between Mediator and RNA polymerase II in Rad2/XPG loading to the chromatin. Nucleic Acids Res. 2019 Sep 26;47(17):8988–9004.

24. Nizon M, Laugel V, Flanigan KM, Pastore M, Waldrop MA, Rosenfeld JA, et al. Variants in MED12L, encoding a subunit of the mediator kinase module, are responsible for intellectual disability associated with transcriptional defect. Genet Med. 2019 Dec;21(12):2713–22.

25. van Weerd JH, Mohan RA, van Duijvenboden K, Hooijkaas IB, Wakker V, Boukens BJ, et al. Trait-associated noncoding variant regions affect TBX3 regulation and cardiac conduction. Moskowitz IP, Stainier DY, Moskowitz IP, Firulli T, editors. eLife. 2020 Jul 16;9:e56697.

26. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell. 2014 Dec 18;159(7):1665–80.

27. Fishilevich S, Nudel R, Rappaport N, Hadar R, Plaschkes I, Iny Stein T, et al. GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. Database (Oxford). 2017 01;2017.

28. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. Genome Res. 2002 Jun;12(6):996–1006.

29. Fant CB, Taatjes DJ. Regulatory functions of the Mediator kinases CDK8 and CDK19. Transcription. 2019 Apr;10(2):76–90.

30. Davis MA, Larimore EA, Fissel BM, Swanger J, Taatjes DJ, Clurman BE. The SCF–Fbw7 ubiquitin ligase degrades MED13 and MED13L and regulates CDK8 module association with Mediator. Genes Dev. 2013 Jan 15;27(2):151–6.

31. Zhang N, Song Y, Xu Y, Liu J, Shen Y, Zhou L, et al. MED13L integrates Mediator-regulated epigenetic control into lung cancer radiosensitivity. Theranostics. 2020;10(20):9378–94.

32. Willis SD, Stieg DC, Ong KL, Shah R, Strich AK, Grose JH, et al. Snf1 cooperates with the CWI MAPK pathway to mediate the degradation of Med13 following oxidative stress. Microb Cell. 2018 Jun 25;5(8):357–70.

33. Li Y-C, Chao T-C, Kim HJ, Cholko T, Chen S-F, Li G, et al. Structure and noncanonical Cdk8 activation mechanism within an Argonaute-containing Mediator kinase module. Science Advances. 2021 Jan 1;7(3):eabd4484.

34. Liu W, Xie Y, Ma J, Luo X, Nie P, Zuo Z, et al. IBS: an illustrator for the presentation and visualization of biological sequences. Bioinformatics. 2015 Oct 15;31(20):3359–61.

35. Muncke N, Jung C, Rüdiger H, Ulmer H, Roeth R, Hubert A, et al. Missense mutations and gene interruption in PROSIT240, a novel TRAP240-like gene, in patients with congenital heart defect (transposition of the great arteries). Circulation. 2003 Dec 9;108(23):2843–50.

36. Savkur R s., Burris T p. The coactivator LXXLL nuclear receptor recognition motif. The Journal of Peptide Research. 2004;63(3):207–12.

37. Ito M, Yuan CX, Malik S, Gu W, Fondell JD, Yamamura S, et al. Identity between TRAP and SMCC complexes indicates novel pathways for the function of nuclear receptors and diverse mammalian activators. Mol Cell. 1999 Mar;3(3):361–70.

38. Meister G. Argonaute proteins: functional insights and emerging roles. Nat Rev Genet. 2013 Jul;14(7):447–59.

39. Duret L, Gasteiger E, Perrière G. LALNVIEW: a graphical viewer for pairwise sequence alignments. Comput Appl Biosci. 1996 Dec;12(6):507–10.

40. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021 Aug;596(7873):583–9.

41. Ravid T, Hochstrasser M. Degradation signal diversity in the ubiquitin-proteasome system. Nat Rev Mol Cell Biol. 2008 Sep;9(9):679–90.

42. Buendía-Monreal M, Gillmor CS. Mediator: A key regulator of plant development. Developmental Biology. 2016 Nov 1;419(1):7–18.

43. Pucheta-Martinez E, D'Amelio N, Lelli M, Martinez-Torrecuadrada JL, Sudol M, Saladino G, et al. Changes in the folding landscape of the WW domain provide a molecular mechanism for an inherited genetic syndrome. Sci Rep. 2016 Jul 26;6:30293.

44. Liu Y, Huang W-J, Lin M-T, Li J-J, Zhang J-Y. Chapter 13 - Proteasome Inhibitors as Sensitizing Agents for Cancer Chemotherapy. In: Chen Z-S, Yang D-H, editors. Protein Kinase Inhibitors as Sensitizing Agents for Chemotherapy [Internet]. Academic Press; 2019 [cited 2021 Oct 17]. p. 207–28. (Cancer Sensitizing Agents for Chemotherapy; vol. 4). Available from: https://www.sciencedirect.com/science/article/pii/B9780128164358000134

45. Stieg DC, Cooper KF, Strich R. The extent of cyclin C promoter occupancy directs changes in stress-dependent transcription. Journal of Biological Chemistry. 2020 Nov 27;295(48):16280–91.

46. Angus SP, Nevins JR. A role for Mediator complex subunit MED13L in Rb/E2F-induced growth arrest. Oncogene. 2012 Nov 1;31(44):4709–17.

47. Utami KH, Winata CL, Hillmer AM, Aksoy I, Long HT, Liany H, et al. Impaired development of neural-crest cell-derived organs and intellectual disability caused by MED13L haploinsufficiency. Hum Mutat. 2014 Nov;35(11):1311–20.

48. Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature. 2009 May;459(7243):108–12.

49. Tomppo L, Ekelund J, Lichtermann D, Veijola J, Järvelin M-R, Hennah W. DISC1 conditioned GWAS for psychosis proneness in a large Finnish birth cohort. PLoS One. 2012;7(2):e30643.

50. Turner TN, Wilfert AB, Bakken TE, Bernier RA, Pepper MR, Zhang Z, et al. Sex-Based Analysis of De Novo Variants in Neurodevelopmental Disorders. Am J Hum Genet. 2019 Dec 5;105(6):1274–85.

51. Kosmicki JA, Samocha KE, Howrigan DP, Sanders SJ, Slowikowski K, Lek M, et al. Refining the role of de novo protein-truncating variants in neurodevelopmental disorders by using population reference samples. Nat Genet. 2017 Apr;49(4):504–10.

52. Lim ET, Uddin M, De Rubeis S, Chan Y, Kamumbu AS, Zhang X, et al. Rates, distribution and implications of postzygotic mosaic mutations in autism spectrum disorder. Nat Neurosci. 2017 Sep;20(9):1217–24.

53. Satterstrom FK, Kosmicki JA, Wang J, Breen MS, De Rubeis S, An J-Y, et al. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. Cell. 2020 Feb 6;180(3):568-584.e23.

54. Heinzen EL, O'Neill AC, Zhu X, Allen AS, Bahlo M, Chelly J, et al. De novo and inherited private variants in MAP1B in periventricular nodular heterotopia. PLoS Genet. 2018 May;14(5):e1007281.

55. Deciphering Developmental Disorders Study, McRae JF, Clayton S, Fitzgerald TW, Kaplanis J, Prigmore E, et al. Prevalence and architecture of *de novo* mutations in developmental disorders. Nature. 2017 Feb;542(7642):433–8.

56. Chen X, Liu F, Mar Aung Z, Zhang Y, Chai G. Whole-Exome Sequencing Reveals Rare Germline Mutations in Patients With Hemifacial Microsomia. Front Genet. 2021;12:580761.

57. Iossifov I, O'Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, et al. The contribution of de novo coding mutations to autism spectrum disorder. Nature. 2014 Nov 13;515(7526):216–21.

58. Monies D, Abouelhoda M, AlSayed M, Alhassnan Z, Alotaibi M, Kayyali H, et al. The landscape of genetic diseases in Saudi Arabia based on the first 1000 diagnostic panels and exomes. Hum Genet. 2017 Aug;136(8):921–39.

59. Cappi C, Oliphant ME, Péter Z, Zai G, Conceição do Rosário M, Sullivan CAW, et al. De Novo Damaging DNA Coding Mutations Are Associated With Obsessive-Compulsive Disorder and Overlap With Tourette's Disorder and Autism. Biol Psychiatry. 2020 Jun 15;87(12):1035–44.

60. Takata A, Nakashima M, Saitsu H, Mizuguchi T, Mitsuhashi S, Takahashi Y, et al. Comprehensive analysis of coding variants highlights genetic complexity in developmental and epileptic encephalopathy. Nat Commun. 2019 Jun 7;10(1):2506.

61. Asadollahi R, Oneda B, Sheth F, Azzarello-Burri S, Baldinger R, Joset P, et al. Dosage changes of MED13L further delineate its role in congenital heart defects and intellectual disability. Eur J Hum Genet. 2013 Oct;21(10):1100–4.

62. Smol T, Petit F, Piton A, Keren B, Sanlaville D, Afenjar A, et al. MED13L-related intellectual disability: involvement of missense variants and delineation of the phenotype. Neurogenetics. 2018 May;19(2):93–103.

63. Snijders Blok L, Hiatt SM, Bowling KM, Prokop JW, Engel KL, Cochran JN, et al. De novo mutations in MED13, a component of the Mediator complex, are associated with a novel neurodevelopmental disorder. Hum Genet. 2018 May 1;137(5):375–88.

64. Schwartz CE, Tarpey PS, Lubs HA, Verloes A, May MM, Risheg H, et al. The original Lujan syndrome family has a novel missense mutation (p.N1007S) in the MED12 gene. J Med Genet. 2007 Jul;44(7):472–7.

65. Vulto-van Silfhout AT, de Vries BBA, van Bon BWM, Hoischen A, Ruiterkamp-Versteeg M, Gilissen C, et al. Mutations in MED12 cause X-linked Ohdo syndrome. Am J Hum Genet. 2013 Mar 7;92(3):401–6.

66. Clark RD, Graham JM, Friez MJ, Hoo JJ, Jones KL, McKeown C, et al. FG syndrome, an X-linked multiple congenital anomaly syndrome: the clinical phenotype and an algorithm for diagnostic testing. Genet Med. 2009 Nov;11(11):769–75.

67. Calpena E, Hervieu A, Kaserer T, Swagemakers SMA, Goos JAC, Popoola O, et al. De Novo Missense Substitutions in the Gene Encoding CDK8, a Regulator of the Mediator Complex, Cause a Syndromic Developmental Disorder. Am J Hum Genet. 2019 Apr 4;104(4):709–20.

68. Chung H-L, Mao X, Wang H, Park Y-J, Marcogliese PC, Rosenfeld JA, et al. De Novo Variants in CDK19 Are Associated with a Syndrome Involving Intellectual Disability and Epileptic Encephalopathy. Am J Hum Genet. 2020 May 7;106(5):717–25.

69. Zarate YA, Uehara T, Abe K, Oginuma M, Harako S, Ishitani S, et al. CDK19-related disorder results from both loss-of-function and gain-of-function de novo missense variants. Genet Med. 2021 Jun;23(6):1050–7.

70. Kaufmann R, Straussberg R, Mandel H, Fattal-Valevski A, Ben-Zeev B, Naamati A, et al. Infantile cerebral and cerebellar atrophy is associated with a mutation in the MED17 subunit of the transcription preinitiation mediator complex. Am J Hum Genet. 2010 Nov 12;87(5):667–70.

71. Vodopiutz J, Schmook MT, Konstantopoulou V, Plecko B, Greber-Platzer S, Creus M, et al. MED20 mutation associated with infantile basal ganglia degeneration and brain atrophy. Eur J Pediatr. 2015 Jan;174(1):113–8.

72. Hashimoto S, Boissel S, Zarhrate M, Rio M, Munnich A, Egly J-M, et al. MED23 mutation links intellectual disability to dysregulation of immediate early gene expression. Science. 2011 Aug 26;333(6046):1161–3.

73. Figueiredo T, Melo US, Pessoa ALS, Nobrega PR, Kitajima JP, Correa I, et al. Homozygous missense mutation in MED25 segregates with syndromic intellectual disability in a large consanguineous family. J Med Genet. 2015 Feb;52(2):123–7.

74. Basel-Vanagaite L, Smirin-Yosef P, Essakow JL, Tzur S, Lagovsky I, Maya I, et al. Homozygous MED25 mutation implicated in eye-intellectual disability syndrome. Hum Genet. 2015 Jun;134(6):577–87.

75. Meng L, Isohanni P, Shao Y, Graham BH, Hickey SE, Brooks S, et al. MED27 Variants Cause Developmental Delay, Dystonia, and Cerebellar Hypoplasia. Ann Neurol. 2021 Apr;89(4):828–33.

76. De Nardi L, Faletra F, D'Adamo AP, Bianco AMR, Athanasakis E, Bruno I, et al. Could the MED13 mutations manifest as a Kabuki-like syndrome? Am J Med Genet A. 2021 Feb;185(2):584–90.

77. Lesca G, Moizard M-P, Bussy G, Boggio D, Hu H, Haas SA, et al. Clinical and neurocognitive characterization of a family with a novel MED12 gene frameshift mutation. Am J Med Genet A. 2013 Dec;161A(12):3063–71.

78. Polla DL, Bhoj EJ, Verheij JBGM, Wassink-Ruiter JSK, Reis A, Deshpande C, et al. De novo variants in MED12 cause X-linked syndromic neurodevelopmental disorders in 18 females. Genet Med. 2021 Apr;23(4):645–52.

79. Donnio L-M, Bidon B, Hashimoto S, May M, Epanchintsev A, Ryan C, et al. MED12-related XLID disorders are dose-dependent of immediate early genes (IEGs) expression. Human Molecular Genetics. 2017 Jun 1;26(11):2062–75.

80. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum Mutat. 2015 Oct;36(10):928–30.

81. Bancerek J, Poss ZC, Steinparzer I, Sedlyarov V, Pfaffenwimmer T, Mikulic I, et al. CDK8 Kinase Phosphorylates Transcription Factor STAT1 to Selectively Regulate the Interferon Response. Immunity. 2013 Feb 21;38(2):250–62.

82. Agostini A, Marchetti D, Izzi C, Cocco I, Pinelli L, Accorsi P, et al. Expanding the phenotype of MED 17 mutations: Description of two new cases and review of the literature. Am J Med Genet B Neuropsychiatr Genet. 2018 Dec;177(8):687–90.

83. Fattal-Valevski A, Ben Sira L, Lerman-Sagie T, Strausberg R, Bloch-Mimouni A, Edvardson S, et al. Delineation of the phenotype of MED17-related disease in Caucasus-Jewish families. Eur J Paediatr Neurol. 2021 May;32:40–5.

84. Nozawa K, Schneider TR, Cramer P. Core Mediator structure at 3.4 Å extends model of transcription initiation complex. Nature. 2017 May 11;545(7653):248–51.

85. Trehan A, Brady JM, Maduro V, Bone WP, Huang Y, Golas GA, et al. MED23-associated intellectual disability in a non-consanguineous family. Am J Med Genet A. 2015 Jun;167(6):1374–80.

86. Lionel AC, Monfared N, Scherer SW, Marshall CR, Mercimek-Mahmutoglu S. MED23-associated refractory epilepsy successfully treated with the ketogenic diet. Am J Med Genet A. 2016 Sep;170(9):2421–5.

87. Hashemi-Gorji F, Fardaei M, Tabei SMB, Miryounesi M. Novel mutation in the MED23 gene for intellectual disability: A case report and literature review. Clin Case Rep. 2019 Feb;7(2):331–5.

88. Monté D, Clantin B, Dewitte F, Lens Z, Rucktooa P, Pardon E, et al. Crystal structure of human Mediator subunit MED23. Nat Commun. 2018 Aug 23;9(1):3389.

89. Haynes D, Pollack L, Prasad C, Goobie S, Colaiacovo S, Wolfinger T, et al. Further delineation of Basel-Vanagaite-Smirin-Yosef syndrome: Report of three patients. Am J Med Genet A. 2020 Jul;182(7):1785–90.

90. Mittler G, Stühler T, Santolin L, Uhlmann T, Kremmer E, Lottspeich F, et al. A novel docking site on Mediator is critical for activation by VP16 in mammalian cells. EMBO J. 2003 Dec 15;22(24):6494–504.

91. van Haelst MM, Monroe GR, Duran K, van Binsbergen E, Breur JM, Giltay JC, et al. Further confirmation of the MED13L haploinsufficiency syndrome. Eur J Hum Genet. 2015 Jan;23(1):135–8.

92. Adegbola A, Musante L, Callewaert B, Maciel P, Hu H, Isidor B, et al. Redefining the MED13L syndrome. Eur J Hum Genet. 2015 Oct;23(10):1308–17.

93. Cafiero C, Marangi G, Orteschi D, Ali M, Asaro A, Ponzi E, et al. Novel de novo heterozygous loss-of-function variants in MED13L and further delineation of the MED13L haploinsufficiency syndrome. Eur J Hum Genet. 2015 Nov;23(11):1499–504.

94. Deciphering Developmental Disorders Study. Large-scale discovery of novel genetic causes of developmental disorders. Nature. 2015 Mar 12;519(7542):223–8.

95. de Ligt J, Willemsen MH, van Bon BWM, Kleefstra T, Yntema HG, Kroes T, et al. Diagnostic exome sequencing in persons with severe intellectual disability. N Engl J Med. 2012 Nov 15;367(20):1921–9.

96. Ockeloen CW, Willemsen MH, de Munnik S, van Bon BWM, de Leeuw N, Verrips A, et al. Further delineation of the KBG syndrome phenotype caused by ANKRD11 aberrations. Eur J Hum Genet. 2015 Sep;23(9):1176–85.

97. Hoyer J, Ekici AB, Endele S, Popp B, Zweier C, Wiesener A, et al. Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability. Am J Hum Genet. 2012 Mar 9;90(3):565–72.

98. Baer S, Afenjar A, Smol T, Piton A, Gérard B, Alembik Y, et al. Wiedemann-Steiner syndrome as a major cause of syndromic intellectual disability: A study of 33 French cases. Clin Genet. 2018 Jul;94(1):141–52.

99. Snijders Blok L, Madsen E, Juusola J, Gilissen C, Baralle D, Reijnders MRF, et al. Mutations in DDX3X Are a Common Cause of Unexplained Intellectual Disability with Gender-Specific Effects on Wnt Signaling. Am J Hum Genet. 2015 Aug 6;97(2):343–52.

100. Helsmoortel C, Vulto-van Silfhout AT, Coe BP, Vandeweyer G, Rooms L, van den Ende J, et al. A SWI/SNF related autism syndrome caused by de novo mutations in ADNP. Nat Genet. 2014 Apr;46(4):380–4.

101. Wolff M, Johannesen KM, Hedrich UBS, Masnada S, Rubboli G, Gardella E, et al. Genetic and phenotypic heterogeneity suggest therapeutic implications in SCN2A-related disorders. Brain. 2017 May 1;140(5):1316–36.

102. Courcet J-B, Faivre L, Malzac P, Masurel-Paulet A, Lopez E, Callier P, et al. The DYRK1A gene is a cause of syndromic intellectual disability with severe microcephaly and epilepsy. J Med Genet. 2012 Dec;49(12):731–6.

103. Roelfsema JH, White SJ, Ariyürek Y, Bartholdi D, Niedrist D, Papadia F, et al. Genetic heterogeneity in Rubinstein-Taybi syndrome: mutations in both the CBP and EP300 genes cause disease. Am J Hum Genet. 2005 Apr;76(4):572–80.

104. Grozeva D, Carss K, Spasic-Boskovic O, Parker MJ, Archer H, Firth HV, et al. De novo loss-of-function mutations in SETD5, encoding a methyltransferase in a 3p25 microdeletion syndrome critical region, cause intellectual disability. Am J Hum Genet. 2014 Apr 3;94(4):618–24.

105. Tørring PM, Larsen MJ, Brasch-Andersen C, Krogh LN, Kibæk M, Laulund L, et al. Is MED13L-related intellectual disability a recognizable syndrome? European Journal of Medical Genetics [Internet]. [cited 2018 Jul 3]; Available from: https://www.sciencedirect.com/science/article/pii/S1769721218301447

106. Al-Dewik N, Mohd H, Al-Mureikhi M, Ali R, Al-Mesaifri F, Mahmoud L, et al. Clinical exome sequencing in 509 Middle Eastern families with suspected Mendelian diseases: The Qatari experience. Am J Med Genet A. 2019 Jun;179(6):927–35.

107. Jiménez-Romero S, Carrasco-Salas P, Benítez-Burraco A. Language and Cognitive Impairment Associated with a Novel p.Cys63Arg Change in the MED13L Transcriptional Regulator. Mol Syndromol. 2018 Feb;9(2):83–91.

108. Sabo A, Murdock D, Dugan S, Meng Q, Gingras M-C, Hu J, et al. Community-based recruitment and exome sequencing indicates high diagnostic yield in adults with intellectual disability. Mol Genet Genomic Med. 2020 Oct;8(10):e1439.

109. Guo H, Wang T, Wu H, Long M, Coe BP, Li H, et al. Inherited and multiple de novo mutations in autism/developmental delay risk genes suggest a multifactorial model. Mol Autism. 2018;9:64.

110. Ziats MN, Ahmad A, Bernat JA, Fisher R, Glassford M, Hannibal MC, et al. Genotype-phenotype analysis of 523 patients by genetics evaluation and clinical exome sequencing. Pediatr Res. 2020 Mar;87(4):735–9.

111. Kosaki R, Kubota M, Uehara T, Suzuki H, Takenouchi T, Kosaki K. Consecutive medical exome analysis at a tertiary center: Diagnostic and health-economic outcomes. Am J Med Genet A. 2020 Jul;182(7):1601–7.

112. Hamdan FF, Myers CT, Cossette P, Lemay P, Spiegelman D, Laporte AD, et al. High Rate of Recurrent De Novo Mutations in Developmental and Epileptic Encephalopathies. Am J Hum Genet. 2017 Nov 2;101(5):664–85.

113. Matthews AM, Blydt-Hansen I, Al-Jabri B, Andersen J, Tarailo-Graovac M, Price M, et al. Atypical cerebral palsy: genomics analysis enables precision medicine. Genet Med. 2019 Jul;21(7):1621–8.

114. Aspromonte MC, Bellini M, Gasparini A, Carraro M, Bettella E, Polli R, et al. Characterization of intellectual disability and autism comorbidity through gene panel sequencing. Hum Mutat. 2019 Sep;40(9):1346–63.

115. Carraro M, Monzon AM, Chiricosta L, Reggiani F, Aspromonte MC, Bellini M, et al. Assessment of patient clinical descriptions and pathogenic variants from gene panel sequences in the CAGI-5 intellectual disability challenge. Hum Mutat. 2019 Sep;40(9):1330–45.

116. Caro-Llopis A, Rosello M, Orellana C, Oltra S, Monfort S, Mayo S, et al. De novo mutations in genes of mediator complex causing syndromic intellectual disability: mediatorpathy or transcriptomopathy? Pediatr Res. 2016 Dec;80(6):809–15.

117. Martínez F, Caro-Llopis A, Roselló M, Oltra S, Mayo S, Monfort S, et al. High diagnostic yield of syndromic intellectual disability by targeted next-generation sequencing. J Med Genet. 2017 Feb;54(2):87–92.

118. Yamamoto T, Imaizumi T, Yamamoto-Shimojima K, Lu Y, Yanagishita T, Shimada S, et al. Genomic backgrounds of Japanese patients with undiagnosed neurodevelopmental disorders. Brain Dev. 2019 Oct;41(9):776–82.

119. Gilissen C, Hehir-Kwa JY, Thung DT, van de Vorst M, van Bon BWM, Willemsen MH, et al. Genome sequencing identifies major causes of severe intellectual disability. Nature. 2014 Jul 17;511(7509):344–7.

120. Asadollahi R, Zweier M, Gogoll L, Schiffmann R, Sticht H, Steindl K, et al. Genotype-phenotype evaluation of MED13L defects in the light of a novel truncating and a recurrent missense mutation. Eur J Med Genet. 2017 Sep;60(9):451–64.

121. Thiffault I, Farrow E, Zellmer L, Berrios C, Miller N, Gibson M, et al. Clinical genome sequencing in an unbiased pediatric cohort. Genet Med. 2019 Feb;21(2):303–10.

122. Yi Z, Zhang Y, Song Z, Pan H, Yang C, Li F, et al. Report of a de novo c.2605C > T (p.Pro869Ser) change in the MED13L gene and review of the literature for MED13L-related intellectual disability. Ital J Pediatr. 2020 Jul 9;46(1):95.

123. Wang T, Guo H, Xiong B, Stessman HAF, Wu H, Coe BP, et al. De novo genic mutations among a Chinese autism spectrum disorder cohort. Nat Commun. 2016 Nov 8;7:13316.

124. Guo H, Duyzend MH, Coe BP, Baker C, Hoekzema K, Gerdts J, et al. Genome sequencing identifies multiple deleterious variants in autism patients with more severe phenotypes. Genet Med. 2019 Jul;21(7):1611–20.

125. Du Y, Li Z, Liu Z, Zhang N, Wang R, Li F, et al. Nonrandom occurrence of multiple de novo coding variants in a proband indicates the existence of an oligogenic model in autism. Genet Med. 2020 Jan;22(1):170–80.

126. Najmabadi H, Hu H, Garshasbi M, Zemojtel T, Abedini SS, Chen W, et al. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. Nature. 2011 Sep 21;478(7367):57–63.

127. Popp B, Ekici AB, Thiel CT, Hoyer J, Wiesener A, Kraus C, et al. Exome Pool-Seq in neurodevelopmental disorders. Eur J Hum Genet. 2017 Dec;25(12):1364–76.

128. Mullegama SV, Jensik P, Li C, Dorrani N, UCLA Clinical Genomics Center, Kantarci S, et al. Coupling clinical exome sequencing with functional characterization studies to diagnose a patient with familial Mediterranean fever and MED13L haploinsufficiency syndromes. Clin Case Rep. 2017 Jun;5(6):833–40.

129. Capalbo A, Valero RA, Jimenez-Almazan J, Pardo PM, Fabiani M, Jiménez D, et al. Optimizing clinical exome design and parallel gene-testing for recessive genetic conditions in preconception carrier screening: Translational research genomic data from 14,125 exomes. PLoS Genet. 2019 Oct;15(10):e1008409.

130. Bessenyei B, Balogh I, Mokanszki A, Ujfalusi A, Pfundt R, Szakszon K. MED13L-related intellectual disability due to paternal germinal mosaicism. Cold Spring Harb Mol Case Stud. 2021 Oct 15;mcs.a006124.

131. Retterer K, Juusola J, Cho MT, Vitazka P, Millan F, Gibellini F, et al. Clinical application of whole-exome sequencing across clinical indications. Genet Med. 2016 Jul;18(7):696–704.

132. Gieldon L, Mackenroth L, Kahlert A-K, Lemke JR, Porrmann J, Schallner J, et al. Diagnostic value of partial exome sequencing in developmental disorders. PLoS One. 2018;13(8):e0201041.

133. Takata A, Miyake N, Tsurusaki Y, Fukai R, Miyatake S, Koshimizu E, et al. Integrative Analyses of De Novo Mutations Provide Deeper Biological Insights into Autism Spectrum Disorder. Cell Rep. 2018 Jan 16;22(3):734–47.

134. Aoi H, Mizuguchi T, Ceroni JR, Kim VEH, Furquim I, Honjo RS, et al. Comprehensive genetic analysis of 57 families with clinically suspected Cornelia de Lange syndrome. J Hum Genet. 2019 Oct;64(10):967–78.

135. Bowling KM, Thompson ML, Amaral MD, Finnila CR, Hiatt SM, Engel KL, et al. Genomic diagnosis for children with intellectual disability and/or developmental delay. Genome Med. 2017 May 30;9(1):43.

136. Wai HA, Lord J, Lyon M, Gunning A, Kelly H, Cibin P, et al. Blood RNA analysis can increase clinical diagnostic rate and resolve variants of uncertain significance. Genet Med. 2020 Jun;22(6):1005–14.

137. Posey JE, Harel T, Liu P, Rosenfeld JA, James RA, Coban Akdemir ZH, et al. Resolution of Disease Phenotypes Resulting from Multilocus Genomic Variation. N Engl J Med. 2017 Jan 5;376(1):21–31.

138. Chérot E, Keren B, Dubourg C, Carré W, Fradin M, Lavillaureix A, et al. Using medical exome sequencing to identify the causes of neurodevelopmental disorders: Experience of 2 clinical units and 216 patients. Clin Genet. 2018 Mar;93(3):567–76.

139. Iossifov I, Ronemus M, Levy D, Wang Z, Hakker I, Rosenbaum J, et al. De novo gene disruptions in children on the autistic spectrum. Neuron. 2012 Apr 26;74(2):285–99.

140. Gordon CT, Chopra M, Oufadem M, Alibeu O, Bras M, Boddaert N, et al. MED13L loss-of-function variants in two patients with syndromic Pierre Robin sequence. Am J Med Genet A. 2018 Jan;176(1):181–6.

141. Baker SW, Murrell JR, Nesbitt AI, Pechter KB, Balciuniene J, Zhao X, et al. Automated Clinical Exome Reanalysis Reveals Novel Diagnoses. J Mol Diagn. 2019 Jan;21(1):38–48.

142. Miller CR, Lee K, Pfau RB, Reshmi SC, Corsmeier DJ, Hashimoto S, et al. Disease-associated mosaic variation in clinical exome sequencing: a two-year pediatric tertiary care experience. Cold Spring Harb Mol Case Stud. 2020 Jun;6(3):a005231.

143. Hamdan FF, Srour M, Capo-Chichi J-M, Daoud H, Nassif C, Patry L, et al. De novo mutations in moderate or severe intellectual disability. PLoS Genet. 2014 Oct;10(10):e1004772.

144. Codina-Solà M, Rodríguez-Santiago B, Homs A, Santoyo J, Rigau M, Aznar-Laín G, et al. Integrated analysis of whole-exome sequencing and transcriptome profiling in males with autism spectrum disorders. Mol Autism. 2015;6:21.

145. Ewans LJ, Schofield D, Shrestha R, Zhu Y, Gayevskiy V, Ying K, et al. Whole-exome sequencing reanalysis at 12 months boosts diagnosis and is cost-effective when applied early in Mendelian disorders. Genet Med. 2018 Dec;20(12):1564–74.

146. Rinaldi B, Race V, Corveleyn A, Van Hoof E, Bauters M, Van Den Bogaert K, et al. Next-generation sequencing in prenatal setting: Some examples of unexpected variant association. Eur J Med Genet. 2020 May;63(5):103875.

147. Liu Y, Liu X, Qin D, Zhao Y, Cao X, Deng X, et al. Clinical Utility of Next-Generation Sequencing for Developmental Disorders in the Rehabilitation Department: Experiences from a Single Chinese Center. J Mol Neurosci. 2021 Apr;71(4):845–53.

148. Bruel A-L, Nambot S, Quéré V, Vitobello A, Thevenon J, Assoum M, et al. Increased diagnostic and new genes identification outcome using research reanalysis of singleton exome sequencing. Eur J Hum Genet. 2019 Oct;27(10):1519–31.

149. Chevarin M, Duffourd Y, A Barnard R, Moutton S, Lecoquierre F, Daoud F, et al. Excess of de novo variants in genes involved in chromatin remodelling in patients with marfanoid habitus and intellectual disability. J Med Genet. 2020 Jul;57(7):466–74.

150. Wang H, Xiao F, Dong X, Lu Y, Cheng G, Wang L, et al. Diagnostic and clinical utility of next-generation sequencing in children born with multiple congenital anomalies in the China neonatal genomes project. Hum Mutat. 2021 Apr;42(4):434–44.

151. Redin C, Gérard B, Lauer J, Herenger Y, Muller J, Quartier A, et al. Efficient strategy for the molecular diagnosis of intellectual disability using targeted high-throughput sequencing. J Med Genet. 2014 Nov;51(11):724–36.

152. Gao C, Wang X, Mei S, Li D, Duan J, Zhang P, et al. Diagnostic Yields of Trio-WES Accompanied by CNVseq for Rare Neurodevelopmental Disorders. Front Genet. 2019;10:485.

153. Kim SH, Kim B, Lee JS, Kim HD, Choi JR, Lee S-T, et al. Proband-Only Clinical Exome Sequencing for Neurodevelopmental Disabilities. Pediatr Neurol. 2019 Oct;99:47–54.

154. Yamamoto T, Shimojima K, Ondo Y, Shimakawa S, Okamoto N. MED13L haploinsufficiency syndrome: A de novo frameshift and recurrent intragenic deletions due to parental mosaicism. Am J Med Genet A. 2017 May;173(5):1264–9.

155. Miao Y-L, Gambini A, Zhang Y, Padilla-Banks E, Jefferson WN, Bernhardt ML, et al. Mediator complex component MED13 regulates zygotic genome activation and is required for postimplantation development in the mouse. Biol Reprod. 2018 Apr 1;98(4):449–64.

156. ACMG Board of Directors. ACMG policy statement: updated recommendations regarding analysis and reporting of secondary findings in clinical genome-scale sequencing. Genet Med. 2015 Jan;17(1):68–9.

157. Ghosh R, Oak N, Plon SE. Evaluation of in silico algorithms for use with ACMG/AMP clinical variant interpretation guidelines. Genome Biol. 2017 Nov 28;18(1):225.

158. Hubisz MJ, Pollard KS, Siepel A. PHAST and RPHAST: phylogenetic analysis with space/time models. Brief Bioinformatics. 2011 Jan;12(1):41–51.

159. Iqbal S, Pérez-Palma E, Jespersen JB, May P, Hoksza D, Heyne HO, et al. Comprehensive characterization of amino acid positions in protein structures reveals molecular effect of missense variants. PNAS. 2020 Nov 10;117(45):28201–11.

160. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014 Mar;46(3):310–5.

161. 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature. 2015 Oct 1;526(7571):68–74.

162. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet. 2013 Jan;Chapter 7:Unit7.20.

163. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009;4(7):1073–81.

164. Feng B-J. PERCH: A Unified Framework for Disease Gene Prioritization. Hum Mutat. 2017 Mar;38(3):243–51.

165. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am J Hum Genet. 2016 Oct 6;99(4):877–85.

166. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016 Aug 17;536(7616):285–91.

167. Silk M, Petrovski S, Ascher DB. MTR-Viewer: identifying regions within genes under purifying selection. Nucleic Acids Research. 2019 Jul 2;47(W1):W121–6.

168. Wiel L, Baakman C, Gilissen D, Veltman JA, Vriend G, Gilissen C. MetaDome: Pathogenicity analysis of genetic variants through aggregation of homologous human protein domains. Human Mutation. 2019;40(8):1030–8.

169. Tørring PM, Larsen MJ, Brasch-Andersen C, Krogh LN, Kibæk M, Laulund L, et al. Is MED13L-related intellectual disability a recognizable syndrome? Eur J Med Genet. 2019 Feb;62(2):129–36.

170. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, et al. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res. 2016 Jan 4;44(D1):D862-868.

171. Smol T, Frénois F, Manouvrier-Hanu S, Petit F, Ghoumid J. Performance of meta-predictors for the classification of MED13L missense variations, implication of raw parameters. Eur J Med Genet. 2021 Nov 16;65(1):104398.

172. Abramovs N, Brass A, Tassabehji M. GeVIR is a continuous gene-level metric that uses variant distribution patterns to prioritize disease candidate genes. Nat Genet. 2020 Jan;52(1):35–9.

173. Tian Y, Pesaran T, Chamberlin A, Fenwick RB, Li S, Gau C-L, et al. REVEL and BayesDel outperform other in silico meta-predictors for clinical variant classification. Sci Rep. 2019 Sep 4;9(1):12752.

174. Hornbeck PV, Zhang B, Murray B, Kornhauser JM, Latham V, Skrzypek E. PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. Nucleic Acids Res. 2015 Jan;43(Database issue):D512-520.

175. Hao B, Oehlmann S, Sowa ME, Harper JW, Pavletich NP. Structure of a Fbw7-Skp1-cyclin E complex: multisite-phosphorylated substrate recognition by SCF ubiquitin ligases. Mol Cell. 2007 Apr 13;26(1):131–43.

176. Poss ZC, Ebmeier CC, Odell AT, Tangpeerachaikul A, Lee T, Pelish HE, et al. Identification of Mediator Kinase Substrates in Human Cells using Cortistatin A and Quantitative Phosphoproteomics. Cell Rep. 2016 Apr 12;15(2):436–50.

177. Alderson TR, Lee JH, Charlier C, Ying J, Bax A. Propensity for cis-Proline Formation in Unfolded Proteins. ChemBioChem. 2018;19(1):37–42.

178. Diella F, Haslam N, Chica C, Budd A, Michael S, Brown NP, et al. Understanding eukaryotic linear motifs and their role in cell signaling and regulation. Front Biosci. 2008 May 1;13:6580–603.

179. Perkins JR, Diboun I, Dessailly BH, Lees JG, Orengo C. Transient Protein-Protein Interactions: Structural, Functional, and Network Properties. Structure. 2010 Oct 13;18(10):1233–43.

180. Aref-Eshghi E, Kerkhof J, Pedro VP, Barat-Houari M, Ruiz-Pallares N, Andrau J-C, et al. Evaluation of DNA Methylation Episignatures for Diagnosis and Phenotype Correlations in 42 Mendelian Neurodevelopmental Disorders. The American Journal of Human Genetics. 2020 Mar 5;106(3):356–70.

181. Kinoshita-Kikuta E, Kinoshita E, Matsuda A, Koike T. Tips on improving the efficiency of electrotransfer of target proteins from Phos-tag SDS-PAGE gel. PROTEOMICS. 2014;14(21–22):2437–42.

182. Mann M, Ong S-E, Grønborg M, Steen H, Jensen ON, Pandey A. Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome. Trends in Biotechnology. 2002 Jun 1;20(6):261–8.

183. Lelieveld SH, Reijnders MRF, Pfundt R, Yntema HG, Kamsteeg E-J, de Vries P, et al. Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. Nat Neurosci. 2016 Sep;19(9):1194–6.

ANNEXES

Tableau 4 ***Variations faux-sens pathogènes et possiblement pathogènes rapportées dans PubMed, ClinVar, DECIPHER et LOVD étudiées dans l'article 2 de la thèse***

Exon	HGVSc	HGVSp	PubMed	ClinVar	DECIPHER	LOVD	MED13	Grantham	verPhyloP	verPhCons
2	c.187T>C	p.Cys63Arg	(107)				1,00	180	5,05	1,00
3	c.337G>C	p.Gly113Arg			307353		1,00	125	5,20	1,00
7	c.995A>G	p.Glu332Gly		VCV000973301			1,00	98	4,59	1,00
10	c.1718C>T	p.Pro573Leu	(112)				0,00	98	4,45	1,00
11	c.2217A>T	p.Lys739Asn			367459		0,50	94	1,06	1,00
14	c.2530C>A	p.Leu844Ile	(118)				0,00	5	1,64	1,00
15	c.2579A>G	p.Asp860Gly	(119)	VCV000427897			1,00	94	4,75	1,00
15	c.2590A>T	p.Met864Leu		VCV000265523		NL_AMC	1,00	15	4,75	1,00
15	c.2591T>A	p.Met864Lys	(121)	VCV000376790			1,00	95	4,75	1,00
15	c.2597C>T	p.Pro866Leu	(62)	VCV000431916	258131		1,00	98	5,73	1,00
15	c.2600C>T	p.Thr867Ile		VCV000488544			1,00	89	5,73	1,00
15	c.2605C>T	p.Pro869Ser	(62)	VCV000978840			1,00	74	5,73	1,00
15	c.2605C>A	p.Pro869Thr		VCV000391487			1,00	38	5,73	1,00
15	c.2606C>T	p.Pro869Leu				NL_AMC	1,00	98	5,73	1,00
15	c.2614G>A	p.Glu872Lys				NL_Groningen	1,00	56	5,73	1,00

15	c.2633C>T	p.Ser878Phe		VCV000421145			0,00	155	5,81	1,00
15	c.2636C>T	p.Pro879Leu		VCV000421338			0,00	98	5,85	1,00
16	c.2930C>T	p.Ala977Val	(62)	VCV000429919			0,50	64	6,17	1,00
17	c.3392G>A	p.Cys1131Tyr	(105,183)		#N/A	NL_Nijmegen	1,00	194	5,55	1,00
17	c.3478T>C	p.Cys1160Arg	(55)		265953		1,00	180	5,00	1,00
17	c.3482C>T	p.Thr1161Ile		VCV000827796	394749	NL_Nijmegen	1,00	89	2,98	1,00
20	c.4358A>C	p.His1453Pro		VCV000978380			1,00	77	4,59	1,00
20	c.4387G>A	p.Gly1463Arg			269224		1,00	125	3,96	1,00
22	c.5033T>A	p.Val1678Asp		VCV000521556			1,00	152	5,18	1,00
22	c.5175G>T	p.Gln1725His		VCV000450904			1,00	24	3,52	1,00
23	c.5282C>T	p.Pro1761Leu	(128)				1,00	98	3,10	1,00
24	c.5505C>G	p.Asp1835Glu			304785		1,00	0	0,17	1,00
25	c.5695G>A	p.Gly1899Arg	(55,117)		262545&323183		1,00	125	5,88	1,00
25	c.5708G>A	p.Arg1903His	(183)				1,00	29	5,88	1,00
26	c.5785A>G	p.Lys1929Glu		VCV000422146			1,00	56	4,49	1,00
26	c.5849T>C	p.Leu1950Pro	(132)				1,00	98	4,49	1,00
27	c.6005C>T	p.Ser2002Leu	(62)		260542		1,00	145	5,73	1,00
28	c.6074T>G	p.Met2025Arg	(55)		262717		0,50	91,00	4,60	1,00
29	c.6280C>A	p.Pro2094Thr	(110)	VCV000520657			1,00	38,00	5,93	1,00

30	c.6485C>T	p.Thr2162Met	(55,135)	VCV000224153	272205	NL_AMC	1,00	81,00	6,00	1,00
30	c.6485C>A	p.Thr2162Lys	(134)				1,00	78,00	6,00	1,00
30	c.6488C>T	p.Ser2163Leu	(62)	VCV000620012			1,00	145,00	5,80	1,00
31	c.6530C>A	p.Ser2177Tyr	(62)				1,00	144,00	5,88	1,00
31	c.6583C>T	p.His2195Tyr		VCV001098385			1,00	83,00	5,88	1,00