

Thèse présentée pour l'obtention du grade de :

**Docteur de l'Université de Strasbourg**

Discipline : Science de la vie et de la santé (ED414)

Spécialité : Aspects moléculaires et cellulaires de la biologie

Par

Salima EL CHEHADEH, née DJEBBAR

**Exploration de familles non-résolues avec déficience  
intellectuelle présumée liée à l'X à l'ère du  
*Whole Genome Sequencing***

Directrice de thèse : Docteur Amélie Piton

Co-directeur de thèse : Docteur Jean Muller

Institut de Génétique et Biologie Moléculaire et Cellulaire, Illkirch-Graffenstaden  
Laboratoire de Génétique Médicale, INSERM UMR-S 1112, Strasbourg

Thèse soutenue publiquement le 30 Novembre 2023 devant le jury :

Directrice de thèse	Dr Amélie Piton, MCU-PH, IGBMC, Illkirch-Graffenstaden
Rapporteur externe	Pr Isabelle Maystadt, Professeure, IPG, Gosselies, Belgique
Rapporteur externe	Dr Frédéric Laumonnier, DR INSERM, UMR_U1253, Tours
Examineur	Dr Valérie Biancalana, MCU-PH, IGBMC, Illkirch-Graffenstaden
Membre invitée	Pr Hélène Dollfus, PU-PH, UMR-S 1112, IGMA, Strasbourg
Membre invité	Pr Jean-Louis Mandel, Professeur émérite, IGBMC, Illkirch-Graffenstaden



*A Paul Eluard,  
Pour ses poèmes sur la paix.*





# Remerciements

---

Cette longue aventure du doctorat a été à l'origine d'un immense enrichissement scientifique et intellectuel (ponctué de quelques remises en question, il faut bien le dire) et n'a pas été parcourue seule, bien au contraire. De nombreuses personnes ont contribué, chacune à sa façon, à l'aboutissement de ce travail qui se matérialise (enfin !) par ce manuscrit. J'aimerais leur rendre hommage par ces quelques mots.

J'aimerais tout d'abord remercier chaleureusement le Professeur Hélène Dollfus, de m'avoir accueillie dans son service il y a bientôt 10 ans (déjà !), et de m'avoir permis, en parallèle de mon activité clinique, de réaliser ce projet périlleux qui me tenait à cœur. Merci de m'avoir fait confiance.

Un grand merci à mes directeurs de thèse, les Docteurs Amélie Piton et Jean Muller, d'avoir accepté de m'encadrer malgré le temps partiel que j'ai pu consacrer à ce travail de thèse. Je vous remercie pour m'avoir donné l'opportunité de travailler avec vous sur un projet difficile, mais extrêmement stimulant, tout en me laissant une grande liberté. Merci de m'avoir appris à interpréter des données de séquençage de génome, et de m'avoir ainsi permis de parfaire mes connaissances en biologie moléculaire et d'améliorer ma pratique clinique en comblant, au moins en partie, cette lacune qui m'encombrait. Amélie, je suis ravie que nous puissions continuer à marcher ensemble sur le chemin de la déficience intellectuelle liée à l'X, malgré ses embûches et ses mystères irrésolus, qui sont autant de trésors à découvrir pour les scientifiques que nous sommes...

Tous mes remerciements aux membres de mon jury, les Docteurs Frédéric Laumonier et Valérie Biancalana ainsi que le Professeur Isabelle Maystadt, pour avoir consacré du temps à la lecture de ce travail et m'avoir fait l'honneur de l'évaluer. Un grand merci au Professeur Jean-Louis Mandel, pour ses commentaires toujours très enrichissants sur ce sujet complexe mais ô combien passionnant qu'est la déficience intellectuelle liée à l'X. Je profite de cette opportunité pour lui témoigner ma profonde considération pour son incommensurable contribution à l'Histoire de cette entité clinique, et au diagnostic des maladies rares en général.

Je remercie l'ensemble des membres des laboratoires de génétique médicale U1112 et de génétique des maladies neurodéveloppementales de l'IGBMC de m'avoir acceptée dans leur équipe. Je pense en particulier à Corinne et Véronique, qui ont bien voulu prendre un peu de leur temps pour m'expliquer avec patience (à moi, une clinicienne...) le maniement des outils de bioinformatique pour que je puisse me lancer dans cette aventure moléculaire. Jean-Baptiste, ton aide a été comment dire, juste indispensable...merci pour ta patience et pour avoir généré autant de scripts et de tableaux Excel que nécessaire pour que je puisse finaliser ce travail. Grâce à toi, j'ai appris des mots de vocabulaire bioinformatique dont je ne soupçonnais pas l'existence et j'ai compris que l'identification et l'analyse des SV en génome sont un réel supplice intellectuel !! Un grand merci à Tarek pour son aide sur les données de génome.

J'adresse toute ma gratitude à l'ensemble du laboratoire de diagnostic génétique, les secrétaires, les techniciens et les biologistes avec qui j'ai la chance de travailler et d'échanger au quotidien. La qualité de votre travail et la fluidité de nos échanges contribuent grandement à rendre optimale la prise en charge des patients. Un merci tout particulier à Sophie, pour m'avoir communiqué des résultats de qPCR en un temps records, afin que je puisse les intégrer *in extremis* dans mon manuscrit ! Je pense aussi à Claire Feger, merci de m'avoir aidée à réaliser les études de ségrégation familiales et d'avoir pris le temps de m'épauler devant la paillasse. Je n'ai pas pu le faire aussi souvent que je l'aurais souhaité, mais c'était un réel plaisir de « manipuler » à tes côtés.

Merci également à Bénédicte Gérard et à Christel Depienne pour leur sagesse et leur enseignement. J'espère vivement que nos chemins se croiseront à nouveau, autour d'un nouveau gène candidat, qui sait ?!

J'adresse ma reconnaissance à tous les membres de l'IGMA et du CARGO, qu'ils soient médecins (sénior ou internes), conseillères en génétique, secrétaires, infirmier(e)s, ARC, assistante sociale ou psychologues (sans oublier nos collègues pédiatres et autres spécialistes), pour leur compréhension et leur soutien pendant cette période durant laquelle j'ai pu avoir tendance à m'isoler un peu, pour tenter tant bien que mal de mener tout de front, au risque de ne pas toujours être aussi disponible que je l'aurais voulu. J'ai l'immense chance de travailler au quotidien parmi vous, au sein d'une équipe de haut niveau mais surtout soudée, bienveillante, et remplie de bonne humeur, ce qui fait sa force. Nous sommes parvenus à rester debout, malgré les tempêtes qui font tourner et s'envoler les secrétaires au beau milieu d'un hôpital lui-même effeuillé et maladif... Continuons sur cette voie, pour les patients !

Je suis clairement redevable à l'association Xtraordinaire et au CREGEMES pour leur soutien financier, ainsi qu'au centre de calcul du laboratoire ICube à Strasbourg et au Centre National de Recherche en Génomique Humaine (Evry) au sein duquel les génomes des patients ont été séquencés.

Je pense également à mes collègues généticiens cliniciens et autres spécialistes des quatre coins de la France – notamment le Professeur Lipsker à Strasbourg – qui m'ont accordé leur confiance en m'adressant l'ADN de leurs patients ainsi que des données cliniques et des photographies, si précieuses pour mon travail d'interprétation. J'adresse une pensée particulière, pleine de respect et d'admiration, au Professeur Lionel Van Maldergem, qui nous a quittés prématurément mais restera à jamais dans mes souvenirs un très grand clinicien et un fin sémiologiste.

Je suis reconnaissante envers mes collègues internationaux pour leur précieuse collaboration, en acceptant de réaliser des études fonctionnelles poussées (Dr J Um, Corée du Sud, Pr B Dermaut, Belgique) ou d'inclure certains de mes résultats dans leur travail de recherche en cours (Dr N Miyake, Japon).

Je ne remercierai jamais assez le Professeur Laurence Faivre, mon Maître devenue amie, qui fut la première à m'ouvrir les portes de la génétique clinique et à me laisser une chance d'entrer dans cette discipline si pointue et singulière, alors même que je savais à peine ce qu'était un chromosome. Laurence, tu m'as communiqué ton énergie et ta passion pour la génétique médicale et la recherche clinique, qui ne m'ont jamais quittée depuis. C'est en te regardant faire qu'est née ma passion pour la dysmorphologie; tu m'as appris à me poser les bonnes questions et à remettre en cause tout ce qui ne me semble pas rationnel, en suivant toujours mon instinct clinique. Je ne serais certainement pas arrivée là où je suis sans toi et je t'en serai éternellement reconnaissante.

Enfin, je remercie bien sûr ma famille, en particulier à Khaled, qui m'a toujours suivie et soutenue depuis 18 ans (cela ne nous rajeunit pas !) dans tous mes projets, avec beaucoup d'enthousiasme (et peut-être un peu d'inconscience aussi...), tout en partageant, bon gré mal gré, mes peines de neurones... Quitter la pédiatrie pour la génétique (pourquoi pas ?), partir faire un master à Paris, enceinte d'Anton et en te laissant une petite Maya âgée de 2 ans sur les bras (un peu rude quand même...), tout lâcher à Dijon pour partir à Strasbourg (l'Alsace, c'est bien aussi!), avoir un 3<sup>ème</sup> enfant (au point où on en est...), puis la thèse durant...5 ans (si j'avais su... !).

Cinq ans, c'est l'âge de Nils. Cette thèse a donc toujours fait partie de son univers familial, il va falloir le préparer à vivre sans... !

Eh bien les enfants, ça y est, elle est bel et bien terminée cette thèse, pour de vrai, si si c'est possible ! Mille mercis pour votre patience, vos encouragements perpétuels et votre vivacité d'esprit (parfois un peu trop débordante, je le conçois) qui m'insufflent chaque jour l'énergie et l'inspiration indispensables pour mener à bien tous mes projets. Je vais ralentir c'est promis...enfin, je vais essayer !

Merci à mes parents et à mes frères pour leur soutien inconditionnel durant toutes ces années, tant sur le plan logistique que spirituel.

Merci aux patients et à leurs familles, qui sont à l'origine de ce travail et au côté desquels je ne cesserai jamais d'apprendre...





# Résumé

---

Salima EL CHEHADEH-DJEBBAR

## Exploration de familles non-résolues avec déficience intellectuelle présumée liée à l'X à l'ère du *Whole Genome Sequencing*

### Résumé

La déficience intellectuelle (DI) concerne 2% de la population et représente la première cause de consultation dans les centres de génétique pédiatriques. Plus de la moitié des causes de DI sont génétiques, avec près de 1500 gènes impliqués, dont 127 gènes situés sur le chromosome X, et 200 syndromes recensés. On estime qu'environ 21% des syndromes avec DI liée à l'X (DILX) restent irrésolus sur le plan moléculaire. Des gènes de DILX et des mécanismes mutationnels encore non décrits sur l'X restent donc probablement à identifier. Ce travail de doctorat a en partie consisté à séquencer le génome (WGS) de patients atteints de DI présumée liée à l'X dont les analyses génétiques antérieures étaient négatives, ce qui a permis de contribuer à la validation de nouveaux gènes candidats de DILX, tel que le gène *SLITRK2*. Il souligne les difficultés rencontrées lors de l'interprétation des données de WGS, en particulier l'identification des variants non-codants et le très grand nombre de variants à analyser. De plus, l'interprétation des variants faux-sens sur des gènes de l'X est délicate en l'absence de ségrégation familiale concluante et d'étude fonctionnelle disponible.

**Mots clés :** Déficience intellectuelle liée à l'X, séquençage de génome entier, gène *SLITRK2*.

### Résumé en anglais

Intellectual disability (ID) affects 2% of the population and is the leading cause of consultation in pediatric genetics centers. More than half the causes of ID are genetic, with almost 1,500 genes involved, including 127 genes on the X chromosome, and 200 syndromes identified. It is estimated that around 21% of X-linked ID (XLID) syndromes remain molecularly unsolved. Therefore, XLID genes and mutational mechanisms that have yet to be described, probably remain to be identified. Part of this PhD work consisted in performing whole genome sequencing (WGS) of patients with presumed XLID whose previous genetic analyses had been negative. This work contributed to the validation of new DILX candidate genes, such as the *SLITRK2* gene. It also highlighted the difficulties encountered when interpreting WGS data, in particular regarding the identification of non-coding variants and the very large number of variants to be analyzed. In addition, the interpretation of missense variants on X-linked genes is tricky in the absence of conclusive familial segregation and available functional studies.

**Key words:** X-linked intellectual deficiency, whole genome sequencing, *SLITRK2* gene

# Table des matières

---

REMERCIEMENTS .....	6
RESUME .....	9
TABLE DES MATIERES .....	10
LISTE DES ABREVIATIONS .....	13
LISTE DES FIGURES .....	16
LISTE DES TABLEAUX .....	19
LISTE DES ANNEXES .....	20
INTRODUCTION.....	22
<b>I. LA DEFICIENCE INTELLECTUELLE (DI).....</b>	<b>22</b>
1. Introduction .....	22
1. Terminologie et prévalence.....	22
2. Les critères de diagnostic de DI.....	22
3. Les principaux outils utilisés pour l'évaluation de la DI.....	23
a) Le déficit des fonctions intellectuelles .....	23
b) Le déficit du fonctionnement adaptatif .....	24
c) Les précautions à prendre dans l'interprétation de ces scores.....	24
4. Les critères de sévérité de la DI .....	25
2. Les étiologies.....	26
1. Les causes prénatales.....	26
2. Les causes périnatales.....	26
3. Les causes postnatales .....	26
3. Le diagnostic de la DI.....	27
1. Les types d'anomalies génétiques impliquées dans la DI.....	27
a) Les anomalies de nombre et de structure des chromosomes .....	28
b) Les micro-réarrangements génomiques déséquilibrés.....	29
c) Les formes monogéniques .....	31
d) Les pathologies liées à l'empreinte.....	32
2. Les modes de transmission .....	33
a) Le mode autosomique dominant.....	33
b) Le mode autosomique récessif .....	33
c) Le mode de transmission lié à l'X .....	34
3. Le diagnostic de la DI : quels besoins? Quels enjeux? .....	35
a) En clinique, pour les familles.....	35
b) En recherche, pour la Science .....	37
c) Au niveau national .....	37
<b>II. LA DI LIEE AU CHROMOSOME X (DILX) .....</b>	<b>38</b>
1. Historique.....	38
2. Le diagnostic de la DILX : que sait-on?.....	40
1. Un mode de transmission particulier .....	40
a) Un conseil génétique pas toujours aisé .....	40
b) Le phénomène d'inactivation de l'X.....	41
2. Les causes monogéniques de DILX .....	44
a) Classification des gènes de DILX .....	45
b) Les protéines d'adhésion synaptique : exemple de la famille SLITRK.....	50

3. Les CNV récurrents sur le chromosome X.....	53
3. Le diagnostic de la DILX : que reste-t-il à découvrir?-----	54
4. Les défis de la DILX -----	55
<b>III. LE SEQUENÇAGE HAUT DEBIT DANS LE DIAGNOSTIC DE LA DI .....</b>	<b>56</b>
1. Historique de l'évolution des techniques -----	56
2. Le principe des technologies de séquençage haut débit (NGS) -----	58
a) L'analyse primaire et secondaire .....	59
b) L'analyse tertiaire : l'annotation, filtre et classification des variants.....	61
3. Les aides à l'interprétation des variants-----	63
1. Les bases de données de variants et d'informations sur les gènes .....	63
2. Les différents outils de prédiction <i>in silico</i> .....	64
a) Les outils de prédiction des effets des variants faux-sens .....	64
b) Les outils de prédiction sur l'épissage .....	66
c) Les outils de prédiction des variants de structure.....	69
3. La classification ACMG.....	69
4. Les études de ségrégation familiale.....	71
5. Les analyses fonctionnelles .....	71
a) Évaluation de l'effet d'un variant sur les transcrits :.....	72
b) Évaluation de l'effet fonctionnel d'un variant sur la protéine .....	73
6. Les réunions de concertation pluridisciplinaires clinico-biologiques.....	75
7. Le partage de données collaboratif national et international .....	75
4. Les techniques de NGS ont fait leurs preuves dans le diagnostic de la DI -----	76
1. L'utilisation de panels ciblés et du séquençage d'exome (ES) .....	76
2. Le séquençage de génome complet (WGS).....	79
a) Le génome humain : gènes, expression, et régions régulatrices.....	79
b) Intérêts et avantages du WGS .....	82
c) Limites du WGS .....	85
d) Le WGS en diagnostic : le Plan France Médecine génomique 2025 .....	87
<b>OBJECTIFS DE LA THESE .....</b>	<b>88</b>
<b>RESULTATS .....</b>	<b>92</b>
<b>I. PREMIERE PARTIE : UNE FORME RARE DE DILX, LE SYNDROME DE DUPLICATION DU GENE <i>MECP2</i> .....</b>	<b>92</b>
1. Syndrome de duplication du gène <i>MECP2</i> chez les filles : que pouvons-nous apprendre pour le diagnostic et le conseil génétique ? -----	93
2. Caractérisation des anomalies à l'IRM cérébrale chez 30 patients atteints de <i>MECP2DS</i> -----	106
3. La cohorte française de 59 patients atteints de <i>MECP2DS</i> : description phénotypique fine avec un focus sur les caractéristiques morphologiques et neurologiques-----	121
<b>II. DEUXIEME PARTIE : LA RECHERCHE DE NOUVEAUX GENES CANDIDATS DE DILX : LES VARIATIONS DU GENE <i>SLITRK2</i> DANS LES TROUBLES DU NEURODEVELOPPEMENT .....</b>	<b>135</b>
<b>III. TROISIEME PARTIE : POURSUITE DES EXPLORATIONS PAR WGS CHEZ DES PATIENTS AVEC DI SUPPOSEE LIEE A L'X NON-RESOLUE PAR LE SEQUENÇAGE CIBLE OU L'ES. ....</b>	<b>155</b>
1. Cohorte des 15 patients avec DI présumée liée à l'X-----	155
1. <i>Matériel et méthodes</i> .....	155
a) Financement, recrutement des patients, circuits des ADN .....	155
b) Analyses bioinformatiques .....	157
2. <i>Résultats</i> .....	159
a) Analyse des variants situés dans les régions codantes .....	159
b) Analyse des variants pouvant avoir un effet sur l'épissage .....	172
c) Analyse des variants structuraux .....	174

2. Patients suivis dans le cadre de ma pratique clinique et ayant bénéficié d'un WGS en <i>trio</i> -----	180
<b>IV. QUATRIEME PARTIE .....</b>	<b>181</b>
1. L'interprétation de séquençage d'exome de patients atteints de DI suivis dans le service de génétique médicale des HUS. -----	181
2. L'interprétation d'ES de patients atteints de maladies dermatologiques très rares -----	182
3. L'interprétation de WGS de patients présentant l'association de mélanomes multiples et lipomatose mésosomatique -----	182
4. La description phénotypique de syndromes rares avec DI -----	182
<b>DISCUSSION .....</b>	<b>184</b>
<b>CONCLUSIONS .....</b>	<b>196</b>
<b>REFERENCES BIBLIOGRAPHIQUES .....</b>	<b>198</b>
<b>ANNEXES .....</b>	<b>228</b>
Annexe 1 : Niveaux de sévérité du trouble du développement intellectuel – DSM-5 -----	228
Annexe 2 : Les variations du gène <i>SLITRK2</i> dans les troubles du neurodéveloppement ( <i>Supplemental data</i> ) -----	231
Annexe 3 : Élargissement du spectre phénotypique associé aux variants pathogènes du gène <i>ARCN1</i> -----	263
Annexe 4 : Syndrome d'Ehlers-Danlos parodontal : description de 13 nouveaux cas et élargissement du phénotype -----	266
Annexe 5 : Description phénotypique du syndrome de Frank-ter Haar -----	273
Annexe 6 : La perte de fonction de <i>PLAAT3</i> provoque un syndrome lipodystrophique et neurologique par altération de la signalisation <i>PPAR<math>\gamma</math></i> -----	284

# Liste des abréviations

---

AA: acides aminés  
AAIDD: American Association on Intellectual and Developmental Disabilities  
AAMR: American Association of Mental Retardation  
ABCR: apparently balanced chromosomal rearrangements  
ACMG: American College of Medical Genetic and Genomics  
ACPA: analyse chromosomique sur puce à ADN  
AD: autosomique dominant  
ADN: acide désoxyribonucléique  
AGVGD: Align-GVGD  
APA: American Psychiatric Association  
ARN: acide ribonucléique  
BAM: Binary Alignment Map  
BDNF: brain-derived neurotrophic factor  
BLR: échelle de Brunet-Lézine révisée  
BSID-IV: échelle de Bayley  
CADD: Combined Annotation Dependent Depletion  
CCMR: centre de compétence maladies rares  
CDG: Congenital Disorders of glycosylation  
CGH-array: Array comparative genomic hybridization, puce d'hybridation génomique comparative  
CHU: Centre hospitalo-universitaire  
CIC: centre d'investigation clinique  
CIM: classification mondiale des maladies  
CNG: Centre national de génotypage  
CNV: copy number variation  
CPDPN: Comité Pluridisciplinaire de Diagnostic Prénatal  
CRM: centre de référence maladies rares  
CRNGH: Centre National de Recherche en Génomique Humaine  
DEFIDIAG: DEFicience Intellectuelle DIAGnostic  
DGV: Database of Genomic Variant  
DI: déficience intellectuelle  
DILX: déficience intellectuelle liée au chromosome X  
DPI : diagnostic pré-implantatoire  
DPN : diagnostic prénatal  
DPS: densité post-synaptique  
DS: déviation standard  
DSM: Manuel diagnostique et statistique des troubles mentaux  
EM: éléments mobiles  
ERN: European Reference Network  
ES : séquençage d'exome  
ESE: exonic splice enhancer  
ESS: exonic splice silencer

FISH: hybridation *in situ* par fluorescence  
FT: facteurs de transcription  
GABA: acide  $\gamma$ -aminobutyrique  
GATK: Genome Analysis ToolKit  
gnomAD: Genome Aggregation Database  
GRCh37: Genome Reference Consortium Human Build 37  
hnRNP: heterogenous nuclear ribonucleoproteins  
HUS: hôpitaux universitaires de Strasbourg  
IGV: Integrative Genome Viewer  
IME: institut médico-éducatif  
*Indel*: petites insertions ou délétions nucléotidiques  
IPSC: Induced Pluripotent Stem Cell  
IRM: imagerie par résonnance magnétique  
ISE: intronic splice enhancer  
ISS: intronic splice silencer  
Kb: kilobases  
KO: knock-out  
Mb: mégabases  
MECP2DS: syndrome de duplication du gène *MECP2*  
MLPA: amplification multiplex par sonde de ligature  
MM: maladies mendéliennes  
NGS: next generation sequencing  
NIH: National Institutes of Health  
NMD: Nonsense Mediated mRNA Decay  
OMIM: Online Mendelian Inheritance in Man  
OMS: Organisation Mondiale de la Santé  
OR: odds ratio  
PARs: régions pseudo-autosomiques  
PC: périmètre crânien  
PCN: périmètre crânien de naissance  
PCR: polymerase Chain Reaction  
PHRC: programme hospitalier de recherche clinique  
PIEV: Pénétrance Incomplète et/ou Expressivité Variable  
pLI: probability of being loss-of-function (LoF) intolerant  
PN: poids de naissance  
PNDS: Protocole National de Diagnostic et de Soins  
PNMR: Plan National Maladies Rares  
QD: quotient de développement  
QI: quotient intellectuel  
QIT: quotient intellectuel total  
qPCR: PCR (polymerase Chain Reaction) quantitative  
RCIU: retard de croissance intra-utérin  
RPTP: récepteurs tyrosine phosphatase  
RT-PCR: reverse transcriptase-PCR (polymerase Chain Reaction)  
SEDp: syndrome d'Ehlers Danlos parodontal

SEGPA: section d'enseignement général et professionnel adapté  
SIFT: Sorting Intolerant From Tolerant  
SNP array: single nucleotide polymorphism – array, puce d'hybridation génomique comparative  
SNP: single nucleotide polymorphism  
SNV: Single Nucleotide Variant  
SR: protéines riches en sérine  
SV: variants de structure  
TAD: domaines topologiquement associés  
TDAH: trouble déficit de l'attention avec ou sans hyperactivité  
TDC: troubles développementaux de la coordination  
TDI: trouble du développement intellectuel  
TN: taille de naissance  
TND: trouble du neurodéveloppement  
TR: répétitions en tandem  
TSA: troubles du spectre de l'autisme  
UTR: untranslated regions  
VCF: Variant Call Format  
WAIS IV: Wechsler Adult Intelligence Scale  
WGS: whole genome sequencing  
WISC V: Wechsler Intelligence Scale for Children  
WPPSI IV: Wechsler Preschool and Primary Scale of Intelligence  
XCI : phénomène d'inactivation du chromosome X

# Liste des Figures

---

Figure 1: Distribution du QI théorique dans la population générale.....	23
Figure 2: Les causes génétiques des TND.....	27
Figure 3: Caryotypes montrant une anomalie chromosomique.....	28
Figure 4: Anomalies congénitales fréquemment observées dans la trisomie 18.....	28
Figure 5: Technique de CGH-array.....	29
Figure 6: Délétion 7q11.23 responsable du syndrome de Williams-Beuren.....	30
Figure 7: Patientes atteintes du syndrome de Wolf-Hirschhorn.....	30
Figure 8: Annotation de la fonction des gènes associés à la DI basée sur l'ontologie.....	31
Figure 9: Réseau d'interaction des gènes associés à la DI.....	32
Figure 10: Proportion des variants pathogènes <i>de novo</i> dans la DI sévère.....	33
Figure 11: Famille princeps décrite en 1943 par Martin et Bell.....	38
Figure 12: Marqueur du syndrome de l'X fragile.....	38
Figure 13: Famille princeps décrite par Renpenning en 1962.....	39
Figure 14: Le corpuscule de Barr.....	39
Figure 15: Famille décrite par Nielsen en 1981.....	40
Figure 16: Mécanismes de compensation de dosage génique.....	41
Figure 17: Ratios d'inactivation de l'X chez 1005 femmes.....	42
Figure 18: Inactivation aléatoire du chromosome X dans un embryon humain.....	42
Figure 19: Gènes échappant au mécanisme d'inactivation de l'X.....	43
Figure 20: Différents niveaux d'échappement à l'XCI.....	43
Figure 21: Gènes associés à des signes neurologiques sur le chromosome X.....	45
Figure 22: Schémas d'une synapse.....	47
Figure 23: Répartition des protéines de la densité post-synaptique selon leur fonction.....	49
Figure 24: Principales protéines d'adhésion et d'architecture synaptique.....	50
Figure 25: Interactions des protéines SLITRK avec les protéines de la famille RPTP.....	51
Figure 26: Représentation schématique du locus Xq27.3.....	52
Figure 27: Expression cérébrale de la protéine SLITRK2 chez l'Homme (GTEx Portal).....	52
Figure 28: Interactions de la protéine SLITRK2 avec PTP $\delta$ et PTP $\sigma$ .....	52
Figure 29: Nouveaux gènes de maladies mendéliennes identifiés par an.....	54
Figure 30: Principes du séquençage selon la méthode de Sanger.....	56
Figure 31: Description de gènes de maladies mendéliennes au fil du temps.....	57
Figure 32: « Plan de métro » de la technologie de séquençage haut débit.....	57
Figure 33: Workflow pour le séquençage d'exome.....	58
Figure 34: Principales étapes de l'analyse bioinformatique des données de NGS.....	59



Figure 35: Alignement des <i>reads</i> (fichier BAM) par le logiciel IGV .....	60
Figure 36: Schématisation des informations contenues dans un fichier VCF (colonnes).....	61
Figure 37: Schématisation du nombre de variants restants après application de filtres .....	62
Figure 38: Nombre de SNP identifiés par WGS chez un individu.....	62
Figure 39: Alignement des acides aminés de la protéine MLH1.....	65
Figure 40: Organisation d'une unité d'épissage.....	67
Figure 41: Schéma des éléments de régulation de l'épissage (A) et des effets de l'altération de l'épissage sur l'ARNm (B).....	68
Figure 42: Localisation cellulaire en microscopie confocale.....	74
Figure 43: Etude d'interactions par immuno-précipitation .....	75
Figure 44: Diagnostics moléculaires associés aux phénotypes (étude DDD).....	76
Figure 45: Identification de gènes de MM par NGS <i>versus</i> approches conventionnelles .....	77
Figure 46: Identification de gènes impliqués dans des formes monogéniques de DI .....	78
Figure 47: Projet Génome humain et projets des espèces modèles .....	80
Figure 48: Répartition des variants dans un seul génome.....	81
Figure 49: Liste des identificateurs d'épigénome .....	81
Figure 50: Méthylation de l'ADN selon le niveau d'expression des gènes .....	82
Figure 51: Rendement diagnostique chez des patients avec DI selon la technique.....	83
Figure 52: Identification en WGS d'un CNV (délétion) impliquant le gène <i>SHANK3</i> .....	84
Figure 53: Délétion au sein de l'exon 4 du gène <i>MECP2</i> détectée en WGS .....	85
Figure 54: Le coût du WGS, de 2001 à 2021 .....	86
Figure 55: Familles recrutées dans le présent travail .....	156
Figure 56: Annotation des variants issus du séquençage de génome.....	158
Figure 57: Nombre de variants rares par patient sur l'X après application de filtres.....	160
Figure 58: Stratégie d'interprétation des variants situés sur le chromosome X .....	160
Figure 59: Distribution des variants rares de l'X selon leur localisation .....	161
Figure 60: Répartition des variants de l'X selon leur type .....	161
Figure 61: Représentation des variants codants selon l'effet protéique.....	162
Figure 62: Arbre généalogique du patient P5. ....	163
Figure 63: Variants hémizygotés potentiellement pathogènes rapportés dans <i>BCORL1</i> .....	164
Figure 64: Arbre généalogique du patient P6. ....	165
Figure 65: vue des variants du gène <i>MED14</i> répertoriés dans ClinVar.....	166
Figure 66: Complexe Mediator.....	167
Figure 67: Interactions de la protéine MED14 avec les autres protéines de la famille MED .....	168
Figure 68: Arbre généalogique du patient PH.....	169
Figure 69: Schéma de la protéine ARHGEF9 et des variants pathogènes publiés. ....	171
Figure 70: Prédictions d'effet sur l'épissage des variants des gènes <i>COL4A5</i> et <i>ARL13A</i> (logiciel Alamut Visual Plus).....	173

Figure 71: Diagramme de Venn illustrant le nombre de CNV identifiés par les outils Manta, Smoove et CNVpytor chez le patient P7 .....	174
Figure 72: Visualisation sur IGV de la duplication Xp11.22 identifiée chez PF .....	175
Figure 73: Arbre de la famille du patient PF .....	176
Figure 74: Duplication Xp11.22 identifiée chez PF et les autres patients de la littérature ...	176
Figure 75: Phénotype des patients porteurs d'une duplication Xp11.22 .....	177
Figure 76: Mise en évidence sur IGV des paires de <i>reads</i> aux points de cassures .....	177
Figure 77: Ratios obtenus en PCR quantitative pour PF, son frère et son cousin maternel atteints concernant la région Xp11.22 .....	178
Figure 78: Vue IGV de la délétion Xq28 de 1,2 Mb détectée chez le patient PD.....	179

# Liste des Tableaux

---

Tableau 1: Classification des niveaux de gravité de la DI selon la CIM-10.....	25
Tableau 2: Critères de gravité de la DI d'après le DSM-5 .....	25
Tableau 3: Principales malformations observées chez les individus atteints de trisomie 18..	28
Tableau 4: Signes cliniques observés dans le syndrome de Wolf-Hirschhorn.....	30
Tableau 5: Gènes de DILX impliqués dans la régulation de la transcription.....	46
Tableau 6 : Gènes de DILX impliqués dans le fonctionnement présynaptique .....	48
Tableau 7: Gènes de DILX impliqués dans l'organisation de la densité post-synaptique.....	50
Tableau 8: Les principaux outils de prédiction in silico des variants faux sens .....	66
Tableau 9: Classification ACMG de pathogénicité des variants.....	71
Tableau 10: Signes cliniques et résultats des examens complémentaires réalisés chez P5..	163
Tableau 11: Signes cliniques et résultats des examens complémentaires réalisés chez P6..	165
Tableau 12: Signes cliniques présentés par le patient PH et ses oncles maternels.....	169
Tableau 13: Récapitulatif des variants pathogènes du gène ARHGEF9 publiés à ce jour .....	171
Tableau 14: scores de prédictions des deux variants identifiés dans les gènes COL4A5 et ARL13A .....	172

# Liste des Annexes

---

Annexe 1 : Niveaux de sévérité du trouble du développement intellectuel – DSM-5 .....	228
Annexe 2 : Les variations du gène <i>SLITRK2</i> dans les troubles du neurodéveloppement ( <i>Supplemental data</i> ).....	231
Annexe 3 : Élargissement du spectre phénotypique associé aux variants pathogènes du gène <i>ARCN1</i> .....	263
Annexe 4 : Syndrome d'Ehlers-Danlos parodontal : description de 13 nouveaux cas et élargissement du phénotype .....	266
Annexe 5 : Description phénotypique du syndrome de Frank-ter Haar.....	273
Annexe 6 : La perte de fonction de <i>PLAAT3</i> provoque un syndrome lipodystrophique et neurologique par altération de la signalisation <i>PPAR<math>\gamma</math></i> .....	284



# Introduction

---

## I. La déficience intellectuelle (DI)

### 1. Introduction

#### 1. Terminologie et prévalence

La déficience intellectuelle (DI) est l'un des principaux troubles du neurodéveloppement (TND) et concerne environ 1 à 2 % de la population (Roeleveld *et al.*, 1997), soit plus d'un million de personnes en France, représentant la première indication de consultation dans les centres de génétique pédiatriques. Elle se définit comme étant caractérisée par une altération du fonctionnement intellectuel et du comportement adaptatif en ce qui concerne les compétences conceptuelles, sociales et pratiques, et survenant avant l'âge de 18 ans (Schalock *et al.*, 2007). Cependant, le terme de « déficience intellectuelle » retranscrivant mal les notions de problèmes adaptatifs et de handicap ou désavantage, notamment l'influence de l'environnement, l'Organisation mondiale de la santé (OMS) décida d'introduire en 2022 le terme de « trouble du développement intellectuel, TDI » (Disorder of Intellectual Development) dans la CIM-11 (classification mondiale des maladies). Nous avons cependant choisi de conserver le terme usuel de déficience intellectuelle (DI) dans ce manuscrit pour plus de clarté.

La DI peut être isolée mais elle s'intègre également très fréquemment à d'autres TND, qui peuvent être multiples et intriqués et comprennent des troubles neurosensoriels (audition, vision), les troubles du spectre de l'autisme (TSA), l'épilepsie, les troubles de développement de la coordination (TDC), les troubles de l'humeur, le trouble déficit de l'attention avec ou sans hyperactivité (TDAH), les troubles du sommeil etc. La prévalence de la DI est plus élevée chez les garçons avec un *sex ratio* évalué à 1,2 à 1,9 selon les études (Buntinx *et al.*, s. d.).

#### 2. Les critères de diagnostic de DI

Les critères de diagnostic positif de DI ont été définis et sont régulièrement réévalués par trois instances internationales, *via* leurs guides respectifs :

- L'Organisation Mondiale de la Santé (OMS): CIM-11 (Classification Internationale des Maladies), Publication de la CIM-11 2022 ([who.int](https://www.who.int)).
- L'American Association on Intellectual and Developmental Disabilities (AAIDD), <https://www.aaid.org/>.
- L'American Psychiatric Association (APA): DSM-5 (Diagnostic and Statistical Manual of Mental Disorders), <https://www.psychiatry.org/psychiatrists/practice/dsm>.

Ces instances ont déterminé les trois critères diagnostiques essentiels suivants :

- **Le déficit des fonctions intellectuelles**, telles que le raisonnement, la résolution de problèmes, la planification, la pensée abstraite, le jugement, l'apprentissage. Ce déficit est confirmé par une évaluation clinique, et mesuré au moyen de tests d'intelligence (psychométriques) standardisés et adaptés à la culture, permettant d'objectiver une différence significative des capacités cognitives d'un individu par rapport à ce qui est attendu pour son âge réel : QI (quotient intellectuel)  $\leq - 2$  écart-types, soit 70 +/- 5.

- **Le déficit du fonctionnement adaptatif**, d'au moins deux écarts-types sous la moyenne, dans au moins une des trois dimensions (habiletés conceptuelles, sociales et pratiques), évaluées à l'aide d'un questionnaire standardisé du fonctionnement adaptatif.

- **Ces déficits doivent apparaître pendant la période développementale**, c'est-à-dire durant l'enfance ou l'adolescence.

### 3. Les principaux outils utilisés pour l'évaluation de la DI

#### a) Le déficit des fonctions intellectuelles

Les batteries les plus utilisées en France pour évaluer le QI sont les trois échelles de Wechsler: **WPPSI IV** (Wechsler Preschool and Primary Scale of Intelligence) (2 ans 6 mois à 7 ans 7 mois), **WISC V** (Wechsler Intelligence Scale for Children) (6 ans à 16 ans 11 mois) et **WAIS IV** (Wechsler Adult Intelligence Scale) (> à 16 ans). Ces batteries sont composées de différents tests regroupés en indices évaluant les diverses fonctions cognitives telles que le raisonnement, la résolution de problèmes, la pensée abstraite, la faculté de jugement, l'apprentissage théorique et l'apprentissage par l'expérience. Les tests de QI comprennent deux volets : l'un est verbal et l'autre porte sur les performances, les résultats de l'une et l'autre partie pouvant diverger chez un même individu.

Les scores de QI suivent une loi de distribution normale dans une population donnée, dont la moyenne est de 100 et la déviation standard (DS) est de 15 (Figure 1).

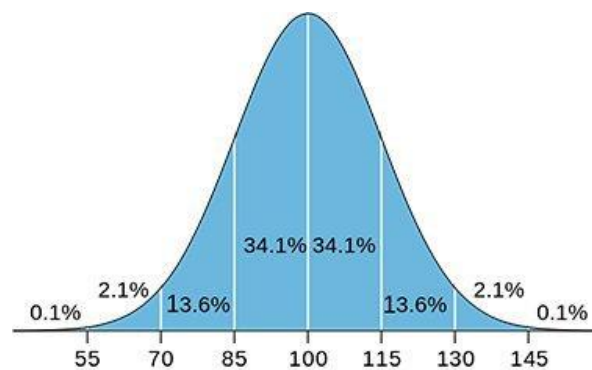


Figure 1: Distribution du QI théorique dans la population générale.

Avant l'âge de 3 à 5 ans, on parle de quotient de développement (QD) plutôt que de QI, et de retard de développement plutôt que de DI, quand le QD est diminué, car le retard de développement peut se normaliser avec l'âge et n'évolue pas systématiquement vers une DI. Le QD peut se mesurer par le biais **d'échelles développementales** telles que l'échelle de **Brunet-Lézine révisée (BLR)** (2 à 30 mois) ou l'**échelle de Bayley (BSID-IV)** (16 jours à 42 mois) qui peuvent être utilisées chez des individus plus âgés en cas de DI profonde. L'échelle BLR évalue le développement précoce en évaluant 4 domaines : posture, coordination, langage et sociabilité. L'échelle BSID-IV évalue 5 compétences : cognitive, langagière (réceptive et expressive), motrice (fine et globale), socio-émotionnelle et adaptative.

### *b) Le déficit du fonctionnement adaptatif*

La seule échelle d'évaluation du comportement adaptatif normée actuellement disponible en France est la Vineland II. Elle évalue le comportement adaptatif à tous les âges de la vie (1 à 90 ans) selon quatre domaines (communication, vie quotidienne, socialisation, motricité) et permet de définir des scores ainsi que les équivalents en âge de développement.

### *c) Les précautions à prendre dans l'interprétation de ces scores*

L'interprétation des tests d'évaluation de QI nécessite cependant de tenir compte de plusieurs paramètres (Des Portes et Héron, 2020):

- Plus l'enfant est jeune, plus la trajectoire développementale est instable ; ainsi, avant l'âge de 5-6 ans, le diagnostic de DI doit être prudent, en dehors des formes sévères ;
- Les scores obtenus ne sont pas toujours le reflet des compétences intellectuelles, dont l'expression peut être entravée par les conditions de passation (temps limité, barrière linguistique, anxiété de performance) ou les troubles associés (audition, vision, motricité, praxies, langage, cognition sociale, attention, régulation émotionnelle, etc.) ;
- Les échelles de Wechsler ne sont pas adaptées aux patients les plus déficients.
- De fortes disparités entre les compétences préservées et déficitaires empêchent le calcul d'un QI total. Ainsi, un profil très hétérogène avec certains subtests dans la norme, doit faire évoquer un trouble spécifique d'apprentissage (dysphasie, dyspraxie, déficit d'attention...) et/ou un trouble psychopathologique plutôt qu'un trouble du raisonnement et justifie des épreuves supplémentaires ciblant les compétences langagières, visuo-spatiales, mnésiques, attentionnelles et exécutives.
- Le QI ne reflète pas le fonctionnement de l'individu dans l'environnement ni l'influence de ce dernier, qu'il soit obstacle ou facilitateur.



#### 4. Les critères de sévérité de la DI

Alors que la CIM-10 définissait des degrés de gravité de la DI sur la base du niveau intellectuel en termes de QI (Tableau 1), le DSM-5, en 2013, s'est basé sur le comportement adaptatif conceptuel, social et pratique. Les niveaux de gravité de la DI restent cependant qualifiés de léger, modéré, grave ou sévère, profond (Tableau 2 et annexe 1).

**Tableau 1: Classification des niveaux de gravité de la DI selon la CIM-10**

Niveau de gravité de la DI	QI (âge mental)	Ecart-types
Léger	50 à 69 (9-12 ans)	Entre -2 et -3
Moyen	35 à 49 (6-9 ans)	Entre -3 et -4
Grave	20 à 34 (3-6 ans)	Entre -4 et -5
Profond	Inférieur à 20 (- 3 ans)	Inférieur à -5

En effet, les scores de QI donnent une valeur approximative du fonctionnement intellectuel mais peuvent s'avérer insuffisants pour évaluer les capacités de raisonnement dans les situations de maîtrise des tâches pratiques (Schalock, 2011). Par exemple, une personne avec un QI au-dessus de 70 peut avoir des problèmes de comportement adaptatif si sévères au niveau de la compréhension sociale, que son fonctionnement réel est comparable à celui de sujets ayant un QI inférieur (Harris, 2010). Plus concrètement, les besoins de soutien d'une personne avec DI sévère qui nécessite une aide permanente, y compris pour son alimentation et son hygiène, sont très différents de ceux d'une personne avec une DI légère, qui a besoin d'une aide humaine pour compter sa monnaie ou se repérer dans les transports en commun.

**Tableau 2: Critères de gravité de la DI d'après le DSM-5**

Gravité	Domaine conceptuel	Domaine social	Domaine pratique
<b>Léger</b>	La personne a une manière plus pragmatique de résoudre des problèmes et de trouver des solutions que ses pairs du même âge...	La personne a une compréhension limitée du risque dans les situations sociales ; a un jugement social immature pour son âge...	La personne occupe souvent un emploi exigeant moins d'habiletés conceptuelles...
<b>Modéré</b>	D'ordinaire, la personne a des compétences académiques de niveau primaire et une intervention est requise pour toute utilisation de ces compétences dans la vie professionnelle et personnelle...	Les amitiés avec les pairs tout-venant souffrent souvent des limitations vécues par la personne au chapitre des communications et des habiletés sociales...	Présence, chez une minorité importante, de comportements mésadaptés à l'origine de problèmes de fonctionnement social...
<b>Grave</b>	La personne a généralement une compréhension limitée du langage écrit ou de concepts faisant appel aux nombres, quantités, au temps et à l'argent...	Le langage parlé est relativement limité sur le plan du vocabulaire et de la grammaire...	La personne a besoin d'aide pour toutes les activités de la vie quotidienne, y compris pour prendre ses repas, s'habiller, se laver et utiliser les toilettes...
<b>Profond</b>	La personne peut utiliser quelques objets dans un but précis (prendre soin de soi, se divertir). Des	La personne peut comprendre des instructions et des gestes simples...	La personne dépend des autres pour tous les aspects de ses soins physiques quotidiens, pour sa santé

Gravité	Domaine conceptuel	Domaine social	Domaine pratique
	problèmes de contrôle de la motricité empêchent souvent un usage fonctionnel...		et pour sa sécurité, quoi qu'elle puisse participer à certaines de ces activités...

On estime la prévalence de la DI sévère à 0,3 à 0,4 % (Mlika *et al.*, 1993; Buntinx *et al.*, s. d.) contre 0,8 à 2,5 % pour la DI légère (David *et al.*, 2014; Des Portes et Héron, 2020). Ainsi, la grande majorité des individus atteints de DI ont une DI légère (85%), 10% une DI modérée, 3-4% une DI sévère et 1 à 2% ont une DI profonde (King *et al.*, 2009). Le contexte socioéconomique joue un rôle certain sur la prévalence de la DI légère, avec une prévalence plus faible de cette dernière dans les milieux socio-économiquement favorisés, ce qui est beaucoup moins observé pour la prévalence de la DI sévère (des Portes, 2020).

## 2. Les étiologies

### 1. Les causes prénatales

Elles comprennent les causes génétiques (variations de séquence ou du nombre de copies, anomalies chromosomiques), les malformations cérébrales, les pathologies maternelles (Huang *et al.*, 2016) et les causes environnementales (alcool, toxiques, agents tératogènes), ces dernières pouvant, lorsque le cerveau en développement y est exposé, conduire à une DI (Majnemer et Shevell, 1995; Niccols, 2007).

### 2. Les causes périnatales

Elles comprennent la très grande prématurité, les incidents survenus pendant le travail et/ou la délivrance à l'origine d'une anoxo-ischémie, l'hypoglycémie néonatale et les infections materno-fœtales (Seidman *et al.*, 2000; Rudolph *et al.*, 2011). Cependant, dans les pays industrialisés où le suivi des grossesses et la réanimation néonatale ont fait d'énormes progrès ces dernières années, la pathologie vasculaire périnatale est rarement responsable d'une déficience intellectuelle isolée : l'absence de signe moteur associé et une imagerie cérébrale normale incitent à remettre en cause une telle hypothèse. En outre, une anoxie aiguë survenant en dehors de tout facteur obstétrical incite à rechercher une pathologie constitutionnelle du fœtus (Des Portes et Héron, 2020).

### 3. Les causes postnatales

Elles comprennent les conséquences d'une hypoxie ischémique (noyade), d'un traumatisme crânien, d'infections du système nerveux central (encéphalite, méningite), d'une démyélinisation, d'une épilepsie sévère et réfractaire au traitement, de syndromes métaboliques toxiques et d'intoxications (plomb, mercure) (Liu *et al.*, 2010).

### 3. Le diagnostic de la DI

On considère que plus de la moitié des causes de DI sont génétiques (Des Portes et Héron, 2020). Parmi ces dernières, on distingue principalement les anomalies de nombre et de structure des chromosomes, visibles sur un caryotype, les micro-réarrangements génomiques déséquilibrés (*copy number variation* – CNV) identifiés par la technique d'ACPA (analyse chromosomique sur puce à ADN), et les variations de séquence génomique identifiées par une technique de séquençage (Figure 2).

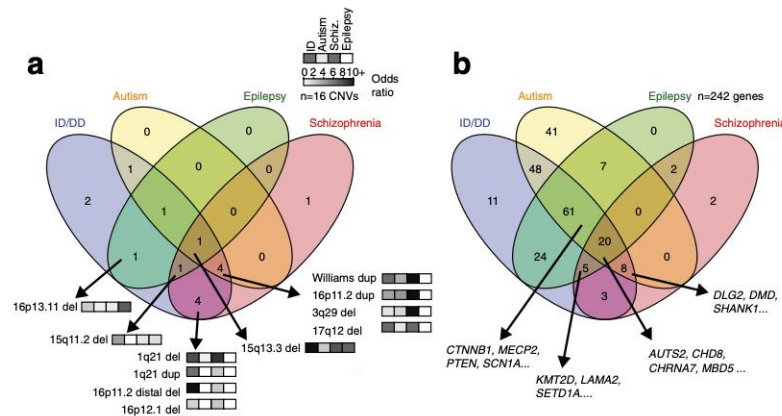


Figure 2: Les causes génétiques des TND  
Exemples de CNV (a) et variations ponctuelles (b) rares impliqués dans les TND (D'après Jensen et Girirajan, 2017).

Actuellement, d'après la base de données SysNDD (<https://sysndd.dbmr.unibe.ch/>), près de 1500 gènes sont impliqués dans environ 1800 maladies rares avec DI, qu'elle soit syndromique ou non syndromique, et 1248 gènes candidats de DI ont été identifiés (Kochinke *et al.*, 2016; Maia *et al.*, 2021). Ces gènes codent pour des protéines jouant un rôle dans de nombreuses voies biologiques et métaboliques, impliquées notamment dans la fonction synaptique, la régulation épigénétique de sa plasticité ou dans le développement embryonnaire cérébral, ce qui rend le choix des stratégies diagnostiques complexe, en particulier dans la DI isolée (sans orientation vers un syndrome clinique précis).

Le taux d'identification de l'étiologie varie de façon très importante selon la sévérité de la DI : une cause est identifiée dans environ 75 % des DI sévères (QI < 50) contre 20 % des DI légères, qui résultent le plus souvent d'une intrication entre des facteurs génétiques, environnementaux et psychosociaux (Des Portes et Héron, 2020).

#### 1. Les types d'anomalies génétiques impliquées dans la DI

Elles sont nombreuses et comprennent les anomalies de l'ADN mitochondrial et leur mode de transmission particulier, qui ne seront pas détaillés dans ce travail volontairement tant ils nécessiteraient de leur consacrer un chapitre entier.

### a) Les anomalies de nombre et de structure des chromosomes

Elles sont révélées par l'étude du caryotype standard, en cytogénétique conventionnelle, dont la résolution est de 5 à 10 mégabases (Mb).

Elles comprennent les aneuploïdies telles que les monosomies et les trisomies et les triploïdies (le plus souvent létales *in utero*) (Figure 3), les anomalies de structures équilibrées (translocations équilibrées) et déséquilibrées. Ces anomalies, lorsqu'elles sont déséquilibrées, sont le plus souvent à l'origine d'un syndrome polymalformatif identifié en période anténatale et, lorsque la grossesse est menée à terme et si l'enfant survit, mènent à un retard de développement global avec DI sévère (Tableau 3, Figure 4).

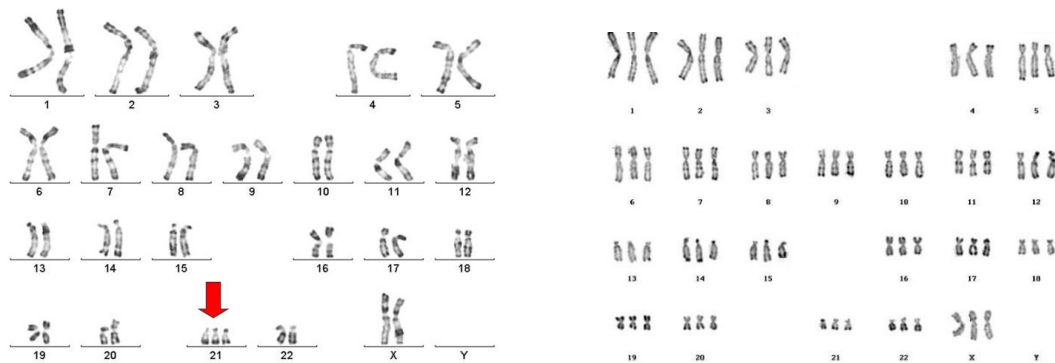


Figure 3: Caryotypes montrant une anomalie chromosomique  
Gauche : trisomie 21 libre et homogène (47, XX+21), Droite : triploïdie (69, XXX)



Figure 4: Anomalies congénitales fréquemment observées dans la trisomie 18  
(A) mains crispées, (B) pieds en piolet, (C) omphalocèle (D'après Cereda et Carey, 2012).

**Tableau 3: Principales malformations observées chez les individus atteints de trisomie 18**

Fréquence	Organe/Système	Malformations
Très fréquentes (>75%)	Cœur	CIV, CIA, pathologie valvulaire
Fréquentes (25-75%)	Génito-urinaire	Reins en fer à cheval
Peu fréquentes (5-25%)	Gastro-intestinal	Omphalocèle, atrésie de l'œsophage, sténose du pylore, diverticule de Meckel
	Système nerveux central	Hypoplasie cérébelleuse, agénésie du corps calleux, polymicrogyrie, <i>spina bifida</i>
	Crânio-facial	Fente orofaciale
	Yeux	Microphtalmie, colobome, cataracte
	Membres	Hypoplasie radiale

(d'après Cereda et Carey, 2012)

Parmi les anomalies de structure des chromosomes on peut citer les isochromosomes, qui sont caractérisés par la présence de deux copies d'un même bras chromosomique et l'absence de l'autre bras, comme dans le syndrome de Pallister-Killian ou tétrasomie 12p.

### b) Les micro-réarrangements génomiques déséquilibrés

Ces anomalies appelées CNV (copy number variation), qui correspondent à un excès (microduplication ou gain de copie) ou un défaut (microdélétion ou perte de copie) de quantité d'ADN, sont mises en évidence par les technique d'ACPA (CGH-array ou SNP array), qui ont progressivement remplacé la cytogénétique conventionnelle permettant de détecter un CNV dans 10 à 15 % des patients avec DI (Vissers *et al.*, 2003) (Figure 5). L'échelle correspond cette fois-ci à des dizaines ou des centaines de kilos bases (kb). Elle a comme avantage une meilleure résolution et la possibilité d'analyser de nombreuses cellules en même temps en s'affranchissant de la culture cellulaire. Sa principale limite est la non détection des réarrangements équilibrés ainsi qu'une faible sensibilité pour détecter les anomalies en mosaïque. La SNP array (single nucleotide polymorphism – array) permet une meilleure détection des faibles mosaïques, d'isodisomies uniparentales et hétérodisomies, ainsi que des régions d'homozygotie en cas de pathologie autosomique récessive.

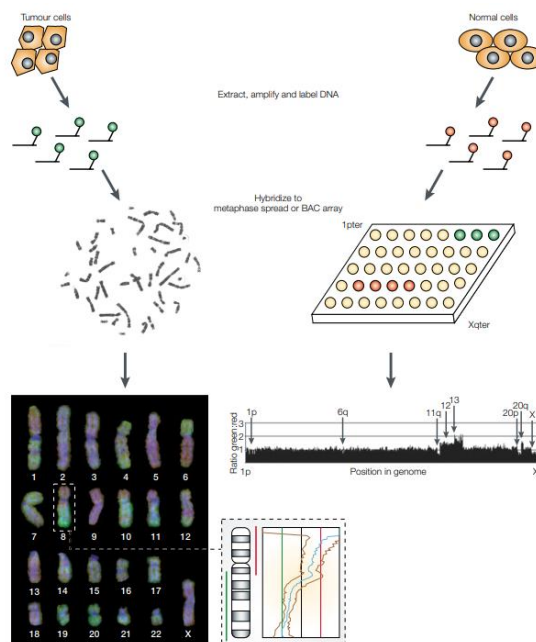


Figure 5: Technique de CGH-array

L'ADN du patient et celui de l'échantillon contrôle sont marqués avec des fluorochromes verts et rouges, respectivement, et mis en compétition pour des sites d'hybridation sur des milliers de segments d'ADN. La perte et le gain de copies apparaissent respectivement en rouge et en vert (D'après Trask, 2002)

Ces CNV sont dits pathogènes s'ils sont responsables d'un phénotype chez l'individu porteur. Ils sont le plus souvent sporadiques et surviennent *de novo*. Certaines d'entre eux sont récurrentes, avec des points de cassure précis et correspondent à des syndromes microdélétionnels ou microduplicationnels bien décrits dans la littérature scientifique, tels

que le syndrome de Wolf-Hirschhorn (délétion 4p16.3), le syndrome de Di George (délétion 22q11.2), ou le syndrome de Williams et Beuren (délétion 7q11.23) (Figure 6).

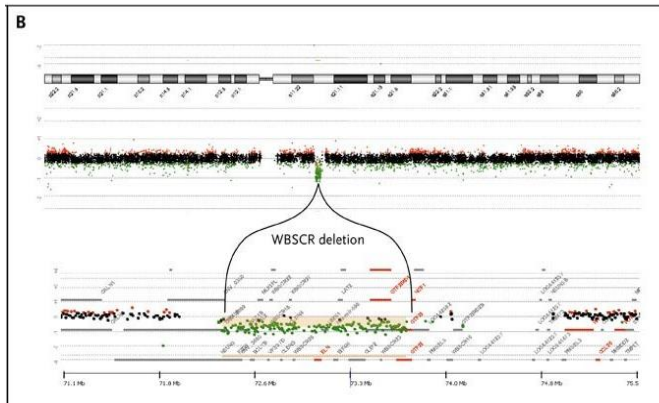


Figure 6: Délétion 7q11.23 responsable du syndrome de Williams-Beuren

L'ACPA (Agilent 244K), révèle la perte d'une copie de la région chromosomique 7q11.23 d'une taille d'environ 1,5 Mb, comme l'indique le groupe de signaux d'hybridation verts (D'après Pober, 2010).

Ces syndromes rares s'accompagnent le plus souvent d'un retard de développement et d'une DI de sévérité variable (Tableau 4, Figure 7) (Pober, 2010).

**Tableau 4: Signes cliniques observés dans le syndrome de Wolf-Hirschhorn**

Branch of medicine	Symtoms
Otolaryngology	dysplastic ears, periauricular tags, deafness (cochlear), infections of respiratory tract, recurrent otitis, protruded eyes
Ophthalmology	ocular hypertelorism, extropia, blepharoptosis, colobomata of the iris, microphthalmia, strabismus
Gastroenterology	swallowing difficulties, gastroesophageal reflux, hepatic adenomas
Cardiology	congenital heart defects
Orofacial Surgery & Dentistry	microcephaly, micrognathia, "Greek warrior helmet" appearance (broad bridge of nose), short philtrum, prominent glabella, high arched eyebrows, retardation in dental development, cranial asymmetry, cleft lip and/or palate, high forehead, wide mouth with short upper lip, cone-shaped teeth, hypodontia
Dermatology	epicanthal folds
Neurology	seizures, epilepsy, mental retardation, muscular hypotonia
Others	growth retardation, hypospadias, renal anomalies (renal agenesis, bladder exstrophy etc.), immunodeficiency (IgG, IgA), retardation of skeletal development, "sacral dimple", scoliosis, high mortality, scalp defects

(d'après Paradowska-Stolarz, 2014).



Figure 7: Patientes atteintes du syndrome de Wolf-Hirschhorn.

Dysmorphie faciale caractéristique comprenant une ensellure nasale proéminente, un nez tubulaire, des sourcils arqués, une grande bouche (d'après Battaglia et Carey, 2021; Levy *et al.*, 2022).

Certains des syndromes sont qualifiés comme étant des syndromes des gènes contigus, car il a été observé une corrélation génotype-phénotype avec notamment un effet additionnel de symptômes reliés spécifiquement et individuellement à un ou plusieurs gènes inclus dans la délétion ou la duplication. D'autres CNV récurrents sont responsables d'un phénotype très variable et dépendant de plusieurs facteurs tels que la taille du CNV, sa localisation dans le génome, et son contenu en gènes (Lesieur-Sebellin *et al.*, 2022; Levy *et al.*, 2022).

Enfin, d'autres CNV sont considérés, dans l'état actuel des connaissances scientifiques comme étant des facteurs de susceptibilité aux TND, à Pénétrance Incomplète et/ou Expressivité Variable (PIEV). C'est le cas par exemple de la délétion récurrente 15q11.2 (BP1-BP2) ou de la délétion 16p11.2 distale. Ces CNV constituent des facteurs de risque génétique, souvent hérités, qui peuvent être associés à des phénotypes variables (Coe *et al.*, 2012), ce qui rend la prédiction du phénotype, et donc le conseil génétique, parfois difficiles. Ces CNV peuvent contribuer au phénotype du patient, mais leur pathogénicité pourrait être influencée par un deuxième événement génétique (double *hit*), épigénétique et/ou environnemental, qui n'est le plus souvent pas identifié.

### c) Les formes monogéniques

Environ 1500 gènes sont actuellement connus comme étant impliqués dans des formes syndromique et non syndromique de DI (Kochinke *et al.*, 2016). Il s'agit ici d'anomalies à l'échelle d'un seul gène, à type de variation ponctuelle touchant un seul nucléotide (Single Nucleotide Variant – SNV) ou de petites insertions ou délétions touchant plusieurs nucléotides (*indel*), pouvant survenir aussi bien dans les régions codantes que non-codantes du génome. Cette description n'est pas exhaustive et inclue d'autres mécanismes mutationnels tels que les expansions et les éléments transposables, qui seront détaillés dans d'autres sous-chapitres. Les gènes impliqués dans la DI codent pour des protéines impliquées dans plus de 30 voies et fonctions cellulaires différentes (Figure 8).

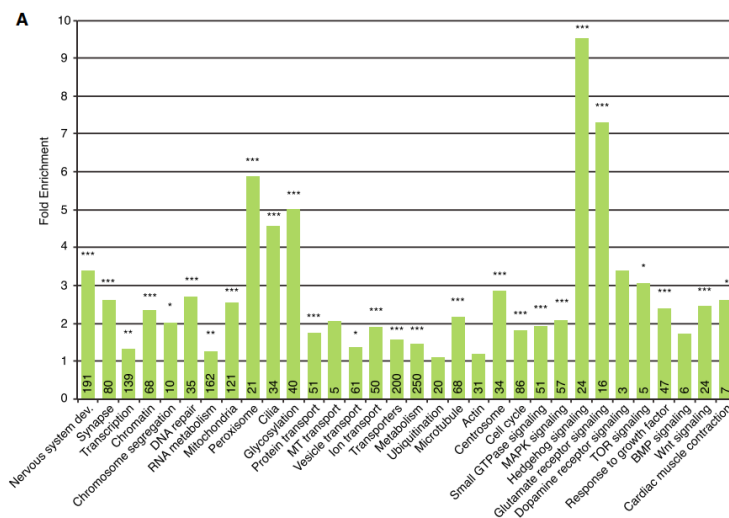


Figure 8: Annotation de la fonction des gènes associés à la DI basée sur l'ontologie

Les histogrammes montrent les enrichissements des gènes associés à la DI dans chacun des groupes basés sur l'ontologie des gènes. Le nombre total de gènes par groupe est affiché dans le diagramme correspondant (D'après Kochinke *et al.*, 2016).

Un réseau de référence a ainsi été modélisé et représente les interactions de plus de 600 gènes de DI entre eux, lesquels sont regroupés en fonction de leur fonction (Figure 9) (Kochinke *et al.*, 2016).

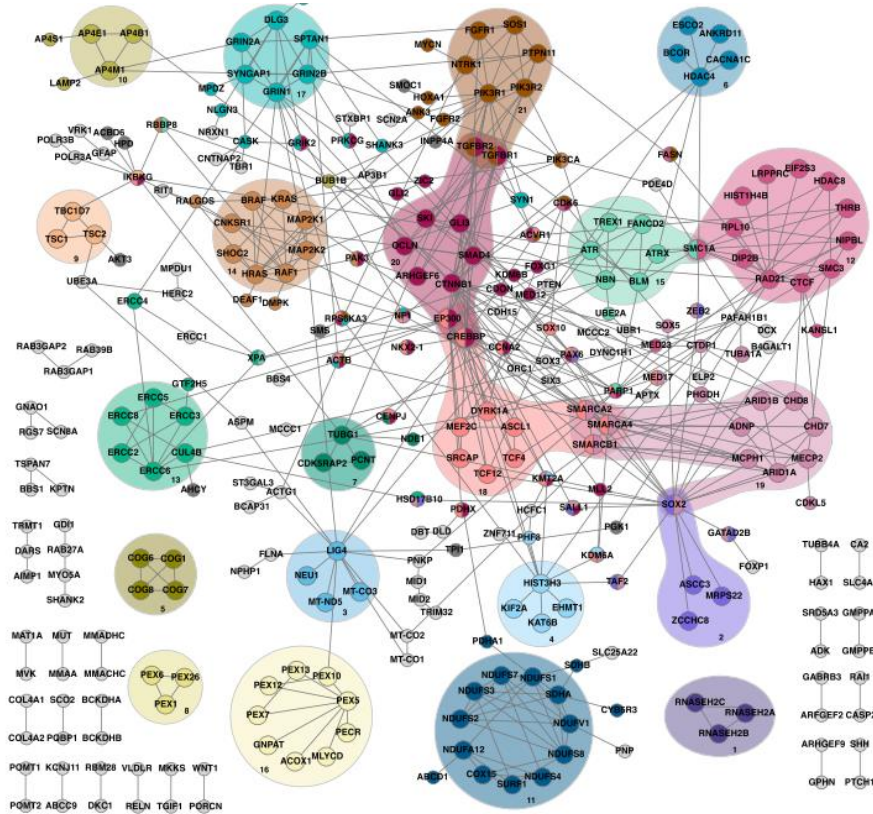


Figure 9: Réseau d'interaction des gènes associés à la DI. Les cercles indiquent les communautés de gènes fortement connectées ; des couleurs similaires illustrent la proximité fonctionnelle (D'après Kochinke *et al.*, 2016).

#### d) Les pathologies liées à l'empreinte

Certains gènes sont soumis à l'empreinte parentale, ce qui signifie que la copie active se situe soit sur l'allèle maternel soit sur l'allèle paternel. Plus d'une centaine de gènes soumis à empreinte sont identifiés (Dulac, 2010) et nombre d'entre eux sont exprimés dans le cerveau. Ces gènes sont localisés dans des régions du génome soumises à empreinte, elles-mêmes sous le contrôle de séquences d'ADN particulières (centre de l'empreinte). Une anomalie génétique entraînant l'inactivation d'un tel gène peut ainsi être responsable de deux syndromes différents selon qu'elle touche l'allèle maternel ou paternel. C'est le cas des anomalies de la région 15q11.2 qui sont responsables du syndrome d'Angelman (inactivation de l'allèle maternel) et du syndrome de Prader-Willi (inactivation de l'allèle paternel).



## 2. Les modes de transmission

Tous les modes de transmission sont représentés dans les maladies rares avec DI.

### a) Le mode autosomique dominant

Une maladie est transmise selon le mode autosomique dominant si l'anomalie génétique en cause est portée par un autosome et si la présence d'un seul allèle porteur de l'anomalie suffit pour que la maladie se manifeste, en tenant compte de la pénétrance et de l'expressivité qui peuvent être incomplète et variable, respectivement. Il y a autant de filles que de garçons atteints et une personne malade a un risque de 50% de transmettre l'allèle anormal à chacun de ses enfants. La transmission est donc verticale et une personne qui ne porte pas l'allèle pathologique ne transmet pas la maladie à ses enfants.

Une variation pathogène peut être héritée d'un parent ou survenir *de novo* (néomutation). Les variations pathogènes *de novo* représentent environ 60% des cas de maladies rares avec DI sévère (Gilissen *et al.*, 2014; Kaplanis *et al.*, 2020; Maia *et al.*, 2021) (Figure 10) telles que les syndromes de Mowat-Wilson (gène *ZEB2* (Mowat *et al.*, 2003)), de Kleefstra (gène *EHMT1* (Kleefstra *et al.*, 2009)) ou de Tatton-Brown-Rahman (gène *DNMT3A* (Tatton-Brown *et al.*, 2014)) etc. Dans certains cas, le variant pathogène identifié chez un enfant atteint de DI peut avoir été transmis par un parent porteur hétérozygote et pauci voire asymptomatique, comme c'est le cas pour le gène *STX1B* (Wolking *et al.*, 2019).

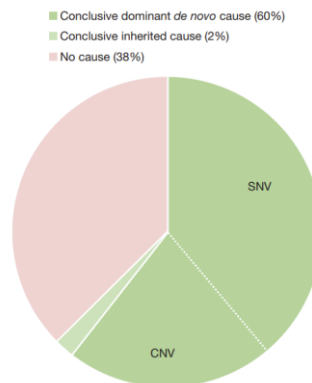


Figure 10: Proportion des variants pathogènes *de novo* dans la DI sévère (D'après Gilissen *et al.*, 2014).

### b) Le mode autosomique récessif

Une maladie est transmise selon le mode autosomique récessif si l'anomalie génétique en cause est portée par un autosome et si la présence de deux allèles porteurs de l'anomalie est nécessaire pour que la maladie se manifeste. Les malades sont homozygotes ou hétérozygotes composites pour le variant en cause. Deux conjoints porteurs sains hétérozygotes ont, à chaque grossesse, un risque de 25% d'avoir un enfant atteint quel que soit son sexe. C'est le cas des syndromes de Smith-Lemli-Opitz (gène *DHCR7*, (Wassif *et al.*, 1998), et de la majorité

des maladies métaboliques. D'autres formes autosomiques récessives de DI sont bien identifiées telles que le syndrome de Cohen (gène *VPS13B*, (El Chehadeh *et al.*, 2011; Kolehmainen *et al.*, 2003) ou le grand groupe des ciliopathies dont font partie les syndromes de Joubert (Parisi et Glass, 1993), Meckel (Knopp *et al.*, 2015) et Bardet-Biedl (Muller *et al.*, 2010), secondaires à des variants pathogènes bi-alléliques (voire tri-alléliques) et caractérisés par une très grande hétérogénéité génétique et phénotypique. L'étude de cohortes de patients avec DI sans orientation clinique par séquençage de génome montre que les causes autosomiques récessives représentent environ 2% des diagnostics (Gilissen *et al.*, 2014).

La transmission est horizontale et peut être favorisée par la présence d'une consanguinité. En effet, deux conjoints apparentés ayant un ancêtre commun, ils sont de ce fait plus à risque que deux individus non apparentés de la population générale d'être porteurs hétérozygotes de la même variation pathogène au sein du même gène (Musante et Ropers, 2014).

### *c) Le mode de transmission lié à l'X*

Les individus de sexe masculin n'ont qu'un seul chromosome X et les femmes en possèdent deux. La question de la dominance ou de la récessivité ne se pose donc pas chez les individus de sexe masculin (soit ils portent le variant pathogène (hémizygotés) et sont atteints, soit ils ne le portent pas et sont sains) mais uniquement chez les individus de sexe féminin.

#### ➤ ***Le mode récessif lié à l'X***

Les femmes hétérozygotes sont dites conductrices et sont le plus souvent asymptomatiques ou présenter de discrets signes. Les femmes homozygotes ou hétérozygotes composites pour l'allèle pathologique étant rares, les individus atteints sont essentiellement, voire exclusivement des hommes. Le risque de transmission des maladies récessives liées à l'X est différent suivant le sexe du parent porteur du variant:

- une femme hétérozygote conductrice a un risque de 50% de transmettre son chromosome X portant l'allèle muté à chacun de ses enfants. Ses filles ont un risque de 50% d'être conductrices et ses fils ont un risque de 50% d'être atteints.
- un homme atteint transmettra l'allèle muté à toutes ses filles qui seront donc toutes conductrices. En revanche, aucun de ses fils ne sera atteint, ni ne transmettra la maladie.

Parmi les maladies récessives liées à l'X avec DI on peut citer l'adrénoleucodystrophie, le syndrome de Menkès, le syndrome de Lesch-Nyhan.

#### ➤ ***Le mode dominant lié à X***

Les deux sexes peuvent être touchés par la maladie, mais en général, les filles hétérozygotes sont moins sévèrement atteintes que les garçons hémizygotés, chez qui la

maladie peut parfois être létale. La transmission est verticale mais diffère de l'hérédité autosomique dominante car il n'y a jamais de transmission père-fils. Comme pour cette dernière, la pénétrance peut être incomplète et l'expressivité peut être variable.

La transmission des maladies dominantes liées à l'X est également différente suivant que le parent atteint est le père ou la mère:

- une mère hétérozygote (atteinte) a un risque de 50% de transmettre son chromosome X portant l'allèle muté à chacun de ses enfants quel qu'en soit le sexe, qui sera alors atteint.
- un père atteint transmet son chromosome X portant l'allèle muté à toutes ses filles qui seront donc toutes atteintes. En revanche aucun de ses fils ne sera atteint.

Parmi les maladies avec DI dominantes liées à l'X on peut citer le syndrome de l'X fragile, le syndrome de Coffin Lowry ou le déficit en OTC (ornithine transcarbamylase).

Néanmoins, nous verrons dans la partie II dédiée à la DI liée au chromosome X, que cette distinction entre modes récessif et dominant lié à l'X n'est pas toujours si aisée. Cette classification est, de ce fait, de moins en moins utilisée et est progressivement remplacée par le terme de transmission liée à l'X.

### *3. Le diagnostic de la DI : quels besoins? Quels enjeux?*

Le diagnostic de la DI représente un enjeu majeur de santé publique puisque touchant 1 à 2% de la population (Roeleveld *et al.*, 1997) avec près de la moitié des individus qui demeurent sans diagnostic (Buntinx *et al.*, s. d.). Les raisons d'identifier la cause moléculaire de la DI sont multiples et peuvent être résumées ainsi :

#### *a) En clinique, pour les familles*

##### **▪ La fin d'une longue odyssée...**

L'augmentation du taux de diagnostic chez les patients atteints de DI de cause rare permet de mettre fin à l'errance diagnostique des familles durement éprouvées et aux analyses et bilans à visée étiologique, notamment les examens invasifs (biopsie musculaire, ponction lombaire etc.). Le diagnostic est souvent vécu par les familles comme étant la fin d'une longue odyssée, et une déculpabilisation pour les parents, soulagés d'avoir une explication, qui plus est, une cause dont ils ne sont pas responsables. Le fait d'établir un diagnostic ne permet pas, la plupart du temps, de mettre en place un traitement curatif, faute malheureusement de thérapies disponibles dans les maladies rares avec DI, mais cela permet d'améliorer la prise en charge globale, notamment socio-éducative, des patients. La connaissance de la maladie rare permet également d'adapter le suivi et la surveillance du patient (dépistage d'un trouble neurosensoriel, sensibilisation des parents en cas de risque important d'apparition

d'une épilepsie) et de prévenir l'apparition d'éventuelles complications. Cela conduit également à affiner le pronostic et, parfois, de mieux appréhender l'évolution de la maladie (progrès croissants ou à l'inverse, risque de régression neurologique). Dans de plus rares cas, l'identification de la cause moléculaire peut mener à une thérapeutique curative précise et ainsi stabiliser l'évolution de la pathologie, notamment quand le diagnostic est établi très rapidement après la naissance. C'est le cas de la phénylcétonurie, qui touche 1 nouveau-né sur 16 000 en France et est dépistée à la naissance depuis 1972. Ce dépistage précoce, d'abord biochimique puis confirmé sur le plan moléculaire, permet la mise en place très précoce chez l'enfant atteint d'un suivi spécialisé et d'un régime alimentaire spécifique, pauvre en phénylalanine, permettant d'éviter l'apparition des manifestations de la maladie telles qu'un retard de développement avec DI et microcéphalie.

- **Un conseil génétique précis...**

Un intérêt majeur de l'identification de la cause moléculaire de la DI est celui du conseil génétique. Elle permet, pour un couple ayant un enfant atteint de DI, de préciser le risque récidive pour une future grossesse, qui peut aller de moins de 1% jusqu'à 50%. L'identification de la cause moléculaire permet également aux couples à risque, selon la pathologie, de bénéficier pour leur projet parental d'un accompagnement vers une démarche diagnostic préimplantatoire (DPI) ou de diagnostic prénatal (DPN).

- **Une aide à la prise de décision en cas de grossesse à risque...**

Il est actuellement possible lorsque cela est indiqué, d'établir très précocement un diagnostic moléculaire chez un fœtus, par le biais d'un prélèvement fœtal (liquide amniotique, biopsie de placenta, sang fœtal voire sang maternel), soit en cas d'anomalies échographiques visualisées, soit après identification d'un variant pathogène chez un premier enfant du couple (ou un parent en cas de pathologie parentale). En France, lorsqu'il est prouvé qu'un fœtus est atteint d'une maladie grave et incurable au moment du diagnostic, telle que la DI, une demande d'interruption médicale de grossesse émanant du couple peut être acceptée par le CPDPN dont il dépend. Cela permet ainsi aux couples concernés, après avoir obtenu des informations sur la pathologie en consultation de conseil génétique, de prendre une décision en connaissance de cause vis-à-vis de la prise en charge ultérieure de la grossesse.

- **Se sentir moins isolés...**

L'obtention d'un diagnostic permet aux familles concernées de communiquer entre elles *via* de nombreux réseaux sociaux et surtout des Associations de patients, et ainsi de partager leurs expériences, de rechercher des conseils et astuces pratiques, de se rassembler voire de collecter des fonds pour la recherche sur une maladie rare donnée.

### *b) En recherche, pour la Science*

Les maladies monogéniques avec DI représentent une approche intéressante pour la génomique fonctionnelle, en permettant de mettre évidence de nouveaux mécanismes moléculaires et cellulaires, avec une immense variété de variants et de phénotypes. De plus, de nombreux modèles cellulaires et animaux sont déjà disponibles et peuvent encore être développés en vue d'aider, lorsqu'ils sont transposables à l'humain, à corrélérer les données génotypiques au phénotype. Ainsi, l'identification d'une variation pathogène permet dans certains cas aux familles de participer à des projets de recherche spécifiques.

Sur le plan clinique, les travaux de recherche sur les maladies rares avec DI permettent d'améliorer de manière considérable les connaissances, notamment par le biais une meilleure description phénotypique des maladies ultra-rares (*reverse phenotyping*). Ils mènent encore régulièrement à la découverte de nouveaux gènes et de nouveaux mécanismes physiopathologiques, pouvant parfois aboutir à de nouvelles pistes thérapeutiques (thérapie génique ou protéique). Ceci souligne le caractère transversal de la médecine génomique, qui, du patient à la paillasse, balaye tous les champs de la médecine contemporaine.

### *c) Au niveau national*

La DI, par les examens diagnostiques, la prise en charge thérapeutique et l'accompagnement qu'elle nécessite, représente 10 % des dépenses de santé et donc un enjeu majeur de santé publique (Pr Battin, « Actualités du déficit intellectuel d'origine génétique », séance du 10 Avril 2018).

En France, la Loi relative à la politique de Santé Publique du 9 août 2004 détermine que « La prise en charge des maladies rares constitue une priorité ». Pour ce faire, un 1er Plan National Maladies Rares (PNMR1, 2005-2008) a permis de structurer l'organisation de l'offre de soins pour les maladies rares et d'améliorer sa lisibilité pour les patients. Pour y parvenir, vingt-trois filières de santé maladies rares ont été créées. Ces filières regroupent des centres de référence (CRMR) et de compétence maladies rares (CCMR), des centres de ressources, des professionnels de santé, des laboratoires de diagnostic et de recherche, des sociétés savantes, des structures médico-sociales, des associations de patients, des partenaires privés et européens (ERN), dans le but de co-construire une réponse aux enjeux posés par les maladies rares (prise en charge, diagnostic, recherche, formation).

La filière DEFISCIENCE (Maladies rares du neurodéveloppement) est la filière de santé maladie rare dédiée aux maladies rares avec TND. Elle comporte 12 CRMR, 11 CCMR, de nombreux laboratoires de diagnostic/recherche et associations de patients et est coordonnée par le Pr Vincent Des Portes (CHU de Lyon) (<https://defiscience.fr>).

## II. La DI liée au chromosome X (DILX)

### 1. Historique

Dans la plupart des espèces eucaryotes, le sexe de l'individu est déterminé à partir de la paire de chromosomes sexuels ou gonosomes. Chez les mammifères, la femelle est homogamétique avec deux chromosomes X et le mâle est hétérogamétique avec un chromosome X et un chromosome Y (Vicoso et Bachtrog, 2009).

La notion de DI liée au chromosome X (DILX) a été évoquée dès la fin des années 1930, notamment par Penrose, qui décrit un excès d'hommes parmi les individus déficients intellectuels institutionnalisés (Haldane, 1938). En 1943, Martin et Bell rapportèrent une grande famille au sein de laquelle plusieurs hommes, reliés par des femmes, présentaient une DI (Martin et Bell, 1943) et donnèrent leur nom à ce syndrome (Figure 11).

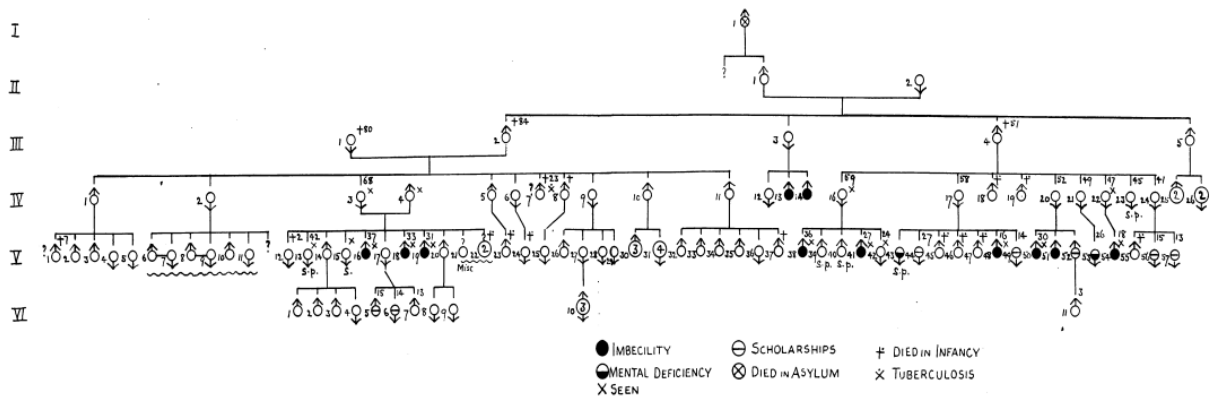


Figure 11: Famille princeps décrite en 1943 par Martin et Bell. (D'après Martin et Bell, 1943).

En 1969, Lubs mit en évidence une zone de fragilité au niveau de l'extrémité du bras long du chromosome X chez ces individus. Cette observation fut à l'origine de la dénomination du syndrome de l'X Fragile, qui représenta alors la première forme de DI localisée sur le chromosome X et, encore aujourd'hui, la première cause de DI héréditaire (Lubs, 1969; Crawford *et al.*, 2001) (Figure 12).

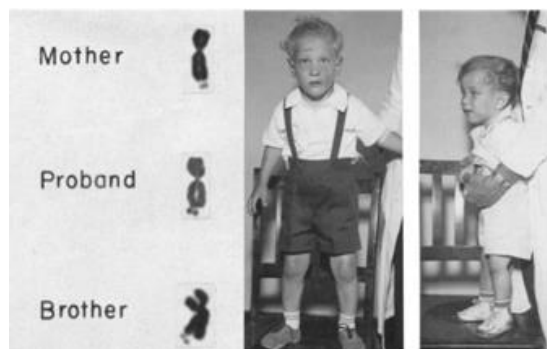


Figure 12: Marqueur du syndrome de l'X fragile.

Gauche : Chromosomes X porteurs du marqueur anormal chez deux cas index décrit par Lubs en 1969 et leur mère. Droite : photographies des deux jeunes cas index (D'après Lubs, 1969).

En 1962, Renpenning reconstitua l'histoire d'une grande famille ayant émigré depuis les Pays-Bas vers l'Ukraine puis au Canada, au sein de laquelle 20 hommes sur 4 générations présentaient une DI avec microcéphalie, petite taille et testicules hypoplasiques, reliés par des femmes saines, lui faisant évoquer un mode de transmission récessif lié à l'X (Renpenning *et al.*, 1962). Ce syndrome, dont on sait maintenant qu'il est secondaire à des variants pathogènes du gène *PQBP1* (Kalscheuer *et al.*, 2003), portera son nom (Figure 13).

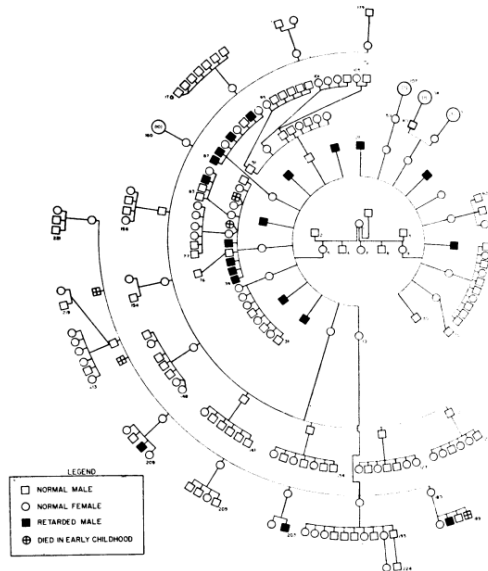


Figure 13: Famille princeps décrite par Renpenning en 1962.

Les carrés noirs représentent les individus masculins atteints de DI (D'après Renpenning *et al.*, 1962).

Concomitamment, en 1961, Mary Lyon décrivait pour la première fois le phénomène d'inactivation du chromosome X (appelé de ce fait lyonisation), mécanisme épigénétique complexe qui maintient un seul chromosome X transcriptionnellement actif par cellule (Lyon, 1961). Peu de temps avant, en 1959, Ohno démontrait que le corpuscule de Barr, satellite nucléolaire d'1  $\mu\text{m}$  de diamètre présent uniquement dans les cellules féminines (Barr et Bertram, 1949), était constitué d'hétérochromatine condensée du chromosome X inactivé (Ohno *et al.*, 1959) (Figure 14).

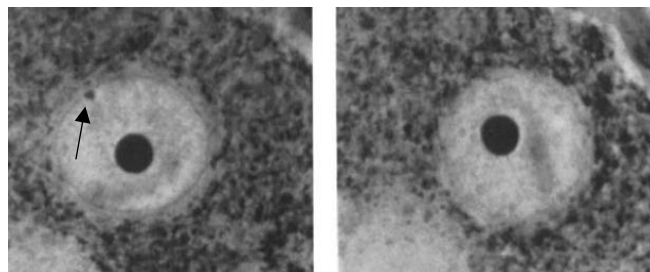


Figure 14: Le corpuscule de Barr.

Photos au microscope standard de neurones de chat femelle (gauche) et mâle (droite). On observe le corpuscule de Barr dans le neurone femelle, situé sous la membrane nucléaire (flèche) (D'après Barr *et al.*, 1950).

En 1977, le locus du syndrome de l'X fragile est localisé à la région Xq27.3 (Sutherland, 1977). Dans les années 1980, on note que les arbres généalogiques des familles concernées

par le syndrome de l’X fragile ne vont pas dans le sens d’une transmission liée à l’X classique puisque que le chromosome X « fragile » peut avoir été transmis par un grand-père sain (Nielsen *et al.*, 1981) (Figure 15). Ceci remet en cause la transmission sur le mode récessif. La transmission apparaît plus complexe... On parle de transmission dominante liée à l’X avec expressivité variable.

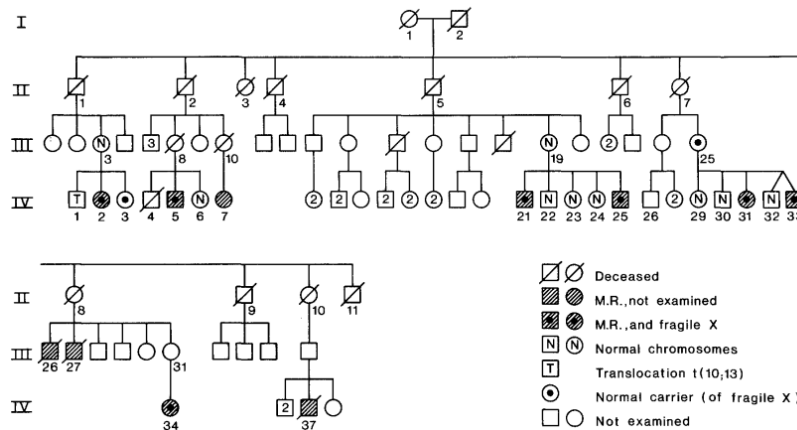


Figure 15: Famille décrite par Nielsen en 1981.

On voit la présence d’hommes et de femmes atteints de DI (génération IV), tous reliés par des femmes conductrices (génération III), porteuses ou non du marqueur du chromosome X, lesquelles sont reliées entre elles par des femmes ou des hommes asymptomatiques (génération II) (D’après Nielsen *et al.*, 1981).

Dans les années 1990, de nombreuses découvertes majeures sont réalisées sur le syndrome de l’X fragile, en grande partie grâce au Professeur Jean-Louis Mandel et son équipe. Le gène *FMR1* fut identifié comme étant à l’origine du syndrome, du fait de variations du nombre de répétitions CGG (amplification de triplets) dans sa région 5’UTR, responsables de son inactivation. Puis on découvrit que la mutation en cause était instable et augmentait en taille au fur et à mesure des générations, introduisant les notions de prémutation et de mutation complète, si spécifiques à ce syndrome et inconnues jusqu’alors (Fu *et al.*, 1991; Oberle *et al.*, 1991; Vincent *et al.*, 1991).

## 2. Le diagnostic de la DILX : que sait-on?

### 1. Un mode de transmission particulier

#### a) Un conseil génétique pas toujours aisé

Le mode de transmission lié à l’X est particulier de par l’existence de femmes conductrices qui sont très fréquemment asymptomatiques dans les formes récessives. On comprend bien que ceci a une importance majeure en termes de conseil génétique. L’identification de l’anomalie génétique en cause sur le chromosome X dans ces familles permet aux femmes à risque de se faire dépister et de bénéficier, lorsqu’elles sont conductrices, d’un accompagnement pour tout projet parental par le biais d’une démarche DPN ou DPI. Dans ces familles, l’annonce diagnostique est souvent particulièrement éprouvante, du fait du



sentiment de culpabilité qu'elle génère fréquemment chez la mère (et parfois grand-mère) conductrice, notamment lorsque plusieurs oncles ou grands-oncles sont atteints, le diagnostic n'ayant pas pu être fait chez ces derniers faute d'analyse génétique disponible à leur époque.

Dans les formes de DI dominantes liée à l'X, les femmes peuvent présenter un phénotype très variable allant de l'absence de symptôme à une DI sévère, comme c'est le cas dans le syndrome de l'X fragile, en passant par une forme pauci symptomatique, comme dans le syndrome de Coffin Lowry (gène *RPS6KA3*) (Pereira *et al.*, 2010), ou le syndrome de Simpson-Golabi-Behmel (gènes *GPC3* et *GPC4*) (Schirwani *et al.*, 2019).

Par ailleurs, de plus en plus de gènes de l'X, initialement décrits comme étant responsables de DI chez les garçons sont rapportés comme étant également impliqués chez des filles atteintes de DI. C'est le cas notamment des gènes *CLCN4* (Palmer *et al.*, 2018), *HUWE1* (Moortgat *et al.*, 2018), *IQSEC2* (Mignot *et al.*, 2019), *KDM5C* (Carmignac *et al.*, 2020), *USP9X* (Jolly *et al.*, 2020) et *NEXMIF* (Stamberger *et al.*, 2021). Pour compliquer encore plus les choses, certains variants dans des gènes de l'X tels que *MECP2*, *CDKL5*, *DDX3X*, ou *PCDH19* affectent préférentiellement les filles hétérozygotes. Ces variants surviennent généralement *de novo* et les garçons hémizygotés ne sont pas viables ou ne survivent que si les variants sont hypomorphes ou en mosaïque post-zygotique (Weaving *et al.*, 2004; Depienne et LeGuern, 2012; Snijders Blok *et al.*, 2015).

Cette complexité a conduit certains auteurs à recommander l'utilisation du terme « maladies liées à l'X » (Dobyns *et al.*, 2004; Lubs *et al.*, 2012) plutôt que de parler de DILX récessive et dominante, qui est cependant toujours le plus souvent de mise actuellement notamment dans la base de données OMIM (Online Mendelian Inheritance in Man) (<https://omim.org/>).

### b) Le phénomène d'inactivation de l'X

#### ❖ Le phénomène d'inactivation aléatoire du chromosome X

Chez les mammifères, et notamment chez l'Homme, un des mécanismes de compensation génique entre mâle et femelle est basé sur l'inactivation de l'un des 2 chromosomes X chez la femelle (Migeon, 2016) (Figure 16).

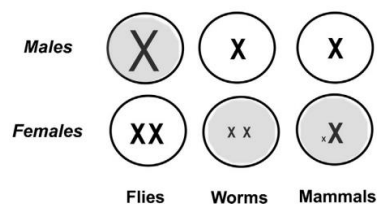


Figure 16: Mécanismes de compensation de dosage génique.

Les cellules en gris sont les cellules qui sont la cible moléculaire de la compensation (D'après Migeon, 2016).

Elle aboutit à la présence d'un seul chromosome X actif (en dehors des régions pseudo-autosomiques (PARs)), la transcription de l'autre chromosome X étant réprimée. C'est un phénomène épigénétique (une régulation de l'expression des gènes sans altération de la séquence nucléotidique) qui opère par une modification des histones, une méthylation de l'ADN ou l'action d'ARN non-codant. Le phénomène d'inactivation du chromosome X (XCI) survient très précocement durant le développement embryonnaire (Monk et Harper, 1979). Chez l'Homme, l'XCI est aléatoire entre le chromosome X d'origine paternelle et celui d'origine maternelle. Cette répartition aléatoire est variable selon les individus, suivant une courbe de Gauss (Amos-Landgraf *et al.*, 2006), et selon l'âge (Knudsen *et al.*, 2007) (Figure 17). La transmission du chromosome X inactivé se fait ensuite de façon clonale et stable aux cellules filles (Balaton *et al.*, 2018).

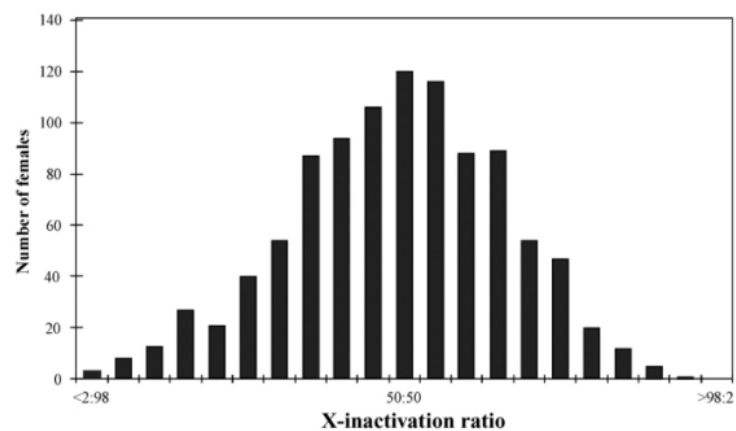


Figure 17: Ratios d'inactivation de l'X chez 1005 femmes. Les ratios suivent une distribution normale allant de 2:98 à 98:2 (D'après Amos-Landgraf *et al.*, 2006).

Par conséquent, les individus de sexe féminin sont des « mosaïques cellulaires », chaque cellule exprimant des gènes provenant du chromosome X maternel ou paternel de manière aléatoire (Figure 18).

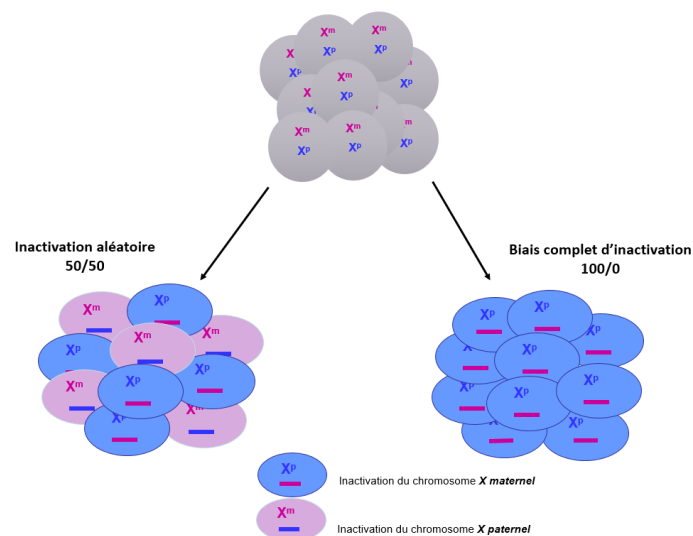


Figure 18: Inactivation aléatoire du chromosome X dans un embryon humain. (D'après S El Chehadeh, 2012, cours aux étudiants sage-femme).

Cependant, un sous-ensemble de gènes, qui peut varier selon les individus et les tissus, échappe à l'inactivation X et continue à s'exprimer à partir des deux chromosomes X, plusieurs processus expliquant ce phénomène (Tukiainen *et al.*, 2017; Navarro-Cobos *et al.*, 2020) (Figure 19, Figure 20).

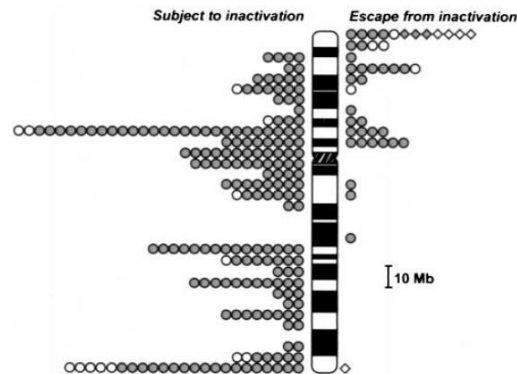


Figure 19: Gènes échappant au mécanisme d'inactivation de l'X. (D'après Carrel *et al.*, 1999).

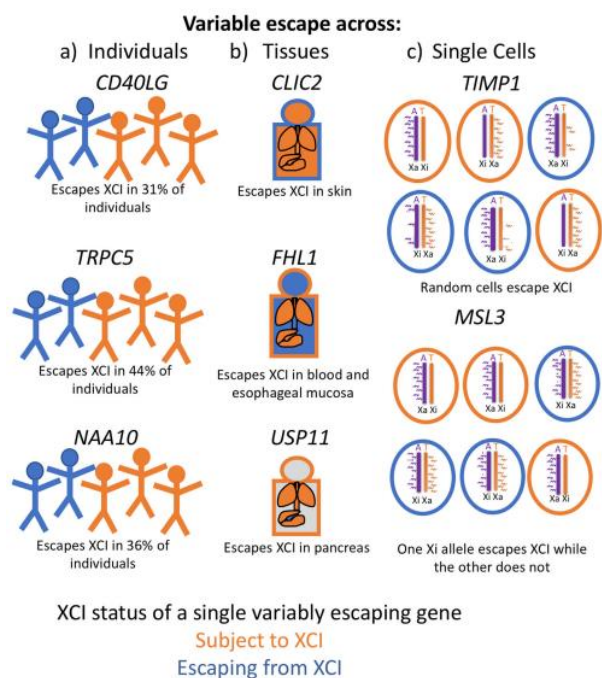


Figure 20: Différents niveaux d'échappement à l'XCI. Certains gènes échappent de manière variable à l'XCI selon les individus (a), selon les tissus (b), ou les types cellulaires (c) (D'après Navarro-Cobos *et al.*, 2020).

### ❖ **Le biais d'inactivation du chromosome X (ou inactivation non aléatoire)**

En théorie, l'inactivation étant aléatoire dans les cellules et transmise de façon clonale aux cellules-filles, l'X inactivé est d'origine paternelle dans 50% des cellules et d'origine maternelle dans les 50 % de cellules restantes. Cependant, chez certaines femmes, il existe un biais d'inactivation, c'est-à-dire que l'un des 2 chromosomes X est préférentiellement inactivé par rapport à l'autre, dans plus de 85% des cellules.

Un biais d'XCI est fréquemment observé chez les femmes conductrices d'une pathologie récessive liée à l'X, qui sont le plus souvent asymptomatique ou paucisymptomatiques, et chez lesquelles il a été observé que le chromosome X porteur du variant pathogène est souvent préférentiellement inactivé. C'est par exemple le cas dans le syndrome de Lesch-Nyhan (gène *HPRT1*), une erreur innée du métabolisme des purines résultant en une hyperuricémie avec hyperuraturie, une DI modérée à sévère, les troubles du comportement et une possible atteinte rénale (Lesch et Nyhan, 1964). Les femmes conductrices sont très majoritairement asymptomatiques (O'Neill, 2004). Cependant, l'excrétion urinaire d'acide urique, xanthine et hypoxanthine peut être légèrement élevée chez ces dernières qui peuvent développer une goutte (Puig *et al.*, 1998). Plus rarement, des filles atteintes de la forme classique et sévère du syndrome de Lesch-Nyhan ont été décrites et présentaient un biais complet d'XCI, responsable d'une inactivation préférentielle du chromosome X non muté (Aral *et al.*, 1996).

Ainsi, dans plusieurs maladies récessives liées à l'X, l'observation d'un phénotype équivalent à celui des garçons atteints chez les filles hétérozygotes a été attribué à la présence d'un biais complet de l'XCI avec inactivation préférentielle du chromosome X non muté. Ces conclusions ont cependant été remises en question après l'observation d'une discordance possible chez certaines femmes, avec une absence de corrélation observée entre le résultat de l'étude de l'XCI et la sévérité du phénotype des filles hétérozygotes (El Chehadeh *et al.*, 2017). Plusieurs hypothèses ont alors été émises pour tenter d'expliquer cette absence de corrélation comme le fait que (1) le phénomène de biais d'inactivation augmente avec l'âge (Amos-Landgraf *et al.*, 2006), (2) la survenue *de novo* de la maladie pourrait avoir un effet sur l'XCI en empêchant la survenue d'un biais « protecteur », par un mécanisme encore méconnu (Grasshoff *et al.*, 2011), (3) des gènes (environ 15% chez l'homme) échapperaient à l'inactivation de l'X (Berletch *et al.*, 2010), (4) l'inactivation de l'X pourrait varier selon les tissus (mosaïque tissulaire), conduisant à des altérations de dosage tissu-spécifiques, avec pour conséquence, une activation préférentielle du chromosome X anormal dans le cerveau des filles atteintes très tôt au cours de l'embryogenèse, entravant ainsi le développement cérébral (Bijlsma *et al.*, 2012; Mayo *et al.*, 2011; Novara *et al.*, 2014).

## 2. Les causes monogéniques de DILX

La DI liée au chromosome X (DILX) a suscité très tôt beaucoup d'intérêt en raison de la surreprésentation des hommes déficients ; les scientifiques ont ainsi consacré beaucoup d'énergie à déchiffrer l'origine génétique de la DILX.

Le chromosome X, séquencé en 2005, est un chromosome d'environ 155 Mb (Ross *et al.*, 2005). Il représente environ 5% de la totalité du génome humain et comporte 829 gènes codants dont 205 sont associés à au moins une maladie monogénique du développement

(Leitao *et al.*, 2022). Le nombre de gènes de l’X impliqués dans des formes monogéniques de DI croît progressivement, allant de 106 en 2013 jusqu’à 127 actuellement. De plus, 55 gènes supplémentaires situés sur le chromosome X sont aujourd’hui considérés comme des gènes candidats de DILX ( Piton *et al.*, 2013; Kochinke *et al.*, 2016; Neri *et al.*, 2018; Leitao *et al.*, 2022) (Figure 21).

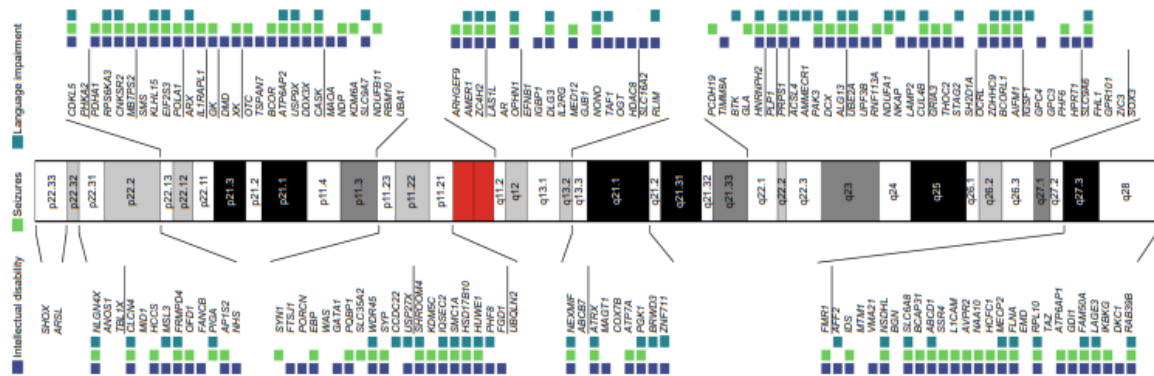


Figure 21: Gènes associés à des signes neurologiques sur le chromosome X. Les carrés illustrent l'association avec la DI (bleu), l'épilepsie (vert), troubles du langage (cyan) (D'après Leitão *et al.*, 2022).

Sur le millier de gènes présents sur le chromosome X, 10% sont décrits comme étant impliqués dans la DI (contre 3% des gènes autosomiques) (Chiurazzi et Pirozzi, 2016). La densité des gènes du chromosome X responsables de DI est considérée à l’heure actuelle comme étant deux fois supérieure à celle des gènes de DI sur les autosomes (Schwartz *et al.*, 2023). Cette importante proportion a justifié le développement d’analyses de séquençage haut débit par panels ciblés sur les gènes de DI situés sur le chromosome X (Hu *et al.*, 2016; Ibarluzea *et al.*, 2020; Redin *et al.*, 2014; Tarpey *et al.*, 2009).

### a) Classification des gènes de DILX

Il est possible de classer ces gènes selon les conséquences phénotypiques, selon qu’il s’agisse de DI non-syndromique (la DI est le symptôme au premier plan et ne s’accompagne pas d’un syndrome malformatif ou dysmorphique franc) ou syndromique (la DI est associée à des éléments dysmorphiques de la face ou d’autres parties du corps, des signes neurologiques et/ou des malformations congénitales). Cette classification peut être utile en pratique clinique en orientant parfois le clinicien et l’interprétation des analyses vers un gène ou un groupe de gènes en particulier. Cependant, nous savons maintenant qu’il peut exister une importante variabilité phénotypique et qu’un nombre croissant de gènes explique à la fois des formes de DI non-syndromique et de DI syndromique. C’est le cas, par exemple, du gène *ARX*, responsable de plusieurs entités phénotypiques appartenant à un continuum clinique qui inclue le syndrome XLAG (DI sévère, lissencéphalie, anomalies génitales) (Bonneau *et al.*, 2002), le syndrome de Proud (DI sévère, agénésie du corps calleux, épilepsie, spasticité)

(Proud *et al.*, 1992), le syndrome de Partington (DI avec dystonie et dyspraxie buccolinguale, troubles de la motricité fine, dysarthrie avec bavage) (Strømme *et al.*, 2002), et à l'extrême du spectre, une DI isolée (Wallerstein *et al.*, 2008). Cette distinction entre formes syndromiques et non-syndromiques est donc parfois peu aisée, notamment du fait d'une possible pénétrance incomplète avec expressivité variable et de l'existence de variants « mild » (Ropers et Hamel, 2005; Stevenson et Schwartz, 2009).

Enfin, il est également possible de classer les gènes de DI selon le mécanisme physiopathologique dans lequel ils sont impliqués (voie de signalisation intra-cellulaire, type cellulaire, tissu etc.). Ceci permet une classification pertinente et cohérente sur le plan fonctionnel (van Bokhoven, 2011). Il faut néanmoins garder à l'esprit que certains gènes ayant une fonction ubiquitaire peuvent apparaître dans plusieurs sous-catégories.

Nous nous baserons sur ce type de classification dans le cadre de cette thèse, en nous focalisant sur les gènes impliqués dans la DILX.

#### ❖ **Les gènes impliqués dans la régulation de la transcription :**

Un grand nombre de gènes de DILX est impliqué dans la régulation de l'expression des gènes. Il s'agit souvent de facteurs de transcription, de protéines intervenant dans la méthylation de l'ADN, la modification des histones, ou le remodelage de la chromatine (Tableau 5). Parmi eux figure le gène *MECP2*, dont je parlerai plus en détail plus loin.

**Tableau 5: Gènes de DILX impliqués dans la régulation de la transcription**

Gènes	Fonction	Phénotype
<i>MED12</i> (Xq13.1)	Initiation de la transcription	Syndrome d'Ohdo, de Lujan-Fryns, d'Opitz-Kaveggia (Maia <i>et al.</i> , 2023)
<i>MECP2</i> (Xq28)	Répression de la transcription	Syndrome de Rett, de duplication du gène <i>MECP2</i> , encéphalopathie néonatale sévère, DILX de type 13 (Van Esch <i>et al.</i> , 2005)
<i>CDKL5</i> (Xp22.13)	Epissage des ARN	Encéphalopathie épileptique (Weaving <i>et al.</i> , 2004)
<i>KDM5C</i> (Xp11.22)	Méthylation des histones	DI avec microcéphalie et paraparésie spastique (Carmignac <i>et al.</i> , 2020)
<i>PHF8</i> (Xp11.22)	Facteur de transcription	DILX avec fente palatine de type Siderius (Laumonier <i>et al.</i> , 2005)
<i>ATRX</i> (Xq21.1)	Remodelage de la chromatine	DILX avec $\alpha$ -Thalassémie (Gibbons <i>et al.</i> , 1995)
<i>BRWD3</i> (Xq21.1)	Modification de la chromatine	DILX avec macrocéphalie (Field <i>et al.</i> , 2007)
<i>PHF6</i> (Xq26.2)	Facteur de transcription	Borjeson-Forssman-Lehmann (Lower <i>et al.</i> , 2002)
<i>HUWE1</i> (Xp11.2)	Ubiquitination des histones	DILX de type Turner (Moortgat <i>et al.</i> , 2018)
<i>RPS6KA3</i> (Xp22.2-22.1)	Phosphorylation des histones Assemblage fuseau mitotique Transmission synaptique	Syndrome de Coffin Lowry (Field <i>et al.</i> , 2006)
<i>KDM5C</i> (Xp11.22-21)	Méthylation des histones	DILX de type Claes-Jensen (Jensen <i>et al.</i> , 2005)

### ❖ **Les gènes impliqués dans la neurogénèse et la migration neuronale :**

Durant le développement embryonnaire, la neurogénèse et la migration neuronale se mettent en place selon un schéma spatio-temporel strictement contrôlé. Ces processus développementaux dépendent à la fois de mécanismes intracellulaires et d'interactions intercellulaires. Une disruption de ces mécanismes entraîne le plus souvent des anomalies morphologiques cérébrales avec DI (et une épilepsie fréquente), telles que:

- Les lissencéphalies : *ARX*, Xp21.3 (Kitamura *et al.*, 2002), *DCX*, Xq23 (Gleeson *et al.*, 1998).
- Les hétérotopies nodulaires périventriculaires : gène *FLNA*, Xq28 (Parrini *et al.*, 2006).
- Les hydrocéphalies : gène *L1CAM*, Xq28 (Weller et Gartner, 2001).
- Les leucodystrophies : gène *PLP1*, Xq22 (Wolf *et al.*, 1993).

### ❖ **Les gènes impliqués dans des voies métaboliques :**

Dans le cas des maladies métaboliques, la DI est causée soit par une accumulation d'un substrat toxique pour le fonctionnement cérébral, soit par un déficit de production d'un substrat nécessaire au développement cérébral, soit par un déficit énergétique, lors d'une période critique du développement cérébral. Elles sont le plus souvent de transmission autosomique récessive mais quelques-unes sont liées à l'X telles que la maladie de Hunter (gène *IDS*, Xq28), le syndrome de Lesch-Nyhan (gène *HPRT*, Xq26.2), la maladie de Menkes (gène *ATP7A*, Xq21.1), le syndrome Allan-Herndon-Dudley (gène *MCT8*, Xq13.2), ou le déficit en transporteur de la créatine (gène *SLC6A8*, Xq28).

### ❖ **Les gènes impliqués dans la fonction et/ou la structure synaptique :**

Les synapses sont des jonctions intercellulaires asymétriques qui permettent la transmission de signaux depuis les neurones présynaptiques vers les neurones postsynaptiques. La transmission de l'information synaptique est coordonnée par des complexes multi-protéiques distincts localisés dans la zone présynaptique, la fente synaptique et la densité postsynaptique (Biederer *et al.*, 2017) (Figure 22).

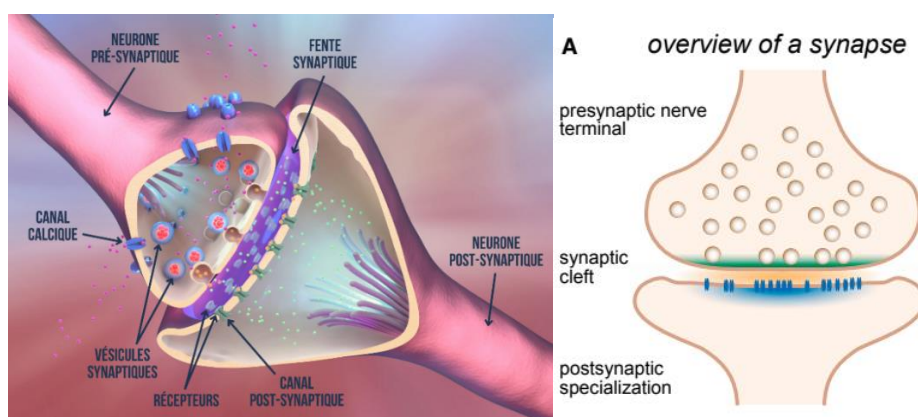


Figure 22: Schémas d'une synapse.

Gauche : neurone présynaptique, fente synaptique, neurone postsynaptique. Droite : densité postsynaptique (bleu), correspondant à un épaissement de la membrane postsynaptique (D'après Biederer *et al.*, 2017).

La majorité des synapses sont dites chimiques car elles fonctionnent par le biais de neurotransmetteurs par opposition aux synapses électriques. Le neurotransmetteur est véhiculé par le biais de vésicules jusqu'à l'extrémité de l'axone du neurone présynaptique (terminaison présynaptique) où il va s'accumuler. L'arrivée de l'influx nerveux ou potentiel d'action entraîne une dépolarisation de la membrane présynaptique responsable d'un afflux massif de calcium par l'ouverture de canaux voltage-dépendants. L'augmentation de la concentration en calcium dans la région présynaptique induit à son tour une cascade de phosphorylations de différentes protéines, telles que les protéines Rab (GTPases) aboutissant à la libération du neurotransmetteur dans la fente synaptique par exocytose. Les molécules de neurotransmetteurs se lient à des récepteurs spécifiques situés sur le neurone postsynaptique induisant une modification de la polarisation de ce dernier avec naissance d'un potentiel d'action au niveau des dendrites (Figure 22).

#### **Fonction présynaptique et cycle vésiculaire :**

De plus en plus de gènes responsables de DILX sont décrits comme étant impliqués dans la fonction présynaptique et le trafic vésiculaire (Tableau 6).

**Tableau 6 : Gènes de DILX impliqués dans le fonctionnement présynaptique**

Gènes	Fonction	Phénotype
<i>GDI1</i> (Xq28)	Régulation du trafic vésiculaire synaptique	DILX type 41 (Bienvenu <i>et al.</i> , 1998)
<i>CASK</i> (Xp11.4)	Régulation de la formation des terminaisons présynaptiques, maintien de la morphologie des épines dendritiques, régulation des canaux ioniques postsynaptiques	DILX avec hypoplasie ponto-cérébelleuses et microcéphalie (Najm <i>et al.</i> , 2008)
<i>SYN1</i> (Xp11.3)	Transmission synaptique, développement neuronal, synaptogénèse, maintien et la plasticité synaptique	DILX, épilepsie, macrocéphalie, troubles du comportement (Garcia <i>et al.</i> , 2004)
<i>RAB39B</i> (Xq28)	Rôle dans le trafic vésiculaire et le développement neuronal	DILX, macrocéphalie, TSA, épilepsie (Giannandrea <i>et al.</i> , 2010)
<i>IL1RAPL1</i> (Xp21.3)	Régulation de la croissance des neurites et de l'exocytose	DILX non-spécifique, TSA (Piton <i>et al.</i> , 2008)
<i>OPHN1</i> (Xq12)	Croissance dendritique, plasticité des synapses excitatrices	DILX, hypoplasie cérébelleuse (Moortgat <i>et al.</i> , 2018)

#### **Architecture dendritique :**

Les épines dendritiques sont de petites protubérances membraneuses qui abritent les composants postsynaptiques essentiels, y compris la densité post-synaptique, le cytosquelette d'actine et de nombreux organelles de soutien. Une altération des protéines régulant ces mécanismes, telles que les protéines de la voie Rho-GTPase, peut entraîner une



DI. C'est le cas du gène *IQSEC2* (Xp11.22), impliqué dans la régulation de l'endocytose vésiculaire et responsable d'une DILX avec possible épilepsie (Mignot *et al.*, 2019), ou du gène *AP1S2* (Xp22.2), impliqué dans le recyclage vésiculaire et la formation du cytosquelette et responsable d'une DILX avec hydrocéphalie et calcifications de la base (Tarpey *et al.*, 2006).

### 🚦 Organisation de la densité post-synaptique :

La densité post-synaptique (DPS) est une zone dense située sous la membrane post-synaptique ayant un rôle clé dans la transmission et la plasticité synaptique puisque comportant la partie intracellulaire des récepteurs et des canaux ioniques, ainsi que de très nombreuses protéines d'adhésion et d'échafaudage (HOMER, SHANK3, PSD-95) (Figure 23). L'intégrité de la DPS est indispensable au bon déroulement de la transduction du signal, qu'il s'agisse d'un courant excitateur (neurones glutamatergiques, récepteurs AMPA ( $\alpha$ -amino-3-hydroxy-5-méthyl-4-isoxazolepropionique acid), NMDA (acide N-méthyl-D-aspartique) ou mGluR5 (récepteur métabotrope du glutamate de type 5)) ou inhibiteur (neurones GABAergiques (acide  $\gamma$ -aminobutyrique)) (Biederer *et al.*, 2017). Les molécules d'adhésion synaptique ont été reconnues comme des composants clés de l'organisation du transfert et du traitement de l'information synaptique (Sudhof, 2018).

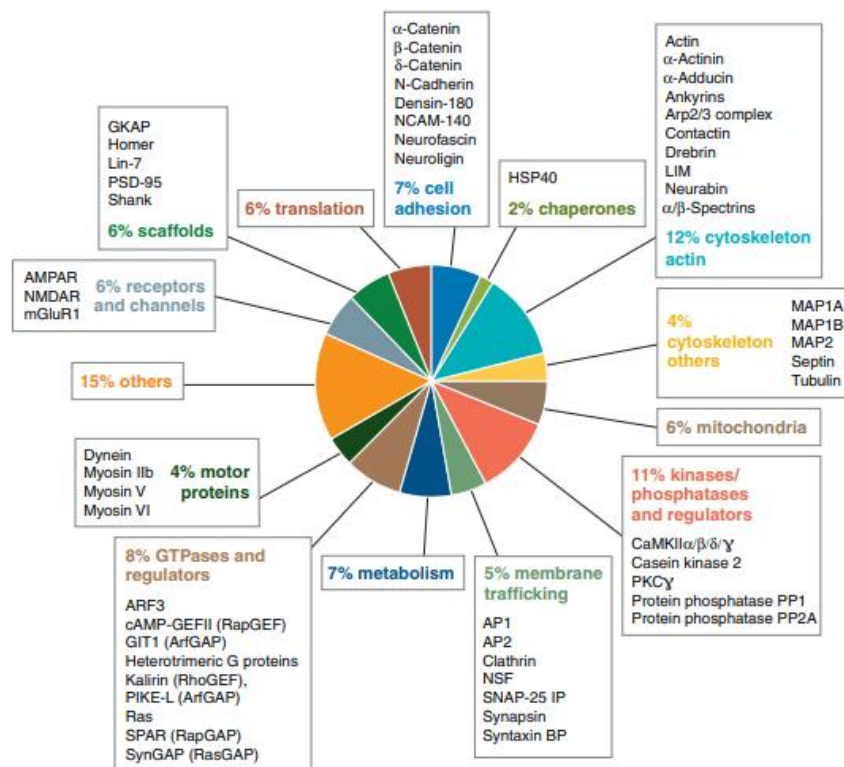


Figure 23: Répartition des protéines de la densité post-synaptique selon leur fonction (D'après Sheng et Hoogenraad, 2007).

Ainsi, les variants pathogènes dans des gènes codant pour des protéines impliquées dans l'organisation de la DPS, telles que les neuroligines, les neurexines et les cadhérines (Figure 24), résultent en une altération de la transduction du signal, avec pour conséquence une DI

pouvant s'associer à une épilepsie (Tableau 7) (Frederic Laumonier *et al.*, 2004; Quartier *et al.*, 2019).

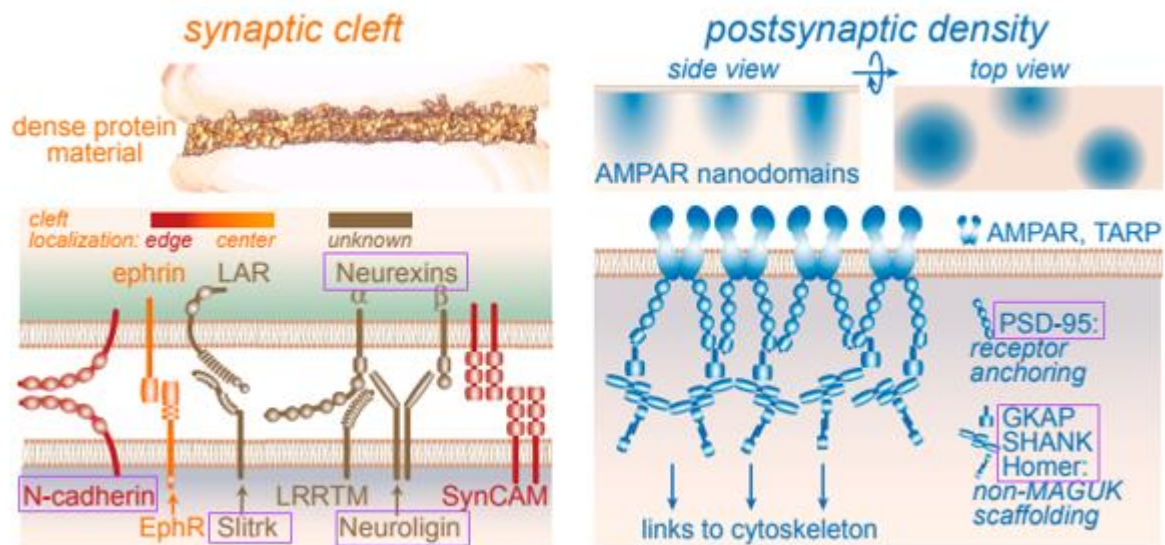


Figure 24: Principales protéines d'adhésion et d'architecture synaptique (Encadrées en rose, d'après Biederer *et al.*, 2017).

**Tableau 7: Gènes de DILX impliqués dans l'organisation de la densité post-synaptique**

Gènes	Fonction	Phénotype
<i>NLGN3</i> ( <i>Xq13.1</i> )	Adhésion cellulaire, interaction avec la neurexine, synaptogenèse	DILX et TSA (Quartier <i>et al.</i> , 2019)
<i>NLGN4</i> ( <i>Xp22.3</i> )	Adhésion cellulaire, interaction avec la neurexine, synaptogenèse	DILX et TSA (Laumonier <i>et al.</i> , 2004)
<i>ARHGEF9</i> ( <i>Xq11.1</i> )	Fonction/localisation des récepteurs au GABA	DILX et épilepsie (Wang <i>et al.</i> , 2018)
<i>DLG3</i> ( <i>Xq13.1</i> )	Fonction synaptique	DILX non-spécifique (Tarpey <i>et al.</i> , 2004)
<i>GRIA3</i> ( <i>Xq25</i> )	Transduction du signal	DILX et épilepsie (Trivisano <i>et al.</i> , 2020)
<i>HUWE1</i> ( <i>Xp11.22</i> )	Développement neuronal et synaptogenèse	DILX type Turner (Moortgat <i>et al.</i> , 2018)
<i>CUL4B</i> ( <i>Xq24</i> )	Turnover des protéines de la densité post-synaptique	DILX type Cabezas ( Tarpey <i>et al.</i> , 2007)

Parmi les gènes d'adhésion synaptique, figure la famille des gènes *SLITRK*, dont le gène *SLITRK2*, qui fait l'objet d'une partie de ce travail de thèse et que je vais donc plus particulièrement développer.

### *b) Les protéines d'adhésion synaptique : exemple de la famille SLITRK*

Les protéines de la famille *SLITRK*, au nombre de six (*SLITRK1*, *SLITRK2*, *SLITRK3*, *SLITRK4*, *SLITRK5*, *SLITRK6*), sont toutes très exprimées dans le cerveau où elles jouent un rôle important dans la fonction synaptique, en étant impliquées dans l'adhésion synaptique et la régulation du développement des synapse excitatrices et inhibitrices (Aruga et Mikoshiba, 2003; Ko, 2012; Um et Ko, 2013; Won *et al.*, 2019). Les six membres de la famille *SLITRK* sont

structurellement similaires, et interagissent avec trois membres de la famille LAR-RPTP (récepteurs tyrosine phosphatase liés à l'antigène commun des leucocytes) que sont PTP $\sigma$ , PTP $\delta$  et LAR, par l'intermédiaire de leurs domaines extra-cellulaires LRR1 et LRR2 (leucin rich repeat domains) pour favoriser l'adhésion transsynaptique et induire la différenciation synaptique (Figure 25) (Takahashi *et al.*, 2012; Um et Ko, 2013; Um *et al.*, 2014).

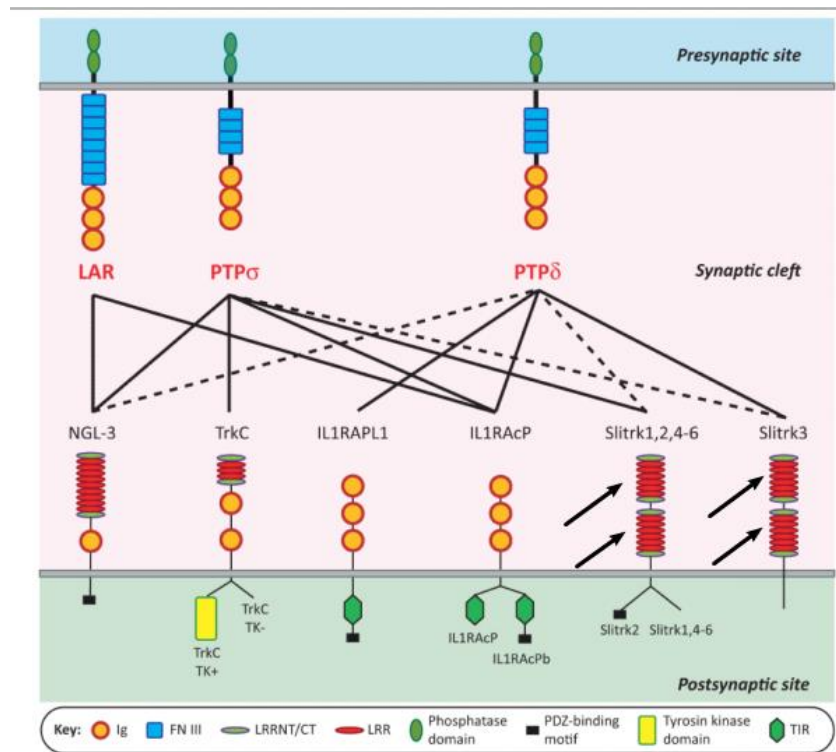


Figure 25: Interactions des protéines SLITRK avec les protéines de la famille RPTP. Leurs domaines de liaison extra-cellulaire LRR1 et LRR2 sont signalés par des flèches (d'après Takahashi et Craig, 2013).

Sur le plan clinique, des variants faux-sens *de novo* dans les gènes *SLITRK1* (13q31.1) et *SLITRK4* (Xq27.3) ont été identifiés chez des patients atteints de troubles neuropsychiatriques tels que la schizophrénie ou le syndrome de Gilles de la Tourette (Kang *et al.*, 2016). Des études fonctionnelles ont permis de démontrer que ces variants ont un impact fonctionnel, caractérisé par une altération de la glycosylation des protéines SLITRK exprimées dans les cellules HEK293T, la rétention de ces dernières dans le réticulum endoplasmique granuleux et le Golgi par défaut d'adressage à la membrane, une abolition de liaison des protéines SLITRK à leur ligand PTP $\delta$  et une altération de la différenciation synaptique (Kang *et al.*, 2016). Par ailleurs, des études antérieures ont montré que SLITRK5 agit comme un corécepteur de TrkB, protéine qui orchestre, *via* sa liaison au BDNF (brain-derived neurotrophic factor), le trafic et la signalisation intra-cellulaire neuronale (Song *et al.*, 2015). Il a également été décrit qu'un variant faux-sens de *SLITRK5*, responsable de troubles obsessionnels compulsifs, altérerait la liaison de SLITRK5 à TrkB (Song *et al.*, 2017).

## Le gène *SLITRK2*

Le gène *SLITRK2* est localisé sur le bras long du chromosome X, en Xq27.3 (Figure 26). Il code pour la protéine SLITRK2, qui, outre les caractéristiques qu'elle partage avec les autres protéines de la même famille (dont une forte expression cérébrale (Figure 27) et des interactions avec les protéines de la famille RTPP (récepteurs tyrosine phosphatase) (Figure 28)), est impliquée dans le contrôle de la différenciation et de la croissance des neurites. Plus particulièrement, elle régule l'organisation des synapses excitatrices des neurones de l'hippocampe par le biais de l'interaction avec 2 protéines clés de l'organisation synaptique que sont SHANK3 et PSD-95 (Han *et al.*, 2019).

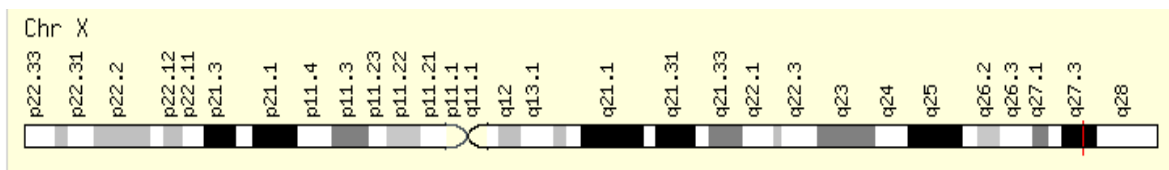


Figure 26: Représentation schématique du locus Xq27.3

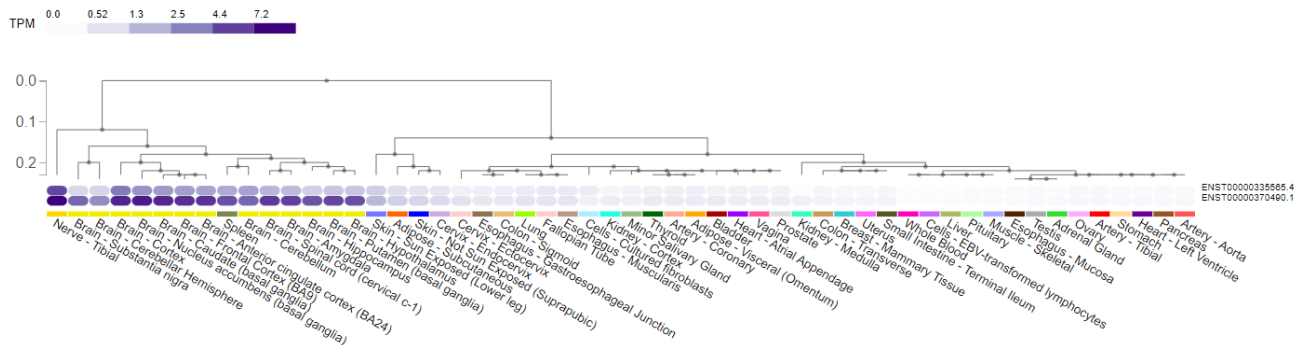


Figure 27: Expression cérébrale de la protéine SLITRK2 chez l'Homme (GTEx Portal)

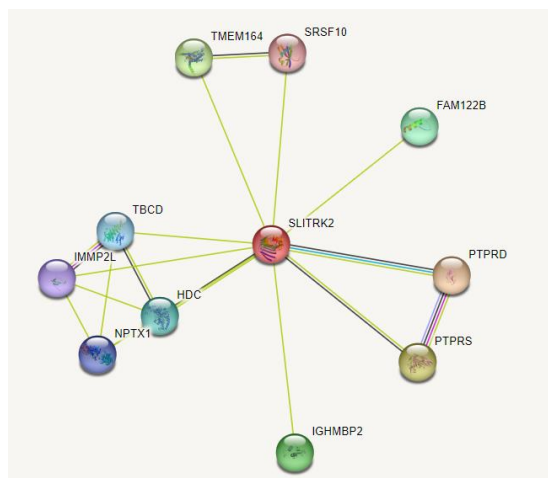


Figure 28: Interactions de la protéine SLITRK2 avec PTP $\delta$  et PTP $\sigma$  (D'après SLITRK2 protein (human) - STRING interaction network (string-db.org)).

Sur le plan clinique, *SLITRK2* a été impliquée dans la maladie bipolaire et la schizophrénie (Kang *et al.*, 2016; Piton *et al.*, 2011). En 2011, Piton *et coll.*, après avoir séquencé 111 gènes de l’X codant pour des protéines synaptiques chez des individus avec TSA ou schizophrénie, ont identifié deux variants faux-sens dans le gène *SLITRK2* à l’état hétérozygote (p.V89M et p.S549F) chez des filles atteintes de schizophrénie et chez leurs frères et sœurs atteints. Ces observations ont fait de *SLITRK2* un gène candidat pour les TND (Piton *et al.*, 2011). Nous avons mis en évidence, par le biais d’un séquençage d’exome (ES), un variant tronquant dans ce gène chez deux frères présentant un TND, et effectué un travail collaboratif dans le but de démontrer son implication dans le phénotype des patients. Ce travail est présenté dans la section Résultats (Partie II).

### 3. Les CNV récurrents sur le chromosome X

L’interprétation des déséquilibres génomiques du chromosome X représente une part importante de l’activité de CGH-array dans le bilan étiologique des TND. En effet, plusieurs CNV pathogènes ont été identifiés sur le chromosome X depuis l’utilisation systématique de la CGH-array dans le bilan étiologique de la DI. Parmi eux, on peut citer la délétion Xp22.3 (Ben Khelifa *et al.*, 2013), la délétion Xp21.1 (syndrome McLeod) (Jung *et al.*, 1993) ou la duplication Xq28 impliquant le gène *MECP2* (Van Esch *et al.*, 2005), syndrome rare auquel je me suis particulièrement intéressée durant ces dernières années et que je vais développer dans ce travail de thèse.

#### **Le syndrome de duplication du gène *MECP2***

Le syndrome de duplication Xq28 impliquant le gène *MECP2*, couramment appelé le syndrome de duplication *MECP2* (*MECP2DS*), a été décrit par Lubs puis Van Esch en 2005 chez des patients de sexe masculin porteurs d'une duplication de la région Xq28 impliquant le gène *MECP2* (et d’autres gènes) et présentant un trouble sévère du développement comprenant un retard de développement global avec hypotonie précoce, une DI sévère avec absence de langage le plus souvent, une épilepsie, des infections récurrentes, une constipation chronique et des mouvements stéréotypés (H. Lubs *et al.*, 1999; Van Esch *et al.*, 2005). Depuis la première description, plus de 300 garçons atteints ont été rapportés, mais seules quelques séries ont été publiées, en particulier en ce qui concerne leur phénotype morphologique et le cas des filles atteintes (Bijlsma *et al.*, 2012; El Chehadeh *et al.*, 2017; Fieremans *et al.*, 2014; Novara *et al.*, 2014; Peters *et al.*, 2019; San Antonio-Arce *et al.*, 2016; Schwoerer *et al.*, 2014; Shimada *et al.*, 2013; Ta *et al.*, 2022). Bien que la taille des duplications diffère d'un patient à l'autre, et malgré l'implication de plusieurs autres gènes contenus dans ces duplications et exprimés dans le cerveau, il a été admis que l'augmentation du dosage de la protéine *MECP2* au niveau cérébral était suffisante pour expliquer le phénotype central de cette affection. A ce jour,

aucune corrélation claire n'a pu être établie entre la taille des duplications, leur contenu en gènes et le phénotype des patients, bien que les individus avec de très grandes duplications visibles en cytogénétique conventionnelle semblent présenter plus fréquemment une microcéphalie, un retard statural et des anomalies génitales (Sanlaville *et al.*, 2005).

### 3. Le diagnostic de la DILX : que reste-t-il à découvrir?

Bien que les techniques de séquençage à haut débit (NGS, next generation sequencing) aient permis d'importants progrès dans le diagnostic de la DILX, qu'il s'agisse de séquençage ciblé de gènes de l'X (Ibarluzea *et al.*, 2020; Tzschach *et al.*, 2015) ou d'« X-exome » (Hu *et al.*, 2016; Niranjana *et al.*, 2015; Philips *et al.*, 2014), on estime qu'environ 21% des syndromes avec DILX restent irrésolus sur le plan moléculaire (Schwartz *et al.*, 2023).

En 2013, le nombre de gènes impliqués dans la DILX étaient de 106 (Piton *et al.*, 2013) contre 127 actuellement (Leitao *et al.*, 2022), dont 21 ont été identifiés depuis 2017, avec environ 200 syndromes recensés avec DILX. Depuis, ce nombre semble relativement stable voire en baisse (Bamshad *et al.*, 2019; Schwartz *et al.*, 2023) (Figure 29).

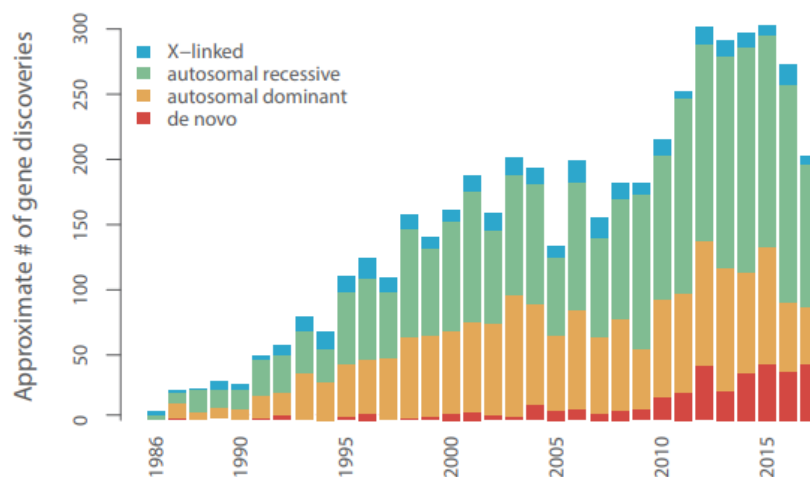


Figure 29: Nouveaux gènes de maladies mendéliennes identifiés par an. (D'après Bamshad *et al.*, 2019).

En effet, le consortium Deciphering Developmental Disorders (DDD) a récemment estimé que les TND liés au chromosome X affectent les hommes et les femmes de la même manière et représentent 6% et 6,9% des TND, respectivement. Cependant, malgré la taille importante de la cohorte (11044 individus atteints), cette étude n'a pas permis d'identifier de nouveaux gènes associés à la DILX (Martin *et al.*, 2021). Bien que les auteurs se soient ici focalisés sur les variants *de novo*, cela confirme que des mécanismes mutationnels encore non ou peu décrits sur l'X restent probablement à découvrir : expansions, variants d'épissage, insertions d'éléments mobiles (rétrotransposons), variants de structure, anomalies épigénétiques

(méthylation ou modification post-traductionnelles d'histones), altérations de TAD (domaines topologiquement associés), etc.

## 4. Les défis de la DILX

L'identification de la cause moléculaire de la DILX dans les familles concernées est associée à de nombreux *challenges* que sont :

- La difficulté d'identifier des familles candidates avec forte suspicion d'une transmission liée à l'X : en effet, lorsqu'une seule génération de garçons est atteinte, comme une fratrie, il est possible qu'il s'agisse d'une transmission autosomique récessive. Ainsi, il est nécessaire d'étudier de grandes familles avec plusieurs garçons atteints sur plusieurs générations, et de surcroît, avec un phénotype similaire (Hu *et al.*, 2016; Tarpey *et al.*, 2009). Par ailleurs, nous avons vu précédemment qu'une DI atteignant à la fois des garçons et des filles dans une même famille peut tout à fait correspondre à une maladie liée à l'X, sous réserve toutefois qu'il n'y ait pas de transmission père fils ; mais ces familles étant rarement très grandes, ce critère d'exclusion ne figure pas souvent sur les arbres.

- L'interprétation d'un variant candidat sur le chromosome X chez des cas sporadiques ou dans des familles où la ségrégation familiale n'est pas possible reste compliquée (Leitao *et al.*, 2022), notamment en ce qui concerne les variants faux-sens. En effet, le caractère hérité d'une mère asymptomatique n'est pas un argument qui aide à l'interprétation car il peut exister un biais d'inactivation chez cette dernière, et par ailleurs, nous avons vu que l'étude de l'XCI n'est pas toujours informative. De la même manière, l'identification d'un variant pathogène dans un gène de DILX récessive chez une fille atteinte ne permet pas de conclure de manière formelle quant à son implication dans le phénotype, la patiente pouvant être conductrice asymptomatique concernant ce variant, sa DI étant alors liée à un autre variant pathogène non encore identifié.

- Pour les mêmes raisons, l'interprétation d'un CNV sur l'X est parfois délicate. De plus, les bases de données de population générale sont moins nombreuses et le sexe du sujet rapporté n'est pas toujours indiqué. Dans ces situations, une étude familiale étendue est le seul recours pour établir une corrélation génotype-phénotype chez les individus masculins porteurs. Elle est cependant parfois peu informative (petite fratrie, enfant unique) ou difficile en pratique (famille éloignée, perdue de vue, apparenté décédé...) ne permettant pas de conclure à la pathogénicité ou non du CNV retrouvé.

- La présence de variants candidats (parfois même publiés comme étant pathogènes) dans des gènes de l'X au sein de grandes bases de données de variants issues de la population générale (donc en théorie, chez des individus non atteints de DI) peut poser question (Piton *et al.*, 2013).

- La question non-encore résolue de la présence d'un phénotype chez des filles dans certaines maladies rares avec DILX (El Chehadeh *et al.*, 2017).

### III. Le séquençage haut débit dans le diagnostic de la DI

#### 1. Historique de l'évolution des techniques

L'âge « d'or » de la description clinique des syndromes rares d'origine génétique correspond aux années 1970, durant lesquelles les techniques de séquençage de 1<sup>ère</sup> génération sont nées et ont valu un prix Nobel à leurs inventeurs, Maxam et Gilbert, et Sanger (Maxam et Gilbert, 1977; Sanger *et al.*, 1977).

La technique Sanger, une fois automatisée, utilisait des ddNTP (didésoxyribonucléotides) fluorescents et l'électrophorèse capillaire pour obtenir une séquence nucléotidique sous la forme d'électrophorégrammes (Smith *et al.*, 1986) (Figure 30).

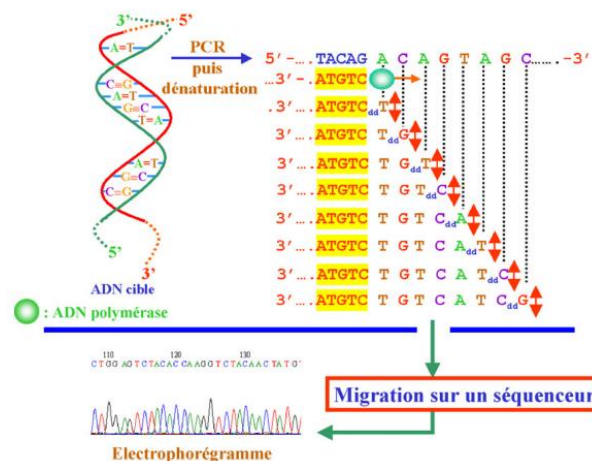


Figure 30: Principes du séquençage selon la méthode de Sanger (D'après Lamoril *et al.*, 2008).

S'en sont suivis des progrès techniques et méthodologiques qui ont favorisé la découverte de gènes comme la technique du clonage positionnel en 1986, avant laquelle une quarantaine de gènes étaient connus (Chong *et al.*, 2015), aidé par le développement de la cartographie génétique au début des années 1990 et la connaissance progressive de la séquence du génome, acquise grâce au projet Génome humain (1990-2001) (Lander *et al.*, 2001).

Entre 1986 et 1997, une accélération a été observée concernant le nombre de maladies mendéliennes (MM) décrites et le nombre de nouveaux gènes correspondants identifiés, qui se sont vus multipliés par 3/an et par 10/an, respectivement. Entre 1998 et 2010, le taux de découverte de gènes pour les MM s'est stabilisé pour ré-augmenter considérablement après 2010, date qui correspond au début du déploiement du séquençage de l'exome (ES) (Ng,



Buckingham, *et al.*, 2010; Vissers *et al.*, 2010), jusqu'en 2015, représentant 36 % (1268/3549) de toutes les découvertes de gènes mendéliens (Bamshad *et al.*, 2019) (Figure 31).

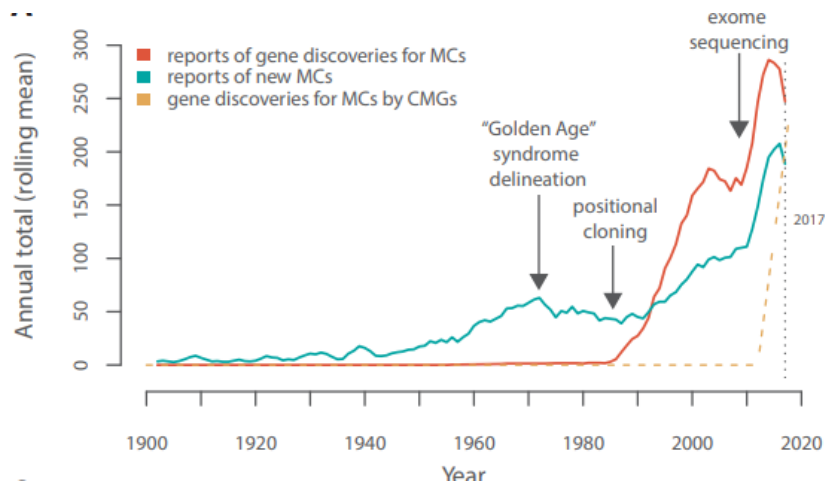


Figure 31: Description de gènes de maladies mendéliennes au fil du temps  
Les gènes découverts sont en rouge et les maladies mendéliennes en bleu (D'après Bamshad *et al.*, 2019).

En parallèle de ces découvertes, le développement d'autres techniques de séquençage haut débit s'est également accéléré ces dernières années notamment concernant l'épigénome (Methyl-Seq) (Brunner *et al.*, 2009), le transcriptome (RNA-Seq) (Ozsolak et Milos, 2011), la régulation génique (Chip-Seq, *Chromatin Immunoprecipitation Sequencing*) (Johnson *et al.*, 2007), le protéome, le lipidome etc. (Figure 32). Ces dernières contribuent considérablement à la description de nouveaux gènes de maladies rares, par les données fonctionnelles complémentaires qu'elles apportent.

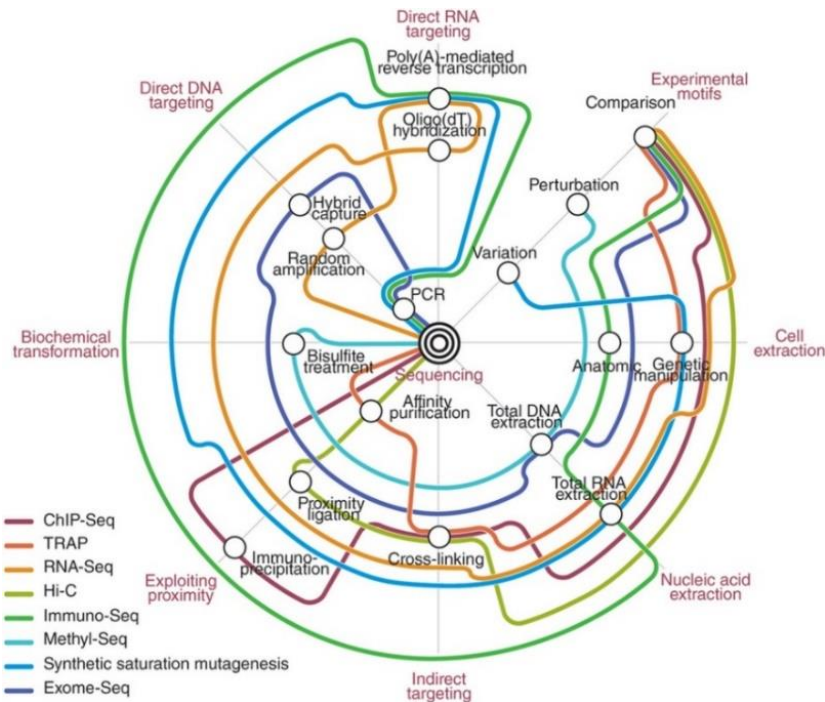


Figure 32: « Plan de métro » de la technologie de séquençage haut débit  
(D'après Shendure et Lieberman Aiden, 2012).

## 2. Le principe des technologies de séquençage haut débit (NGS)

Contrairement au séquençage de 1<sup>ère</sup> génération, dont le rendement en nombre de bases est limité, les séquençages de 2<sup>ème</sup> et 3<sup>ème</sup> générations sont basés sur une parallélisation du séquençage de millions de molécules, fragmentée ou non, et d'une taille plus ou moins longue. Le produit de ce séquençage est appelé *read*. Cette technologie est avantageuse de par la masse de *reads* qu'elle produit et son faible coût par base séquencée (Ng *et al.*, 2009; Metzker, 2010). Le procédé de séquençage d'exome est précédé d'une étape d'enrichissement au cours de laquelle les exons sont capturés par hybridation (technique la plus couramment utilisée). Cette technique permet donc de séquencer uniquement les exons connus et ciblés par les sondes.

Le séquençage haut débit d'exome (ES) et de génome (WS) est subdivisé en plusieurs étapes une fois l'extraction d'ADN réalisée: la préparation des bibliothèques, l'amplification de l'ADN (et la capture pour l'exome), le séquençage de l'ADN et l'analyse bioinformatique (Figure 33).

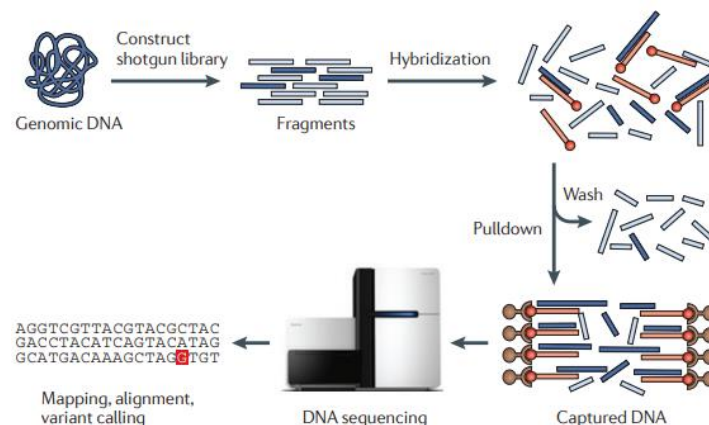


Figure 33: Workflow pour le séquençage d'exome (D'après Bamshad *et al.*, 2011).

Le séquençage de 2<sup>ème</sup> génération réalisé dans le cadre de ce travail utilise la technologie « *short read* » d'Illumina (*paired-end* 2 x 150 b, NovaSeq 6000, avec une couverture de 30 X) (Bentley *et al.*, 2008).

L'analyse bioinformatique est indispensable au traitement des données générées par le séquenceur et permettra *in fine* l'interprétation de ces dernières, qui sont massives et complexes. Ceci nécessite l'utilisation d'un domaine de compétence en plein essor qu'est la bioinformatique, dont un des rôles cruciaux sera de convertir les données brutes issues du séquençage en des fichiers « lisibles et interprétables » par le biologiste moléculaire qui interprétera et validera les résultats.

Le processus bioinformatique d'analyse des données issues du séquençage haut débit se décompose en trois principales étapes : **l'analyse primaire**, qui génère les séquences, **l'analyse secondaire**, qui mène à l'identification des variations par rapport à la séquence de référence, et **l'analyse tertiaire**, qui correspond à l'annotation des variants permettant leur interprétation (Figure 34).

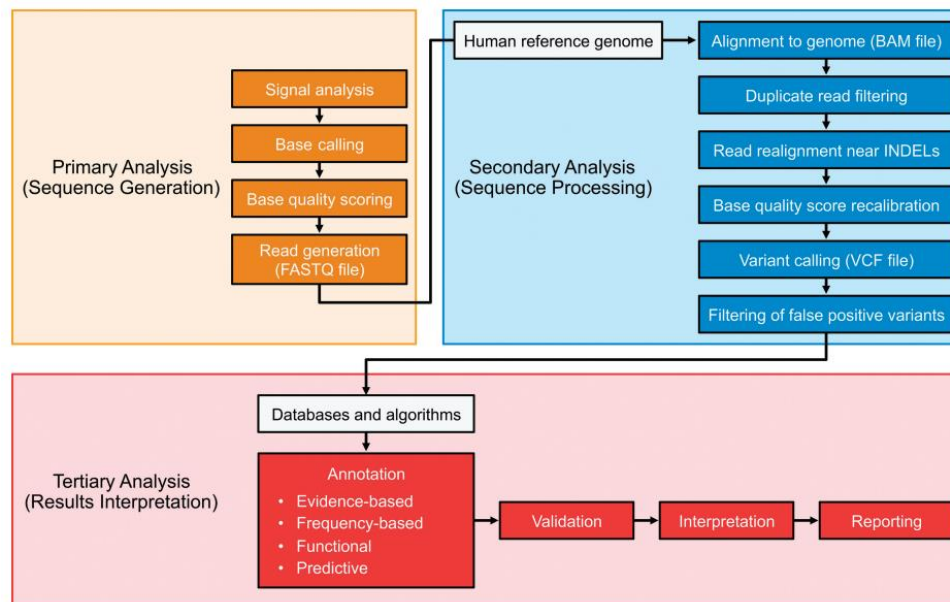


Figure 34: Principales étapes de l'analyse bioinformatique des données de NGS (D'après (Oliver *et al.*, 2015).

### a) L'analyse primaire et secondaire

**Le base calling** permet de convertir le signal brut du séquenceur en séquences, et d'estimer la qualité des bases obtenues. Un fichier FASTQ (FASTA Quality) est alors généré (Cock *et al.*, 2010); il est composé d'un bloc de trois lignes pour chaque séquence : l'identifiant de la séquence, la séquence nucléotidique et le score de qualité pour chaque base (score Phred).

**Le trimming** des *reads* permet de supprimer les séquences d'amorces et les bases « de mauvaise qualité ». Les séquences sont ensuite **alignées** sur une séquence de référence et des **fichiers BAM** (Binary Alignment Map) sont générés. Ils peuvent être lus directement avec des outils informatiques tels que IGV (Integrative Genome Viewer) (Thorvaldsdottir *et al.*, 2013), ce qui permet de visualiser à la fois la longueur des *reads*, leur sens, la couverture et la profondeur, ainsi que les variations par rapport à la séquence de référence (Robinson *et al.*, 2011; Thorvaldsdottir *et al.*, 2013) (Figure 35). La *longueur* de lecture correspond au nombre de bases successives obtenues par une lecture. La *couverture* est le pourcentage de bases séquencées pour une cible initialement choisie (ex. un gène, un exon en particulier...) à une

profondeur donnée. La *profondeur* correspond au nombre de lectures à une position donnée (20X correspondant à une profondeur de 20 lectures à une position donnée).

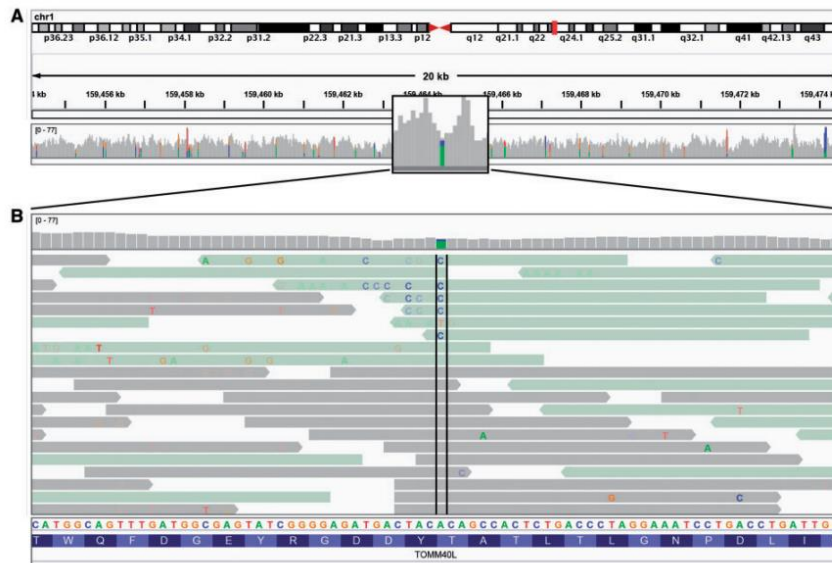


Figure 35: Aligement des reads (fichier BAM) par le logiciel IGV

(A) Les reads apparaissent sous forme d'un graphique de couverture. (B) Le sens des reads apparaît en vert ou gris. Les variations par rapport à la séquence de référence sont directement visualisées (D'après Thorvaldsdottir *et al.*, 2013).

Un certain nombre d'étapes bioinformatiques telles que l'élimination des reads dupliqués, multi-alignés, la calibration des bases et le réaligement autour des *indel* (*postprocessing*) permettent d'améliorer encore la qualité des données et assurent une meilleure détection des variants.

Le **variant calling** a pour but d'identifier les variants détectés et peut être réalisé par des outils comme GATK (Genome Analysis ToolKit) (DePristo *et al.*, 2011; Van der Auwera et O'Connor, 2020), SAMtools (Danecek *et al.*, 2021) ou VarScan (Koboldt *et al.*, 2009). Les variations structurales sont détectées en utilisant des outils spécifiques tels que Smoove (<https://github.com/brentp/smoove>), Manta (X. Chen *et al.*, 2016) et CNVpytor (Suvakov *et al.*, 2021).

Le format de fichier obtenu est le **VCF (Variant Call Format)**. Celui-ci possède huit colonnes comportant les informations suivantes : le chromosome sur lequel le variant est appelé (CHROM), la position du variant (POS), son identifiant (ID), la référence à la position donnée (REF), la base alternative (ALT), la qualité moyenne (QUAL), les critères de passage du filtre prédéfini (PASS). Les résultats sont visualisés généralement *via* un tableur de type Excel qui permet de trier et filtrer les variants (Danecek *et al.*, 2011) (Figure 36).

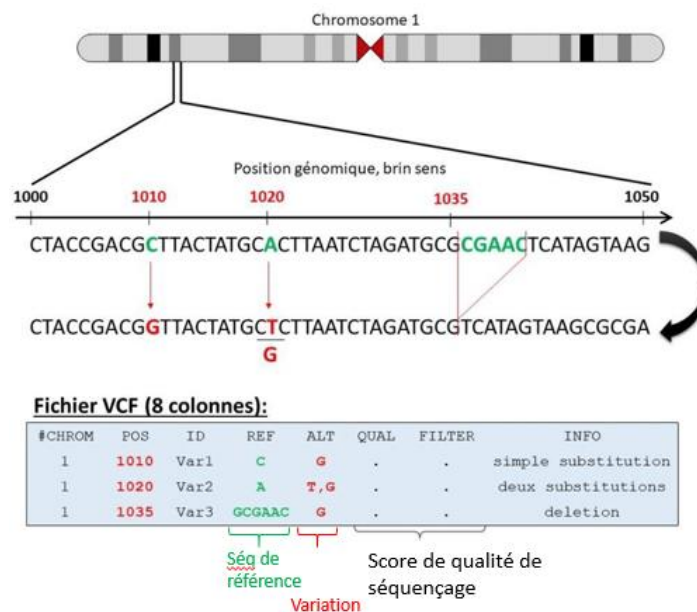


Figure 36: Schématisation des informations contenues dans un fichier VCF (colonnes).

### b) L'analyse tertiaire : l'annotation, filtre et classification des variants

L'analyse tertiaire permet l'annotation des variants issus du fichier VCF après interrogation de plusieurs bases de données et outils de prédiction *in silico* (cf chapitre « Aides à l'interprétation »). Plusieurs outils permettent de réaliser cette annotation, et dans le cadre de ma thèse nous avons utilisé l'outil développé au laboratoire, à savoir le logiciel VaRank (Geoffroy *et al.*, 2015). VaRank permet, en combinaison avec le logiciel Alamut Batch (<https://www.sophiagenetics.com/platform/alamut-visual-plus/>) de compiler de très nombreuses informations pour chaque variant identifié (SNV et *indel*) telles que le statut du variant (bénin, probablement bénin, de signification incertaine, probablement pathogène, pathogène) dans les bases de données reliant les variants au phénotype telles que ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/> (Landrum *et al.*, 2018, 2020)), sa fréquence dans la population générale, sa position génomique/protéique/génique, les scores de prédiction de pathogénicité qui lui sont reliés etc. Contrairement aux SNV et *indel*, les variants de structure sont annotés par un autre logiciel, AnnotSV, également développé au laboratoire (Geoffroy *et al.*, 2018).

Une fois annotés, les variants peuvent être **filtrés** selon divers critères, comme le fait d'éliminer les variants fréquents dans les bases de données publiques, incluant gnomAD (Genome Aggregation Database) (Lek *et al.*, 2016) ou 1000 Genomes (Auton *et al.*, 2015), les variants rapportés comme bénins, les variants avec un faible nombre de lectures etc... Cette étape de filtre est primordiale car elle permet de diminuer significativement le nombre de variants d'intérêt à interpréter (Figure 37).

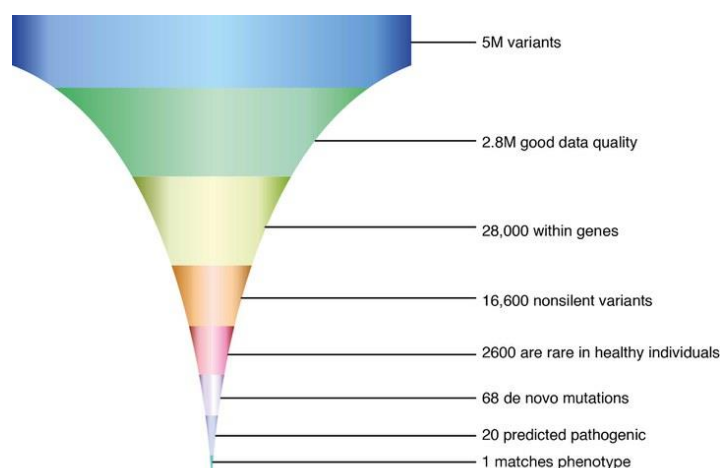


Figure 37: Schématisation du nombre de variants restants après application de filtres (D'après Garvan institute of medical research)

L'ultime phase correspond à l'**interprétation des variants**, qui permet de confronter les informations relatives aux variants au phénotype et d'aboutir *in fine* à leur **classification**. Cette étape correspond à l'étude de l'effet potentiel des variants et des gènes concernés, en s'aidant des informations existantes sur leur implication en pathologie humaine, leur fonction, leur profil d'expression (par ex son niveau d'expression au niveau cérébral en ce qui concerne la DI ou au niveau des lymphocytes en cas d'étude de transcrit), les éventuelles interactions avec des protéines d'intérêt en lien avec la pathologie étudiée etc. Cependant, plusieurs dizaines de milliers de variations sont détectées dans un exome et plusieurs millions dans un génome (Lupski *et al.*, 2010; Wright *et al.*, 2015) (Figure 38). L'enjeu est donc de parvenir à identifier LA ou LES variation(s) dans LE ou LES gène(s) responsable(s) du phénotype du patient. Pour cela, il est nécessaire de faire appel à différents outils d'aide à l'interprétation.

SNP Type	No. of SNPs
Nongene	2,255,102
Gene	1,165,204
Intron	1,064,655
Promoter	60,075
3' UTR	16,350
5' UTR	3,517
Splice regulatory site	2,089
Splice site	112
Synonymous	9,337
Stop→stop	17
Nonsynonymous	9,069
Stop→gain	121
Stop→loss	27
Total	3,420,306

Figure 38: Nombre de SNP identifiés par WGS chez un individu.

SNP : polymorphisme d'un seul nucléotide présent à une fréquence supérieure à 1% dans une population donnée (D'après Lupski *et al.*, 2010).

### 3. Les aides à l'interprétation des variants

#### 1. Les bases de données de variants et d'informations sur les gènes

L'identification de très nombreux variants chez un individu nécessite d'appliquer différents filtres, dont leur fréquence dans la population générale, dans le but de les trier en fonction de leur pertinence clinico-biologique. Pour cela, les laboratoires font appel, de manière automatisée, à l'utilisation de plusieurs bases de données, telles que:

- **Bases de données de variants en population générale :**
  - **1000 genomes** (1000 Genomes Project Consortium, 2010) (Auton *et al.*, 2015) (<http://www.1000genomes.org/>): recense les données génomiques de 2504 individus provenant de 26 populations différentes.
  - **gnomAD** (Genome Aggregation Database) (Lek *et al.*, 2016; Karczewski *et al.*, 2019): regroupe environ 125 000 données d'ES et 15 000 données de WGS, réalisés sur individus non apparentés et ne comportant théoriquement pas de TND sévère.
  - **DGV** (Database of Genomic Variant) (<http://dgv.tcag.ca/dgv/app/home>) regroupant les anomalies du nombre de copies identifiées dans de larges études de populations.
- **Bases de données de variants associés à un phénotype :**
  - **DECIPHER** : Base de données de remaniements chromosomiques submicroscopiques (délétions, duplications et inversions) et phénotypiques chez l'Homme, utilisant les ressources d'Ensembl (Ensembl) (<https://www.deciphergenomics.org/>).
  - **ClinVar** : recense des variants identifiés chez des patients atteints de maladies rares, en précisant leur pertinence et leur caractère pathogène, polymorphique ou incertain (<https://www.ncbi.nlm.nih.gov/clinvar/>) (Landrum *et al.*, 2020).
  - **LOVD** base de données gène spécifique dont les données sont validées par des experts du gène (<https://www.lovd.nl/>)
  - **denovo-DB** : recensement de variants *de novo* (<https://denovo-db.gs.washington.edu/denovo-db/>).
  - **HGMDpro** (Human Gene Mutation Database): compilation manuellement validée de variants identifiés dans la littérature scientifique par leur association à des maladies génétiques (Stenson *et al.*, 2017) (<https://www.hgmd.cf.ac.uk/>).
- **Outils d'interprétation et de visualisation de variants :**
  - **VarSome** (<https://varsome.com/>)
  - **Alamut Visual** (<https://www.sophiagenetics.com/platform/alamut-visual-plus/>)
- **Autres bases de données :**

- **OMIM** (Online Mendelian Inheritance in Man) : recense plus de 3800 maladies secondaires à des variants pathogènes dans plus de 3400 gènes (<http://www.ncbi.nlm.nih.gov/omim/>) (Amberger *et al.*, 2015).
- **SysNDD** : recense les gènes associés aux TND (<https://sysnnd.dbmr.unibe.ch/>)
- **Genatlas** : cartographie des gènes et les maladies génétiques (Frezal, 1998).
- **STRING** : cartographie des réseaux protéiques ([string-db.org](http://string-db.org)) (Snel *et al.*, 2000).
- **dbRIP** : base de données d'insertion d'éléments mobiles polymorphiques humain (<http://dbrip.brocku.ca/>) (J. Wang *et al.*, 2006).
- **GeneHancer** : base de données sur les éléments régulateurs (*enhancers* et promoteurs) (Fishilevich *et al.*, 2017)
- **UniProt** : données de séquences et de fonction protéiques (UniProt) (uniprot consortium 2023, nucleic acid research 2023)
- **GENCODE** : annotation du génome humain et murin (Frankish *et al.*, 2019)

## 2. Les différents outils de prédiction *in silico*

Si l'effet d'un variant *frameshift* ou non-sens localisé avant le dernier exon peut généralement être considéré comme étant délétère, l'effet d'autres variants comme les faux-sens ou les variants localisés à proximité d'un site d'épissage, requière l'utilisation de programmes pour prédire leur effet potentiel sur la protéine. Ainsi, dans le but d'aider à l'interprétation de ces nombreux variants, plusieurs outils de prédiction de l'effet fonctionnel de ces derniers ont été développés et sont devenus indispensables pour interpréter les données de séquençage haut débit. On peut distinguer principalement deux catégories d'outils : ceux qui prédisent si un variant faux-sens est préjudiciable à la fonction ou à la structure de la protéine concernée, et ceux qui prédisent les éventuels effets sur l'épissage.

### a) Les outils de prédiction des effets des variants faux-sens

Une variation faux-sens est une altération ponctuelle d'un nucléotide qui aboutit à un changement d'acide aminé (AA) lors de la traduction. Ce changement d'AA peut avoir ou non des conséquences sur la fonction et la structure de la protéine produite, notamment s'il est localisé dans un domaine fonctionnel clé et qu'il entraîne une modification de la structure tridimensionnelle de la protéine (domaine de liaison protéique pour un récepteur ou de liaison à l'ADN pour un facteur de transcription par exemple), et ainsi avoir des effets délétères (expression phénotypique) ou non (Mattioli *et al.*, 2019).

Par ailleurs, un AA conservé au cours de l'évolution a de grandes chances d'appartenir à un élément fonctionnel jouant un rôle important dans la fonction de la protéine. Les procédés



d'alignement multiple permettent de visualiser la conservation des résidus entre les espèces (Tavtigian *et al.*, 2008) (Figure 39).



Figure 39: Alignement des acides aminés de la protéine MLH1. A partir de l'acide aminé K140 humain (D'après Tavtigian *et al.*, 2008).

Il est donc important de pouvoir statuer sur l'éventuelle modification de fonction ou de structure de la protéine traduite. Plusieurs outils de prédiction sont couramment utilisés pour l'interprétation des variants faux-sens (Tableau 8), tels que :

- SIFT (Sorting Intolerant From Tolerant): se base sur ces alignements multiples et la **conservation des AA** à des positions précises pour prédire le caractère délétère d'un variant faux-sens (Ng et Henikoff, 2003; Kumar *et al.*, 2009). Il tient compte des **propriétés physico-chimiques des AA** telles que la polarité et l'encombrement stérique. La conséquence prédite d'une variation impliquant des AA très différents d'un point de vue physico-chimique (AA à caractère hydrophobe *versus* AA chargé ou polaire) sera plus délétère que si elle concerne des AA proches.
- PolyPhen-2 (Polymorphism Phenotyping) : utilise des données d'alignements multiples, des annotations connues de séquences (ex. site de fixation d'un cofacteur) et des informations structurales de manière à prédire si la substitution engendre un effet sur la fonction de la protéine (Adzhubei *et al.*, 2010).
- AGVGD (Align-GVGD): programme en ligne qui combine les **caractéristiques biophysiques** des AA et **des alignements de séquences** de protéines pour prédire où se situent les substitutions faux-sens dans les gènes d'intérêt.
- Grantham : se base sur les **propriétés physico-chimiques** et la **distance évolutive** des AA (composition, polarité et volume moléculaire). Les scores de distance vont de 5 à 215. Plus le score est faible plus l'écart séparant les AA est faible (et les propriétés physico-chimiques de la protéine sont respectées), et plus il est élevé plus l'écart est grand, ce qui est en faveur d'un effet délétère (Grantham, 1974).
- CADD (Combined Annotation Dependent Depletion): propose une intégration de multiples annotations fonctionnelles (**conservation en séquence, signaux de modifications d'histone, prédictions de changement d'AA, îlots CpG**, etc.) et des prédictions précédemment listées pour générer un score global d'effet délétère potentiel

(Kircher *et al.*, 2014; Rentzsch *et al.*, 2019). Le score CADD est compris entre 0 et 35. Un score CADD de 20 signifie que le variant fait partie des 1% des variants prédits les plus délétères du génome, un score de 30 des 0,1%. On considèrera généralement un effet délétère peu probable pour un score CADD inférieur à 20 et un effet délétère très probable pour un score CADD supérieur à 25.

- MutationTaster : intègre des informations provenant de différentes bases de données biomédicales et utilise des outils d'analyse concernant la conservation évolutive, les changements de sites d'épissage, la perte de caractéristiques protéiques et les changements susceptibles d'affecter la quantité d'ARNm (Schwarz *et al.*, 2010).

**Tableau 8: Les principaux outils de prédiction *in silico* des variants faux sens**

Outil de prédiction	Site web
SIFT	<a href="https://sift.com">https://sift.com</a>
POLYPHEN	<a href="http://genetics.bwh.harvard.edu/pph2/">http://genetics.bwh.harvard.edu/pph2/</a>
CADD	<a href="https://cadd.gs.washington.edu/">https://cadd.gs.washington.edu/</a>
MutationTaster	<a href="https://www.mutationtaster.org">https://www.mutationtaster.org</a>
PhastCons / PhyloP	<a href="http://compgen.bscb.cornell.edu/phast/">http://compgen.bscb.cornell.edu/phast/</a>
AGVGD	<a href="http://agvgd.hci.utah.edu">http://agvgd.hci.utah.edu</a>

### *b) Les outils de prédiction sur l'épissage*

L'épissage des ARN pré-messagers est la suppression d'introns réalisée par un complexe appelé spliceosome qui reconnaît les séquences bien spécifiques localisées aux frontières intron-exon (Brody et Abelson, 1985; Frenthewey et Keller, 1985). C'est l'établissement d'une continuité entre les exons consécutifs qui permet la lecture de la séquence servant de matrice à la machinerie de synthèse protéique (Dujardin *et al.*, 2016).

Plusieurs déterminants de séquence, présents dans les introns, permettent à la machinerie d'épissage de « définir » les exons. Il s'agit des sites 5', ou sites donneurs d'épissage, des sites 3', ou sites accepteurs d'épissage, du point de branchement (PB) vers l'extrémité 3' de l'intron, et d'une étendue de pyrimidines située entre le PB et le site 3'. Plusieurs types de séquences activatrices (ESE : exonic splice enhancer, ISE : intronic splice enhancer) ou inhibitrices (ESS : exonic splice silencer, ISS : intronic splice silencer) participent également à la reconnaissance de l'exon en interagissant avec des protéines riches en sérines et en arginines (SR) ou avec des protéines hnRNP (heterogenous nuclear ribonucleoproteins) qui, respectivement, favorisent ou réfrènt l'inclusion normale des exons (Caceres et Kornblihtt, 2002) (Figure 40).

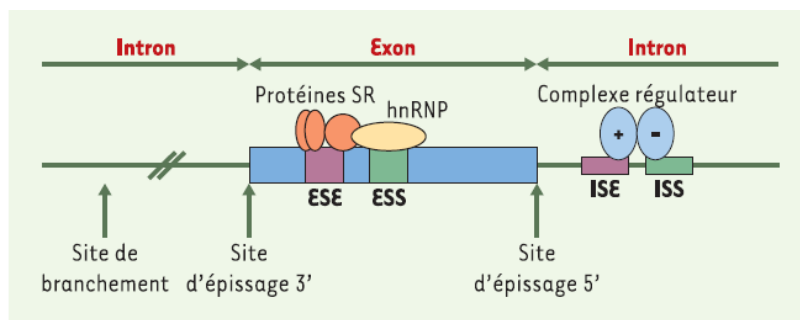


Figure 40: Organisation d'une unité d'épissage.  
(D'après Corcos et Solier, 2005).

Une importante proportion des variants pathogènes identifiés chez l'homme sont susceptibles d'affecter l'épissage, estimée actuellement à environ un tiers (allant de 3,8% à 65%) de tous les variants pathogènes (Ars *et al.*, 2000; Cheung *et al.*, 2019; Julien *et al.*, 2016; Lim *et al.*, 2011; Mueller *et al.*, 2015), ayant fait émettre l'hypothèse que ce mécanisme était la cause majoritaire des maladies héréditaires (Lopez-Bigas *et al.*, 2005; Sterne-Weiler *et al.*, 2011). A titre d'exemple, Soemedi *et coll.* étudièrent en 2017 4964 variants ponctuels exoniques pathogènes et observèrent que 10% d'entre eux (513/4964) altéraient l'épissage *in vitro* et *in vivo* (Soemedi *et al.*, 2017).

Les variants impliquant les sites « canoniques » d'épissage donneur (+1/+2 GT, en 5') et accepteur (-2/-1 AG, en 3'), mais également le *tract* de pyrimidines, le PB et les séquences régulatrices renforçant (ESE) ou réprimant l'épissage (ESS), sont susceptibles d'entraîner une altération de l'épissage et donc des modifications importantes sur la protéine en aval.

Ces variants peuvent être responsables de diverses conséquences sur l'épissage incluant (Figure 41) (Wimmer *et al.*, 2007; Lord et Baralle, 2021; Leman *et al.*, 2022):

- Un saut d'exon par la perte de reconnaissance du site d'épissage,
- L'utilisation d'un nouveau site d'épissage par altération d'un site préexistant et utilisation en remplacement d'un autre site ou l'activation d'un nouveau site par le variant,
- L'apparition d'un nouvel exon par l'activation de l'un de ses sites d'épissage,
- La rétention complète de l'intron par altération de l'un des motifs canoniques,
- L'apparition d'un nouveau site d'épissage, *de novo* si le variant crée un site d'épissage qui n'existait pas auparavant, ou cryptique s'il renforce un site préexistant.

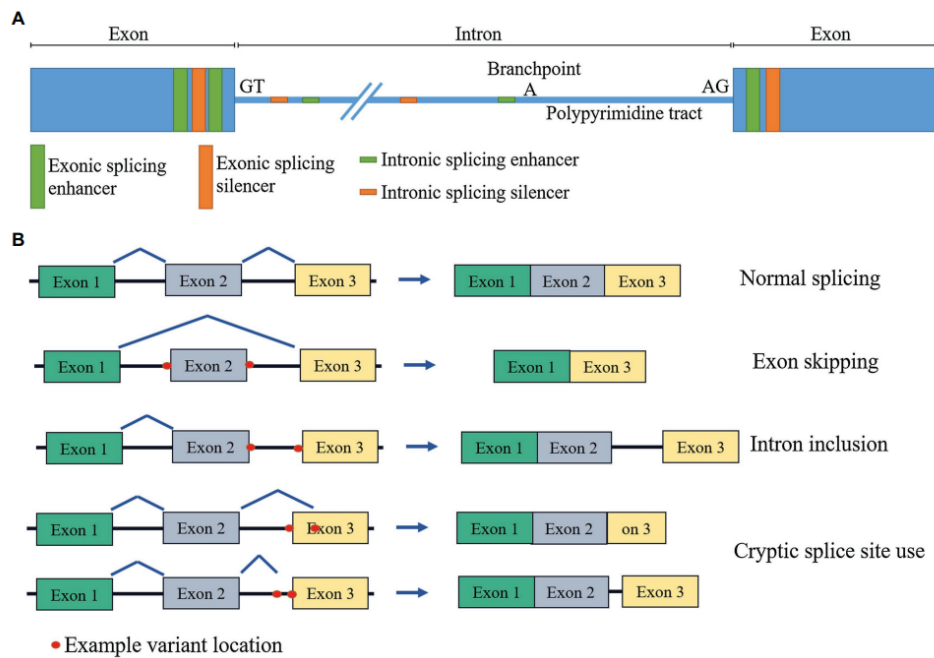


Figure 41: Schéma des éléments de régulation de l'épissage (A) et des effets de l'altération de l'épissage sur l'ARNm (B)  
(D'après Lord et Baralle, 2021)

La perturbation de l'épissage entraîne souvent un décalage du cadre de lecture du transcrite et l'introduction d'un codon STOP prématuré, conduisant à la traduction de protéines tronquées avec un potentiel impact délétère par perte de fonction ou bien par effet dominant négatif. Lorsque cela se produit, les transcrits sont généralement ciblés par la désintégration médiée par Nonsense-Mediated Decay (NMD), un processus par lequel les cellules éliminent les transcrits anormaux afin d'éviter la production de protéines tronquées (Corcos et Solier, 2005). Il est donc intéressant de pouvoir prédire l'impact de l'effet d'un variant sur l'épissage.

Cinq prédicteurs d'épissage sont historiquement utilisés (Khan *et al.*, 2020): **MaxEntScan (MES)** (Yeo et Burge, 2004), **NNSplice (NNS)** (Reese *et al.*, 1997), **Human Splicing Finder (HSF)** (Desmet *et al.*, 2009, p. 200), **SpliceSiteFinder (SSF)** (Shapiro et Senapathy, 1987), **GeneSplicer prediction (GSP)** (Pertea *et al.*, 2001). Ils permettent la prédiction d'une altération ou d'une perte du site d'épissage exprimée par le biais du calcul d'un pourcentage de variation entre la force du site sauvage et la force du site intégrant le variant. Les seuils sont dépendants des programmes et on considère un effet lorsque l'on observe une différence de -5% à -15% : SSF: -5%, NNS: -10%, MES: -15% (Houdayer *et al.*, 2012). Ils se révèlent assez efficaces lorsque le variant est proche des sites consensus mais les prédictions peuvent être plus variables lorsque le variant s'en éloigne. Dans le cadre de ce travail de thèse, j'ai considéré qu'il existait un effet potentiel d'un variant sur l'épissage lorsqu'au moins 2 scores sur 3 avaient un pourcentage de variation inférieur au seuil. De même, la création de sites cryptiques d'épissage en plus des sites canoniques sont évalués de la même manière.

J'ai par ailleurs également utilisé les outils SpliceAI (Jaganathan *et al.*, 2019) et SPiP (Splicing Prediction Pipeline) (Leman *et al.*, 2022), plus récemment développés, et qui permettent d'améliorer la détection des variants d'épissage dans les introns et l'identification d'effets autres comme par exemple des variants impliquant les points de branchement.

SPiP utilise un ensemble d'outils bioinformatiques complémentaires (SPiCE pour l'altération des sites consensus, BPP pour les points de branchement,  $\Delta$ tESRseq pour les ESRs) pour prédire un défaut d'épissage, et hiérarchiser les études sur ARN, quelle que soit la position du variant au sein du gène (exonique, intronique, intronique profond) et le motif affecté (Leman *et al.*, 2022). Il fournit un score entre 0 et 1, un score proche de 1 étant en faveur d'un effet délétère sur l'épissage.

SpliceAI est basé sur le *deep learning* et prédit si chaque position dans un pré-ARNm transcrit est un donneur d'épissage, un accepteur d'épissage ou aucun des deux, en utilisant comme donnée uniquement la séquence génomique du pré-ARNm transcrit. Il fournit 2 scores, un pour la perte d'un site d'épissage et un pour le gain d'un autre site d'épissage (Jaganathan *et al.*, 2019).

### *c) Les outils de prédiction des variants de structure*

Il existe actuellement plusieurs outils de prédiction des variants de structure, qui ne font pas tous consensus. Dans le cadre de cette thèse nous avons utilisé l'outil AnnotSV v3.1, développé au sein de notre laboratoire (Geoffroy *et al.*, 2018) et qui permet d'annoter les CNV en rapportant les données de fréquence en population générale et de prioriser ceux qui incluent un ou plusieurs gènes morbides. Ce dernier permet d'intégrer plus de 60 annotations par SV incluant le type de SV (délétion, duplication, insertion, inversion etc..) et leurs points de cassure, le chromosome et le(s) gène(s) concernés, des annotations spécifiques à chaque gène (OMIM, protéine, pLI, haploinsuffisance), la fréquence dans différentes bases de données de population générale (gnomAD, 1000 Génomes) et de variants pathogènes (ClinVar), la classe ACMG/ClinGen prédite etc. (Clinical Genome Resource, <https://clinicalgenome.org/>).

## **3. La classification ACMG**

A l'échelle d'un génome entier, on identifie plus de trois millions de variations dont la grande majorité n'aura cependant pas d'effet pathogène. Il est donc nécessaire d'identifier parmi l'ensemble de ces variations celle(s) associée(s) au phénotype du patient.

La classification de l'ACMG (American College of Medical Genetic and Genomics) a émis des recommandations en 2015 indiquant de ne plus utiliser les termes « mutation » ou « polymorphisme » mais seulement le terme « variant » qui sera interprété selon le niveau de pathogénicité (Richards *et al.*, 2015). L'ACMG propose une méthode de classification des

variants en pondérant les principaux critères utilisés pour aboutir à une classification en cinq classes : bénin (classe 1), probablement bénin (classe 2), de signification inconnue (classe 3), probablement pathogène (classe 4) et pathogène (classe 5) (Richards *et al.*, 2015; Amendola *et al.*, 2016). Cette classification permet une harmonisation dans l'interprétation des variants mais certains critères, moins objectifs, peuvent être déterminés différemment en fonction du biologiste en charge de l'interprétation, expliquant pourquoi le même variant peut parfois être classé différemment selon le laboratoire ayant réalisé l'analyse. Cette classification a été également adaptée aux SV avec le concours de ClinGen (Clinical Genome Resource, <https://clinicalgenome.org/>).

Les critères utilisés pour cette classification sont regroupés dans les catégories suivantes (Tableau 9):

- Les données de population : selon la fréquence du variant dans la population générale.
- Les données bioinformatiques et prédictives : selon l'impact potentiel du variant (sur l'expression, la fonction et l'intégrité de la protéine) mais également le type de variant (modification du même AA rapportée comme pathogène, variant non-sens dans un gène ou la perte de fonction est un mécanisme pathogène rapporté).
- Les données fonctionnelles : selon l'effet fonctionnel du variant lorsqu'il est connu.
- Les données de ségrégation familiale : selon la présence ou non du variant chez des apparentés du cas index atteints ou sains et le caractère *de novo* d'un variant.
- Les données alléliques : selon que le variant apparaisse en *trans* ou en *cis* d'un autre variant pathogène ou non (pathologies autosomiques récessives).
- Les données phénotypiques : selon la bonne ou mauvaise concordance clinique, et le degré de spécificité du phénotype vis-à-vis du gène concerné.

**Tableau 9: Classification ACMG de pathogénicité des variants**

	Benign			Pathogenic		
	Strong	Supporting	Supporting	Moderate	Strong	Very Strong
<b>Population Data</b>	MAF is too high for disorder <i>BA1/BS1</i> OR observation in controls inconsistent with disease penetrance <i>BS2</i>			Absent in population databases <i>PM2</i>	Prevalence in affecteds statistically increased over controls <i>PS4</i>	
<b>Computational And Predictive Data</b>		Multiple lines of computational evidence suggest no impact on gene /gene product <i>BP4</i> Missense in gene where only truncating cause disease <i>BP1</i> Silent variant with non predicted splice impact <i>BP7</i>	Multiple lines of computational evidence support a deleterious effect on the gene /gene product <i>PP3</i>	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before <i>PM5</i> Protein length changing variant <i>PM4</i>	Same amino acid change as an established pathogenic variant <i>PS1</i>	Predicted null variant in a gene where LOF is a known mechanism of disease <i>PVS1</i>
<b>Functional Data</b>	Well-established functional studies show no deleterious effect <i>BS3</i>		Missense in gene with low rate of benign missense variants and path. missenses common <i>PP2</i>	Mutational hot spot or well-studied functional domain without benign variation <i>PM1</i>	Well-established functional studies show a deleterious effect <i>PS3</i>	
<b>Segregation Data</b>	Non-segregation with disease <i>BS4</i>		Co-segregation with disease in multiple affected family members <i>PP1</i>	Increased segregation data →		
<b>De novo Data</b>				<i>De novo</i> (without paternity & maternity confirmed) <i>PM6</i>	<i>De novo</i> (paternity & maternity confirmed) <i>PS2</i>	
<b>Allelic Data</b>		Observed in <i>trans</i> with a dominant variant <i>BP2</i> Observed in <i>cis</i> with a pathogenic variant <i>BP2</i>		For recessive disorders, detected in <i>trans</i> with a pathogenic variant <i>PM3</i>		
<b>Other Database</b>		Reputable source w/out shared data = benign <i>BP6</i>	Reputable source = pathogenic <i>PP5</i>			
<b>Other Data</b>		Found in case with an alternate cause <i>BP5</i>	Patient's phenotype or FH highly specific for gene <i>PP4</i>			

Le tableau organise chacun des critères en fonction de leur caractère bénin (côté gauche) ou pathogène (côté droit). BS, benign strong; BP, benign supporting; FH, family history; LOF, loss-of-function; MAF, minor allele frequency; path., pathogenic; PM, pathogenic moderate; PP, pathogenic supporting; PS, pathogenic strong; PVS, pathogenic very strong (D'après Richard *et al.*, 2015).

#### 4. Les études de ségrégation familiale

Selon le mode de transmission suspecté pour un variant candidat identifié chez un cas index, le clinicien peut être amené à proposer aux parents de réaliser une étude de ségrégation familiale, qui sera réalisée par séquençage Sanger pour les SNV et par qPCR ou FISH (*Fluorescence In Situ Hybridization*) pour les CNV, notamment pour valider l'absence ou la présence du variant chez d'autres individus sains ou atteints de la même famille, et ainsi guider le raisonnement pour l'interprétation de ce dernier.

#### 5. Les analyses fonctionnelles

Dans certains cas, il n'y a pas suffisamment de preuves pour déterminer si les variants identifiés sont responsables de la pathologie. L'utilisation d'outils de prédiction *in silico* peut aider le biologiste dans l'interprétation, cependant, ils ne suffisent pas à eux seuls et sont

parfois peu informatifs. Le variant peut alors être considéré comme étant un variant « de signification incertaine » (classe 3 selon la classification ACMG), y compris après une discussion en réunion de concertation pluridisciplinaire (RCP) et une étude de ségrégation familiale correctement menée. C'est notamment souvent le cas de variants faux-sens identifiés dans des gènes de l'X, en particulier lorsqu'ils sont hérités de femmes asymptomatiques et qu'aucun apparenté masculin sain de la famille ne peut être testé. Afin de sortir de cette impasse, le recours à des études fonctionnelles est souvent nécessaire pour mieux comprendre l'impact du variant sur la transcription (étude sur ARNm) (Courraud *et al.*, 2021), ou sur l'expression ou la fonction de la protéine (études protéiques) (Mattioli *et al.*, 2019; Karam *et al.*, 2023), et donc son implication dans le phénotype du patient. Il est possible dans certains cas de réaliser l'étude directement chez le patient, *in vivo*, mais cela représente une minorité de cas. Il est ainsi souvent nécessaire de « mimer » l'effet de la variation dans un environnement contrôlé *in vitro*. Ces tests fonctionnels peuvent également être utilisés pour confirmer l'implication dans le phénotype d'un gène encore non-décrit en pathologie humaine. Il existe différents types d'études fonctionnelles couramment réalisées dans le domaine du diagnostic de la DI (liste non-exhaustive).

#### *a) Évaluation de l'effet d'un variant sur les transcrits :*

Nous avons vu qu'un variant peut entraîner un effet sur l'épissage pouvant aboutir soit à la synthèse d'un transcrit anormal (saut d'exon, inclusion de tout ou partie d'un intron) soit à une absence de transcrit par mécanisme de *Nonsense Mediated mRNA Decay* (NMD), qui élimine les ARNm comportant un codon stop prématuré. La multiplicité des altérations d'épissage imputables à un variant d'épissage rend d'autant plus complexe l'association entre un variant et phénotype. Ainsi, caractériser ces transcrits aberrants peut devenir un réel défi.

#### ❖ **Séquençage des transcrits après RT-PCR:**

Pour évaluer l'impact potentiel d'un variant sur l'épissage, il est possible de séquencer l'ARNm extrait à partir des globules blancs du patient, si le gène est exprimé dans ces cellules. Il est également possible de quantifier par une technique de PCR quantitative les transcrits mutés par rapport aux transcrits sauvages et de comparer l'expression de l'allèle porteur du variant avec celle de l'allèle sauvage.

#### ❖ **Technique du mini-gène :**

Le minigène est une version simplifiée d'un gène, cloné dans un vecteur (Baralle et Baralle, 2005). Ceux utilisés pour analyser l'épissage sont classiquement constitués de 2 exons, séparés par un intron dans lequel est introduit l'exon d'intérêt additionné de séquences introniques flanquantes, créant alors un mini-gène à 3 exons. S'en suit une transfection dans des cellules puis une analyse de l'ARN. Certaines substitutions ponctuelles localisées dans des éléments de régulation d'épissage ont pu être ainsi identifiées (Di Giacomo *et al.*, 2013;



Steffensen *et al.*, 2014) et des variations initialement classées comme VSI, ont pu être reclassées (Bonnet *et al.*, 2008; Théry *et al.*, 2011).

### *b) Évaluation de l'effet fonctionnel d'un variant sur la protéine*

#### ❖ ***In vivo*** :

##### ▪ A partir de prélèvements réalisés chez le patient :

###### ➤ Dosages biochimiques et mesure de l'activité enzymatique :

Un déficit enzymatique dans une voie métabolique entraîne une accumulation des métabolites ou substrats en amont de l'enzyme déficiente et une diminution des produits en aval. L'activité de l'enzyme déficiente et les métabolites en excès peuvent être mesurés.

###### ➤ Signatures biologiques :

Plusieurs signatures biologiques ont été mises au point dans des pathologies avec DI, notamment une signature « interféron » dans les interféronopathies (Rice *et al.*, 2013), une signature épigénétique dans le syndrome de Tatton-Brown-Rahman, le syndrome lié aux variants du gène *DYRK1A* ou des syndromes liés à des CNV (Courraud *et al.*, 2021; Rooney et Sadikovic, 2022).

##### ▪ Utilisation de modèles cellulaires et de modèles animaux :

Différents modèles animaux ont massivement contribué à la compréhension physiopathologique des maladies rares, en particulier le modèle murin et le poisson zèbre. Dans le champ de la DI, il peut s'agir de valider l'implication d'un gène ou de variants candidats dans le phénotype des patients en démontrant des altérations phénotypiques chez l'animal en lien avec les signes cliniques chez l'homme (El Chehadeh *et al.*, 2022), de disséquer les mécanismes physiopathologiques pouvant expliquer l'existence d'un phénotype chez l'homme (Montillot *et al.*, 2023), de déterminer le rôle de protéines clés au niveau de certains organes et d'étudier des pistes thérapeutiques (di Blasio *et al.*, 2018; Delvallee et Dollfus, 2023). Ces modèles permettent d'observer l'effet de variants candidats au niveau d'un organe cible comme le cerveau et de prouver l'implication d'un gène dans une pathologie.

##### ▪ Cellules souches pluripotentes induites (iPSC : *Induced Pluripotent Stem Cell*) et organoïdes :

Elles sont créées artificiellement en laboratoire en « reprogrammant » les fibroblastes du patient, puis sont différenciées en cellules d'intérêt pour permettre, par exemple, l'étude *ex vivo* des effets du variant du patient sur la protéine (Gelinas *et al.*, 2017). Cette différenciation peut être réalisée selon notre tissu d'intérêt, 2 ou 3 dimensions comme c'est le cas des organoïdes, version miniature et simplifiée d'un organe, fabriquée *in vitro* en 3 dimensions et

qui présente une micro-anatomie réaliste. Un modèle d'organoïde de rein a par exemple montré son intérêt dans la validation de variants impliquée dans un type de ciliopathie (Forbes *et al.*, 2018). Les organoïdes, pour lesquels le terme de phénotype est également utilisé, peuvent cependant présenter certaines limites, notamment pour récapituler la complexité de compartimentation de certains organes ou d'organisation cellulaire. Différentes approches de *readout* fonctionnel sont régulièrement évaluées (Hofer et Lutolf, 2021; Zhao *et al.*, 2022). Il est également possible d'introduire des variants d'intérêt au sein des iPSC par le biais de la technologie CRISPR/Cas9, technique permettant de « couper » l'ADN au niveau d'une séquence cible d'un gène d'intérêt, pour ensuite le « réparer » en y intégrant le variant étudié. Il est ainsi possible de créer des cellules KO en vue de caractériser des variants difficiles à interpréter (Jinek *et al.*, 2012).

❖ ***In vitro* :**

▪ Etude d'expression protéique et localisation cellulaire en microscopie confocale :

Une mutagenèse dirigée permet d'exprimer un variant *in vitro* dans une lignée cellulaire d'intérêt de manière à surexprimer la protéine mutante, laquelle peut ensuite être marquée par un marqueur fluorescent afin de localiser son expression au niveau cellulaire (nucléaire, cytoplasmique ou membranaire). Si le variant étudié entraîne une perte de fonction induisant un défaut d'adressage cellulaire, ceci est un argument en faveur de l'implication du variant dans le phénotype du patient (El Chehadeh *et al.*, 2022; Mattioli *et al.*, 2019) (Figure 42).

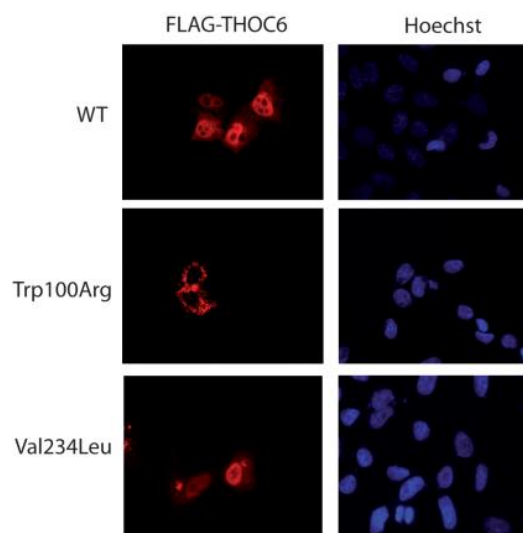


Figure 42: Localisation cellulaire en microscopie confocale.

Panel de gauche : localisation cellulaire de la protéine THOC6 (FLAG rouge) normale (noyau) ou mutée (cytoplasme). Panel de droite : noyaux cellulaires marqués en bleu (D'après Mattioli *et al.*, 2019).

▪ Etude d'interactions avec différents ligands:

Par exemple, par immuno-précipitation (Figure 43)

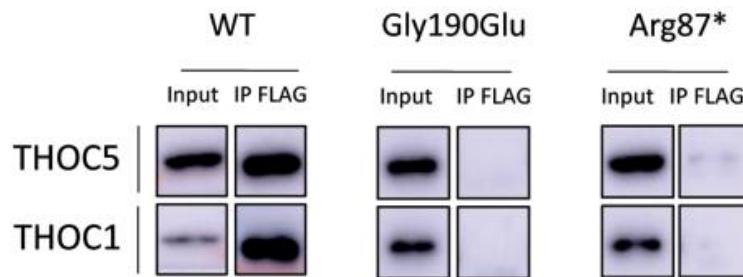


Figure 43: Etude d'interactions par immuno-précipitation

Interactions entre la protéine THOC6 de type sauvage (WT) ou mutante (Gly190Glu et Arg87\*) et les sous-unités THOC5 et THOC1. Les protéines mutantes sont présentes dans l'immuno-précipitat de la protéine THOC6 sauvage et absentes pour les mutantes (D'après Mattioli *et al.*, 2019).

## 6. Les réunions de concertation pluridisciplinaires clinico-biologiques

Les réunions de concertations pluridisciplinaires clinico-biologiques réunissent périodiquement plusieurs experts et sont devenues un élément incontournable dans le champ du diagnostic des maladies rares. Elles permettent de discuter de la pertinence des variants candidats identifiés en adéquation avec le phénotype des patients, des éventuels examens complémentaires à prescrire ou de la nécessité de réaliser une étude de ségrégation familiale pour avancer dans l'interprétation de ces derniers. Cette complémentarité indispensable entre biologiste et clinicien permet des échanges réguliers et fluides autour des dossiers des patients, toujours dans le but d'améliorer leur prise en charge et d'apporter aux familles une réponse la plus claire possible.

## 7. Le partage de données collaboratif national et international

Nous faisons régulièrement appel à de outils de partage de données génomiques nous permettant d'identifier à travers la France, voire le monde, d'autres patients porteurs de variants candidats dans le même gène que celui que nous étudions et dont l'interprétation est complexe. C'est par exemple le cas de l'outil GeneMatcher (Sobreira *et al.*, 2015) ([GeneMatcher \(GM\)](#)), dont l'utilité n'est plus à prouver et qui a beaucoup contribué aux avancées rapides récentes dans le domaine du diagnostic de la DI, en permettant à de très nombreuses reprises la mise en commun de patients suivis par des généticiens cliniciens de toutes nationalités et ainsi la formation de cohortes menant à l'identification de nouveaux gènes de DI en un temps réduit.

## 4. Les techniques de NGS ont fait leurs preuves dans le diagnostic de la DI

Avant l'avènement des techniques de NGS et leur utilisation dans le cadre du diagnostic de la DI, le bilan étiologique chez ces patients consistait principalement à rechercher un syndrome de l'X fragile, une anomalie de nombre/de structure d'un chromosome ou un CNV par le biais d'une étude en CGH-array. Si le généticien clinicien avait une orientation diagnostique forte du fait du phénotype du patient, une analyse ciblée d'une région chromosomique précise ou un séquençage des régions codantes du gène suspecté pouvaient être effectués, par FISH ou séquençage Sanger, respectivement.

### 1. L'utilisation de panels ciblés et du séquençage d'exome (ES)

Aujourd'hui, la technologie NGS est largement utilisée pour le diagnostic de la DI. Le séquençage ciblé permet d'analyser un panel de 1600 gènes de DI pour un rendement diagnostique d'environ 25% (Ibarluzea *et al.*, 2020; Redin *et al.*, 2014). L'ES en *trio*, qui permet l'étude de l'ensemble des parties codantes du génome, est un outil de diagnostic puissant ayant mené à l'identification de très nombreux gènes (Ng, Bigham, *et al.*, 2010) avec un taux diagnostique d'environ 30% (de Ligt *et al.*, 2012; Rauch *et al.*, 2012). L'apport de l'ES dans le diagnostic moléculaire a notamment été démontré dans un projet publié en 2014 dans le cadre des anomalies du développement (projet DDD : Deciphering Developmental Disorders) (Figure 44).

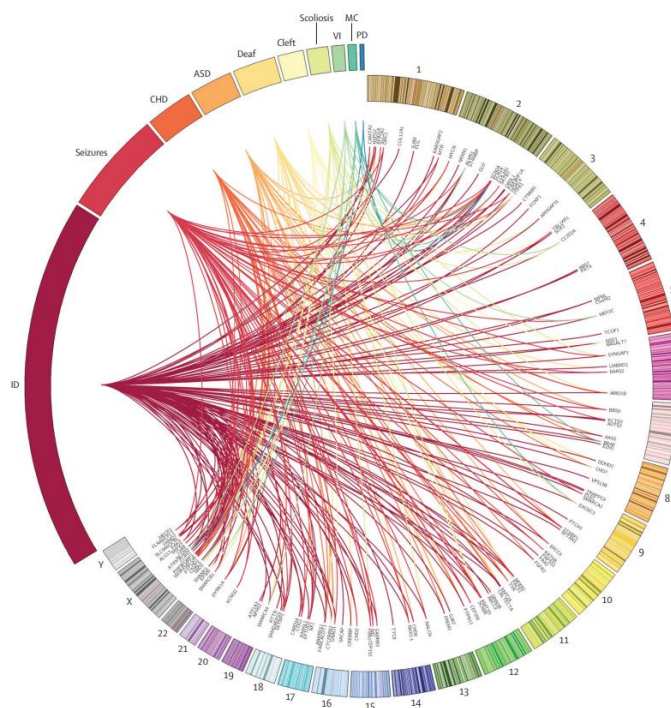


Figure 44: Diagnostics moléculaires associés aux phénotypes (étude DDD)

Base de données qui relie l'emplacement génomique de chaque gène aux phénotype. ID = déficience intellectuelle. CHD = malformation cardiaque congénitale. TSA = troubles du spectre autistique. Deaf=surdité. Cleft = fente orale. VI = déficience visuelle. MC=nanisme microcéphalique. PD=polydactylie. (D'après Wright *et al.*, 2015).

Mille cent treize exomes en *trio* de patients recrutés par les 24 centres de génétique clinique du Royaume-Uni, ont été séquencés par le Wellcome Trust Sanger Institute. L'ES a permis l'identification d'un diagnostic dans 27% des cas, ce qui triplait alors le rendement diagnostique par rapport aux panels de gènes (Wright *et al.*, 2015) (Figure 44).

En ce qui concerne la DILX, des techniques de NGS ont également fait leurs preuves en termes de rendement diagnostique, qu'il s'agisse de séquençage ciblé de gènes de l'X (Ibarluzea *et al.*, 2020; Tzschach *et al.*, 2015) ou « d'X-exome » (Hu *et al.*, 2016; Niranjan *et al.*, 2015; Philips *et al.*, 2014), avoisinant 25 % de taux de diagnostic.

L'adoption rapide du NGS durant la dernière décennie a rapidement fait progresser nos connaissances sur le contenu de notre génome, transformé la pratique clinique de la génétique médicale et remis en question notre compréhension des concepts fondamentaux de la génétique humaine (le risque, la pénétrance, l'expressivité variable, etc.). En particulier, en génétique médicale, l'usage du NGS a permis d'accélérer de manière significative le taux de découverte de nouveaux gènes de MM et l'identification de plus de 1000 nouvelles MM. Cela a mené à une meilleure classification des MM avec TND dont certaines peuvent ne pas être reconnaissables cliniquement, faire partie d'un spectre phénotypique large et/ou présenter une importante hétérogénéité génotypique avec de nombreux gènes impliqués dans une même entité (Figure 45).

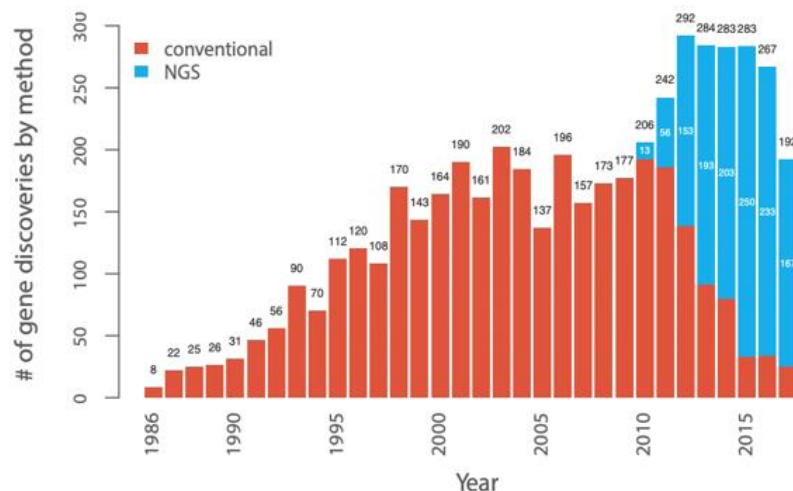


Figure 45: Identification de gènes de MM par NGS *versus* approches conventionnelles

Depuis 2010, les approches de NGS (séquençage cible et exome) (en bleu) ont été utilisées pour la quasi-totalité les découvertes de gènes de MM par rapport aux approches conventionnelles (rouge). (D'après Bamshad *et al.*, 2019).

Cela a également permis d'élargir notre compréhension des conséquences phénotypiques de milliers de variants (Bamshad *et al.*, 2019). En effet, un bon nombre de syndromes rares avec DI ont été décrits bien avant l'avènement du NGS et étaient décrits par

des généticiens cliniciens ou des pédiatres uniquement sur la base de leur phénotype, qui, quand il était récurrent et caractéristique, rendait ces syndromes reconnaissables cliniquement. L'arrivée du NGS a poussé les limites de la clinique, en identifiant des variants pathogènes dans des gènes connus, chez patients dont le phénotype n'avait pas été a priori reconnu. L'analyse du phénotype *a posteriori* a souvent permis de confirmer la pathogénicité du variant identifié et a conduit à la notion de *reverse-phénotyping*. De la même manière, certains gènes dits « syndromiques » (comme par exemple le gène *CREBP* dans le syndrome de Rubinstein Taybi) ont été identifiés chez des patients présentant une DI isolée non syndromique, et donc non identifiables cliniquement, donnant naissance au terme de « *Genotype first* ». Ces constatations, d'abord déroutantes pour bon nombre de cliniciens, ont permis d'augmenter considérablement notre compréhension et notre vision de la syndromologie, et ont conduit à changer notre manière de réfléchir lors de l'analyse d'un phénotype. Cela a également permis de renforcer plus encore les interactions entre biologistes et cliniciens, actuellement indispensable à une analyse de qualité des données issues du NGS.

Ceci est particulièrement vrai en ce qui concerne la DI, l'avènement du NGS ayant permis l'identification de mutations dans de nouveaux gènes de DI ou de confirmer l'implication de gènes candidats (Mattioli *et al.*, 2016, 2017). A l'heure actuelle, environ 1500 gènes ont pu être associé à une forme monogénique de DI, dont 127 sur le chromosome X (Kochinke *et al.*, 2016; Neri *et al.*, 2018; Leitao *et al.*, 2022) (Figure 46).

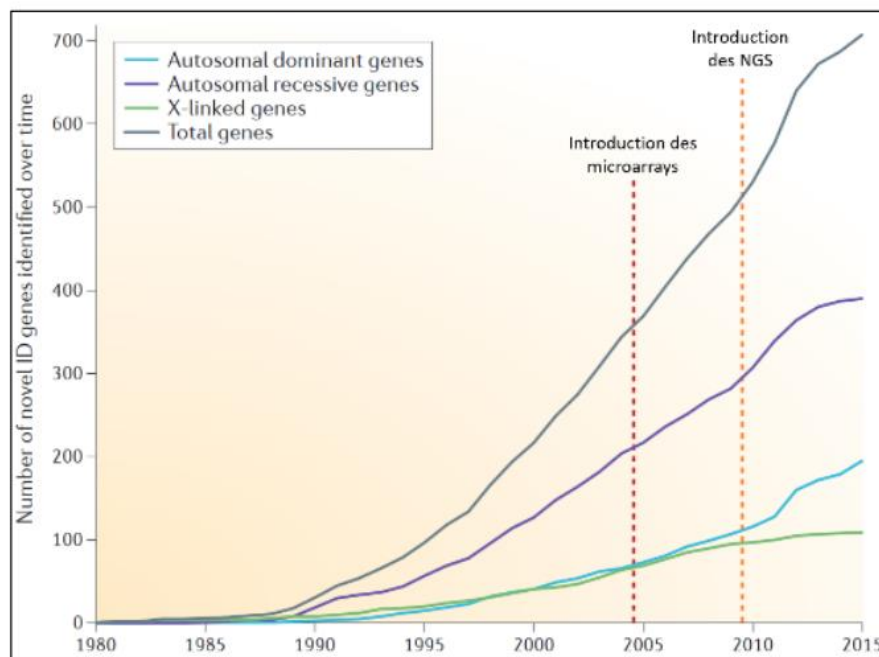


Figure 46: Identification de gènes impliqués dans des formes monogéniques de DI (D'après [Vissers et al., 2016](#)).

Bien que le bénéfice de l'ES ait été démontré essentiellement pour la détection des SNV et *indel*, plusieurs études ont montré qu'il permet également, dans une certaine mesure, de détecter de nombreux autres types de variations, comme des CNV (Pfundt *et al.*, 2017), certains variants introniques (Guo *et al.*, 2012), des disomies uniparentales (Yauy *et al.*, 2020), des variations mitochondriales (Pronicka *et al.*, 2016; Theunissen *et al.*, 2018) (Schlegel *et al.*, 1992), ou des insertions d'éléments mobiles (Torene *et al.*, 2020; Garret *et al.*, 2023).

Plusieurs études ont démontré que la prescription de l'ES entraîne une réduction considérable des coûts des soins de santé, en particulier lorsqu'elle est réalisée très tôt dans le bilan étiologique (« *genetics first* »). En particulier, l'intérêt de l'ES « rapide » ou *fast exome* (résultats en 15 jours environ) dans la prise en charge des patients en situation aiguë, tels que les nouveau-nés avec détresse neurologique ou des fœtus avec anomalies malformatives échographiques n'est plus à démontrer (Chung *et al.*, 2020; McDermott *et al.*, 2022; Monies *et al.*, 2023). Ainsi, de nombreux auteurs s'accordent à dire que l'ES peut remplacer les technologies « traditionnelles » sans compromettre le rendement diagnostique (Monroe *et al.*, 2016; Stark *et al.*, 2017; Thevenon *et al.*, 2016).

## 2. Le séquençage de génome complet (WGS)

### a) Le génome humain : gènes, expression, et régions régulatrices

Déchiffrer le code de la vie : tel était l'objectif du projet collaboratif mondial "Génome Humain" (*The Human Genome Project*, s. d.), visant à décoder la séquence nucléotidique de notre génome (Dulbecco, 1986; Watson, 1990). Ce projet international, débuté en 1989 pour une durée de 15 ans avec un budget estimé à 3 milliards de dollars, était piloté par le National Institutes of Health (NIH) et dirigé par le Dr. J Watson, qui a découvert la conformation de l'ADN (WATSON et CRICK, 1953). Il avait pour objectif de répondre à deux problématiques :

- déterminer la séquence la plus complète possible du génome humain, représentant plus de 3 milliards de positions, réparties dans 23 paires de chromosomes ;
- localiser et définir les unités fonctionnelles du génome, plus particulièrement les gènes.

En parallèle, le génome d'autres espèces que l'Homme est également séquencé (« Genome sequence of the nematode *C. elegans*: a platform for investigating biology », 1998). C'est en 2001, 3 ans avant la date annoncée par le NIH, que les premiers résultats de ce projet ont été publiés (Lander *et al.*, 2001) (Figure 47).

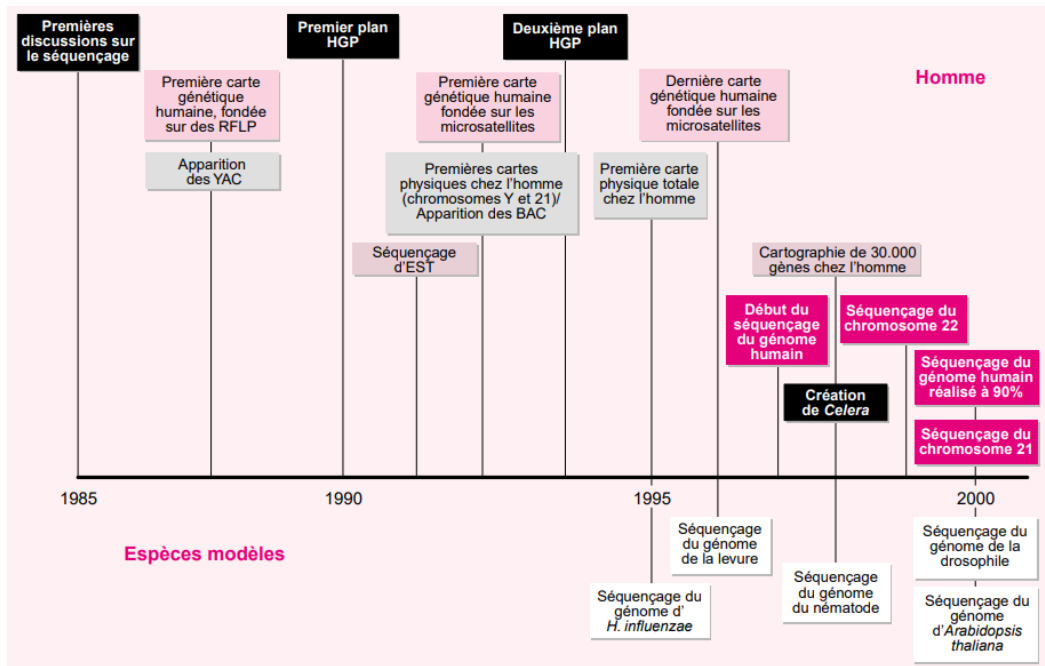


Figure 47: Projet Génome humain et projets des espèces modèles

HGP: human genome project; YAC: yeast artificial chromosome ; BAC: bacterial artificial chromosome ; EST: expressed sequence tags (D'après <https://www.ipubli.inserm.fr/handle/10608/1915>).

En 2003, le séquençage du génome humain était annoncé comme achevé (*Finishing the euchromatic sequence of the human genome | Nature*, s. d.) mais ce n'est finalement qu'en 2022, puis 2023 pour le chromosome Y, que le consortium T2T (Telomere to Telomere) (*Telomere-to-Telomere*, s. d.) a révélé la séquence intégrale et sans aucune base manquante du génome humain (Nurk *et al.*, 2022; Rhie *et al.*, 2023). Ces travaux ont permis de caractériser plus finement la composition génique du génome humain. A l'issue du projet, le nombre de gènes codants détectés dans les séquences a été réduit à 20 000. Ces analyses ont permis d'observer que tous les gènes présentent des transcrits alternatifs, et comportent de très larges introns, tandis que la séquence codante représente une très faible partie du génome (environs 1,5%). Le séquençage du génome complet de la souris, obtenu peu de temps après (*Initial sequencing and comparative analysis of the mouse genome | Nature*, s. d.), a été l'occasion d'explorer la conservation de l'architecture et de la séquence du génome humain avec un autre représentant des mammifères. Cette comparaison a permis d'identifier qu'au total, 5% de la séquence du génome humain sont conservés chez la souris (Guénet, 2005).

En 2004, la première vague de développement est lancée dans le but de proposer un WGS pour le plus grand nombre et de baisser les coûts de façon à passer sous les 1000 \$ pour un WGS humain (Church, 2006). À partir de 2005, les plateformes de séquençage «de seconde génération» émergent et permettent une diminution importante des coûts (Mardis, 2011).

A la suite des résultats obtenus sur l'étude de la séquence du génome humain, de larges consortiums se sont constitués pour explorer les propriétés régulatrices associées au génome non-codant : projet ENCODE (de Souza, 2012; Dunham *et al.*, 2012), projet Roadmap



Epigenomics (Bernstein *et al.*, 2010; Kundaje *et al.*, 2015), projet FANTOM (Andersson *et al.*, 2014). Ces consortiums ont permis l'édition d'atlas recensant et cartographiant l'ensemble des éléments de régulation connus de notre génome que sont les facteurs de transcription, la structure de la chromatine, les modifications des histones, les *enhancers* et plus largement, l'épigénome, à travers les différents tissus et cellules humains (Figure 48, Figure 49 et Figure 50). Cela a également permis de faire un lien entre ces éléments régulateurs et des variants de séquences identifiés dans les maladies humaines. Surtout, ces données ont permis d'attribuer des fonctions biochimiques à environ 80% du génome, en particulier en dehors des régions codantes, et de constituer une vaste source d'annotation fonctionnelle, particulièrement utile en recherche biomédicale (de Souza, 2012; Dunham *et al.*, 2012).

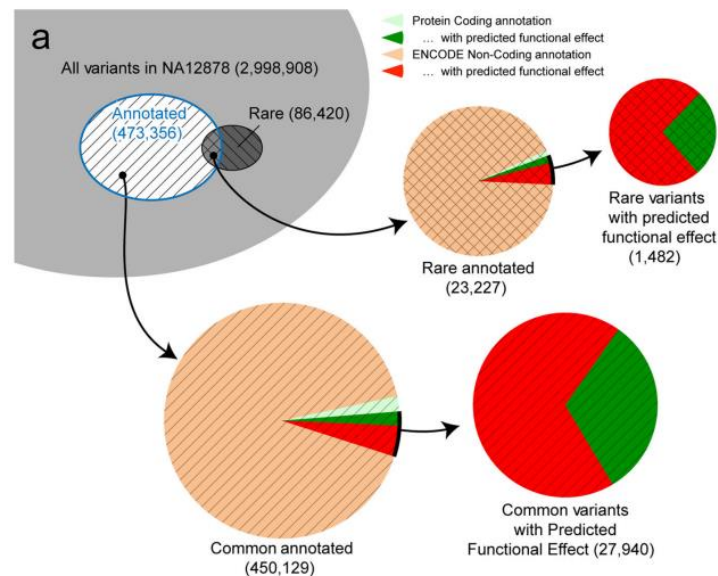


Figure 48: Répartition des variants dans un seul génome Répartition par fréquence, commune ou rare, et par annotation ENCODE, y compris les gènes codant pour des protéines et les éléments non-codants (D'après Dunham *et al.*, 2012).

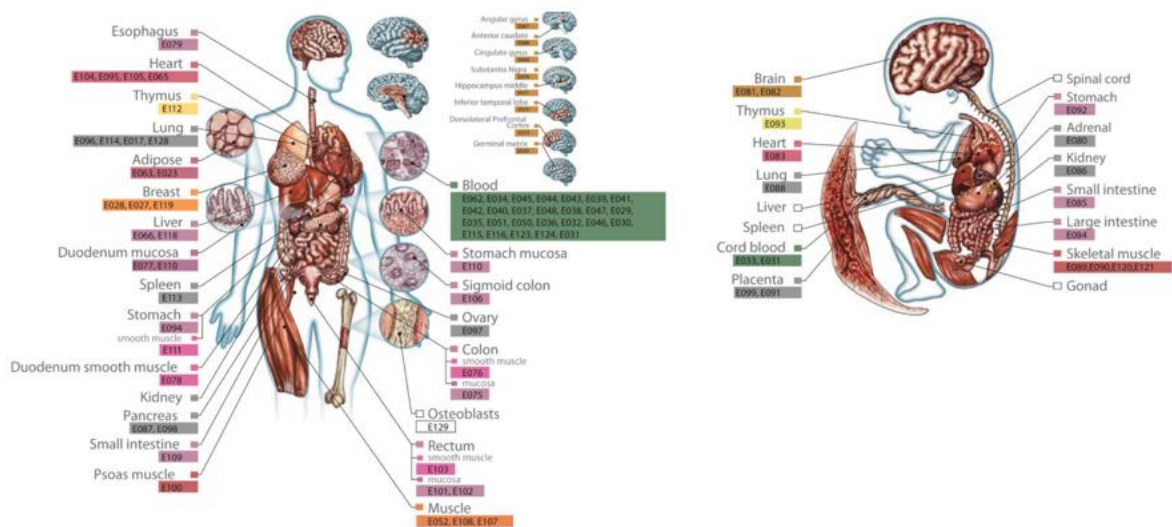


Figure 49: Liste des identificateurs d'épigénome

Gauche : échantillons adultes ; droite : échantillons fœtaux (D'après Kundaje *et al.*, 2015).

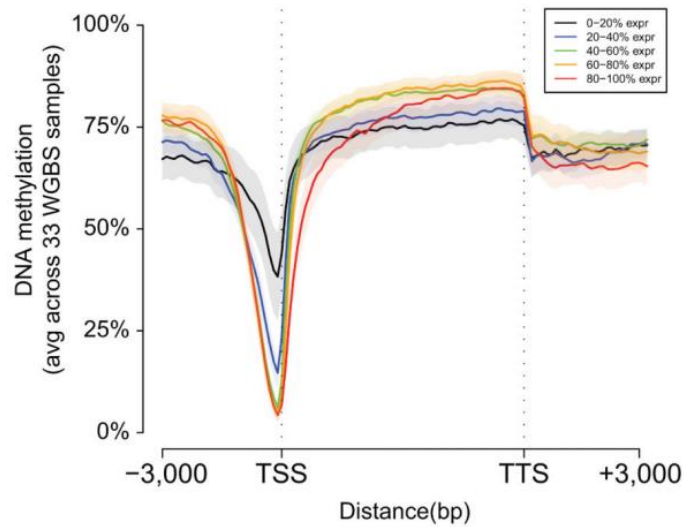


Figure 50: Méthylation de l'ADN selon le niveau d'expression des gènes

La diminution de la méthylation de l'ADN dans les régions promotrices et son enrichissement dans les régions transcrites sont tous deux plus prononcés pour les gènes fortement exprimés (D'après Kundaje *et al.*, 2015).

### b) Intérêts et avantages du WGS

Le WGS, qui était peu utilisé en France dans le cadre du diagnostic jusqu'en 2020, en raison essentiellement de son coût et des difficultés de stockage de données, représente actuellement l'approche la plus complète dans le diagnostic des maladies rares avec DI, avec un taux d'identification de variants pathogènes d'environ 42% des cas (Gilissen *et al.*, 2014) (Figure 51). Tout comme pour l'ES, plusieurs équipes ont prouvé l'intérêt du WGS rapide dans des situations urgentes comme chez des enfants hospitalisés en service de réanimation néonatale, pour permettre de guider ou lever des thérapeutiques, éliminer des diagnostics différentiels et éviter des examens complémentaires invasifs inutiles (Saunders *et al.*, 2012). Le WGS est maintenant déployé en diagnostic en France pour la DI dans le cadre du Plan France Médecine Génomique, et certaines équipes projettent de le mettre en place dans le cadre du DPN.

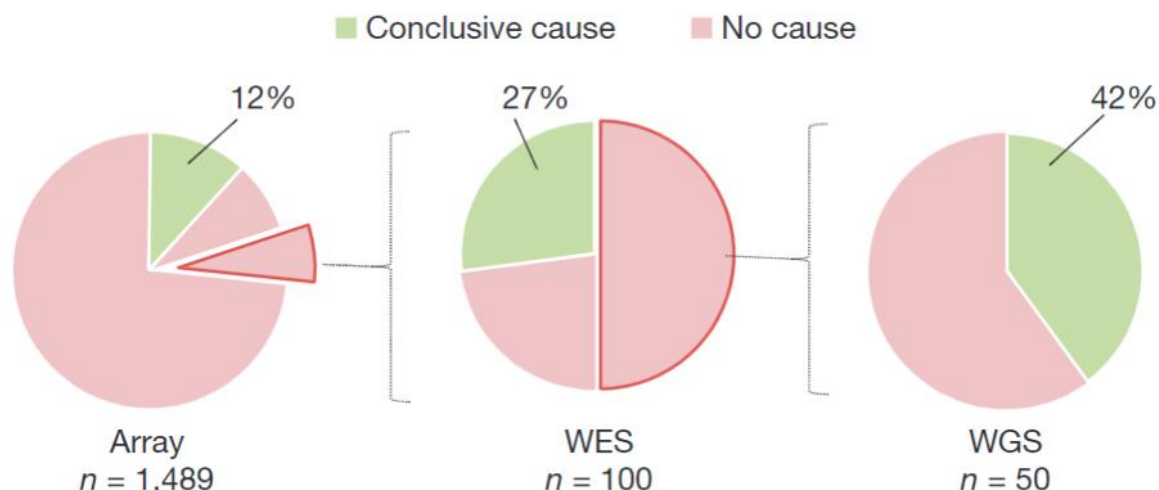


Figure 51: Rendement diagnostique chez des patients avec DI selon la technique

Les pourcentages indiquent le nombre de patients chez qui une cause concluante a été identifiée. Les accolades indiquent le groupe de patients chez lesquels aucune cause génétique n'a été identifiée et dont l'ADN a ensuite été analysé à l'aide de la technique suivante. (D'après Gilissen *et al.*, 2014).

Si les études de séquençage ciblé et d'ES conduisent à l'identification d'une variation causale dans une portion importante de patients, un peu plus de la moitié des patients avec DI restent sans diagnostic, y compris après une analyse par WGS (Gilissen *et al.*, 2014; Mainali *et al.*, 2023). L'ES représente 60 millions de paires de bases et seulement 1,5 à 2% du génome, soit une infime partie, bien qu'on considère cependant que ces parties codantes contiennent plus de 85 % des anomalies responsables de pathologies génétiques chez l'homme (Bamshad *et al.*, 2019). L'immense majorité du génome, essentiellement composée des introns et des régions inter-géniques, n'est donc pas couverte par l'ES, contrairement au WGS. Or, on sait maintenant qu'elles peuvent être le siège de variants pathogènes, notamment des variants introniques profonds, impliqués dans la DI (Coursimault *et al.*, 2022; Enomoto *et al.*, 2022; Chang *et al.*, 2023), notion qui a conduit à parler de « missing heritability » pour les maladies monogéniques (Manolio *et al.*, 2009). Depuis quelques années, plusieurs études ont remis en cause l'intérêt de l'ES au profit du WGS, notamment car la profondeur de lecture des régions codantes séquencée est plus importante dans le cas du WGS ce qui permet une meilleure détection des variants exoniques (Belkadi *et al.*, 2015; Lelieveld *et al.*, 2015; Meienberg *et al.*, 2016). De plus, l'ES a plus de difficultés à détecter des variants situés dans des régions riches en GC, contenant des séquences répétées (Meienberg *et al.*, 2016; Weissbach *et al.*, 2021).

Depuis quelques années des « erreurs de l'ES » sont mises en avant démontrant ses limites malgré ses nombreux avantages : variant intronique profond responsable de l'inclusion d'un pseudo-exon dans le transcrit du gène (Corominas *et al.*, 2022), erreur de « calling » comme une délétion homozygote apparaissant hétérozygote ou une petite duplication « manquée » par le logiciel GATK, variants dans des gènes dont l'implication en maladie humaine n'était pas connue au moment de l'analyse, variants non détectés par les

pipelines locaux (variants dans des régions hors cible, filtres de faible qualité etc.) (Corominas *et al.*, 2022; Denomme-Pichon *et al.*, 2023; Zurek *et al.*, 2021). Plusieurs auteurs ont ainsi publié les « leçons » retenues suite à ces « erreurs », de manière à les éviter à l’avenir et améliorer ainsi leur taux de détection de variants en exome (Corominas *et al.*, 2022).

Ainsi, bien que l’ES soit encore à l’heure actuelle le choix privilégié dans la majorité des études, plusieurs auteurs s’accordent à dire que la réduction des coûts de séquençage et du stockage des données, pourrait permettre prochainement au WGS de remplacer totalement l’ES ainsi que l’ensemble des techniques impliquant la capture de séquences ciblées (Meienberg *et al.*, 2016). Par ailleurs, son rendement diagnostique (dans la DI) est estimé à 60-68 % (Lejeune *et al.*, 2022), contre 43 % avec l’ES (Gilissen *et al.*, 2014) et environ 32 % avec les panels de gènes (Redin *et al.*, 2014; Chérot *et al.*, 2018, expérience du laboratoire de diagnostic de Strasbourg, A Piton).

Le WGS présente d’autres avantages par rapport à l’ES tels que:

- Une meilleure détection des variants structuraux (SV) qui sont des grands réarrangements avec (CNV) ou sans (ABCR, *apparently balanced chromosomal rearrangements*: translocations, inversions, insertions) perte/gain de nombre de copies (Gilissen *et al.*, 2014; Geoffroy *et al.*, 2018; Dutta *et al.*, 2019; Tan *et al.*, 2020; Karam *et al.*, 2023), et dont certains seraient potentiellement « manqués » en exome (Figure 52). Ainsi, Palmer *et coll.* ont pu identifier des SV complexes chez 3 patients sur 15 avec encéphalopathie épileptique et ayant bénéficié d’un exome négatif (Palmer *et al.*, 2021).

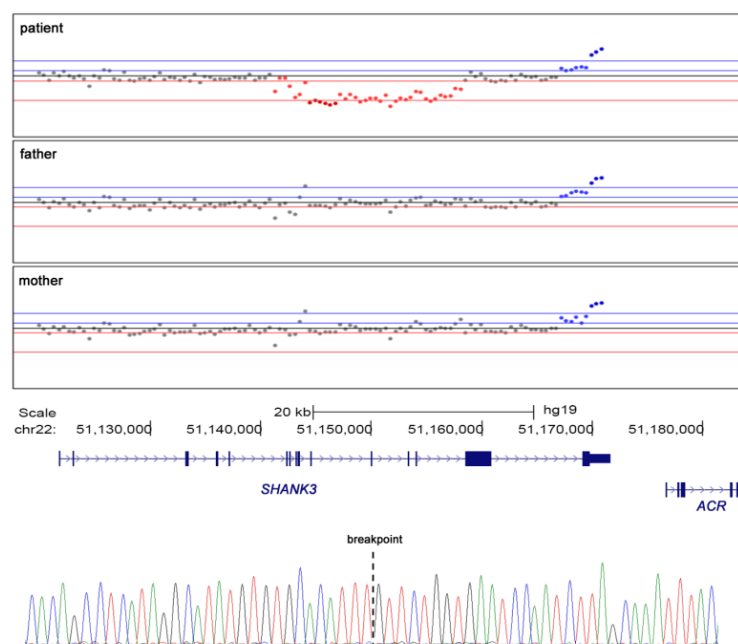


Figure 52: Identification en WGS d’un CNV (délétion) impliquant le gène *SHANK3* (D’après Gilissen *al.*, 2014).

- La détection de variants situés hors des régions codantes :
  - Dans les régions 3'UTR et 5'UTR, promoteur, introns, site d'épissage non-canoniques, site de liaison aux micro ARNs (Gilissen *et al.*, 2014; Devanna *et al.*, 2018; Paolo Devanna *et al.*, 2018). En effet, les données d'ES sont susceptibles de n'identifier que les variants UTR qui se trouvent à proximité des premiers et derniers exons des gènes alors que le WGS est en mesure d'analyser les variants distaux ou mal définis (Wright *et al.*, 2021).
  - Variants de type expansions/répétitions en tandem (Depienne et Mandel, 2021; Dolzhenko *et al.*, 2020) (Dolzhenko *et al.*, 2017, Depienne et Mandel 2021) et insertion d'éléments mobiles (Biquet *et al.*, 2022).
- Une meilleure détection des variants de l'ADN mitochondrial, même si le WGS n'est pas encore aussi performant que les panels dédiés à cette analyse à ce jour.
- La détection de variants dans les régions codantes difficiles à identifier par d'autres techniques. Ceci a été très bien illustré par Gilissen *et coll.* qui a pu identifier en WGS une délétion intra-exonique de *MECP2* qui avait été « loupée » lors du séquençage Sanger initial et de l'analyse par MLPA (amplification multiplex par sonde de ligature) car les sites de liaison des amorces étaient situés juste à l'extérieur de la région délétée (Gilissen *et al.*, 2014) (Figure 53).

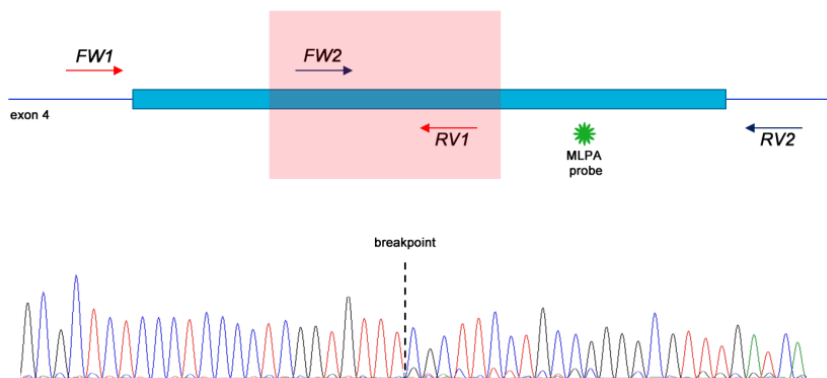


Figure 53: Délétion au sein de l'exon 4 du gène *MECP2* détectée en WGS (D'après Gilissen *et al.*, 2014).

### c) Limites du WGS

#### **Le nombre important de variants**

Un génome humain compte environ 4 à 5 millions de variations, dont plus de 99,9% sont des SNP et petites *indel* et 10 000 sont localisées dans les séquences codantes. Par ailleurs, un génome contient environ 2100 à 2500 variants structuraux (1000 grandes délétions, 160 CNV, 915 insertions Alu, 51 insertions SVA (SVA-retrotransposons), 4 NUMT (nuclear mitochondrial DNA segment), et 10 inversions, affectant 20 millions de bases de la séquence (Auton *et al.*,

2015; Sudmant *et al.*, 2015). Encore plus que pour l'ES, le nombre de variants générés à l'issue du WGS est donc colossal avec souvent plusieurs variants possiblement candidats.

### ***Les difficultés de stockage des données***

Une des difficultés principales rencontrées avec le WGS concerne les difficultés en termes de volume et de coût du stockage des données issues du séquençage ainsi que leur traitement bioinformatique (Hsi-Yang Fritz *et al.*, 2011; Krumm et Hoffman, 2020).

### ***Le challenge des variants non-codants***

Bien que le WGS nous donne accès aux séquences des régions non-codantes, notamment introniques, la quantité importante de variants contenus dans ces dernières rend leur interprétation encore difficile à l'heure actuelle. Ainsi, la majorité des résultats positifs issus du WGS en diagnostic résulte de l'étude des régions codantes (Gilissen *et al.*, 2014).

### ***Le coût***

Le coût du WGS, bien qu'ayant diminué significativement ces cinq dernières années (Figure 54), reste une préoccupation majeure de santé publique à l'échelle de la population d'un pays. Schwartze, en 2020, estimait le coût d'un WGS en *trio* à 7 050 £ (9 330 \$ US) (Schwarze *et al.*, 2020). D'autres résultats déjà publiés présentaient diverses estimations allant de 1 421 € (1 602 \$ US) (K. van Nimwegen *et al.*, 2016) à 6 435 \$ CAN (4 975 \$ US) (Jegathisawaran *et al.*, 2020), s'expliquant principalement par des différences de méthodologie et de système de santé locaux.

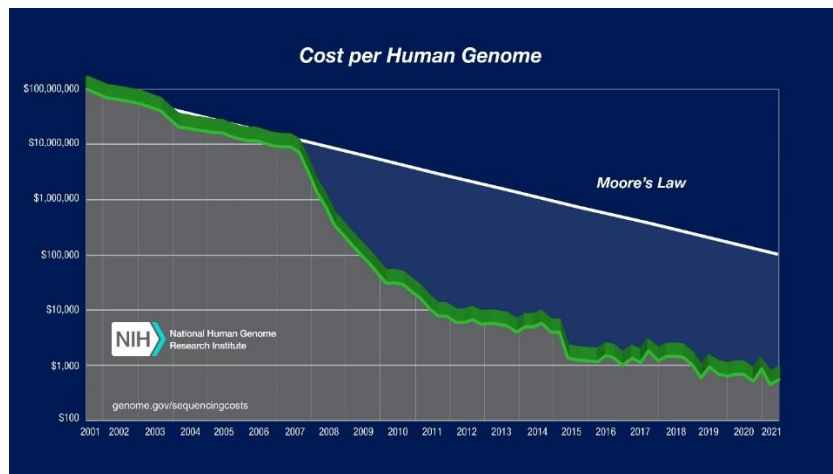


Figure 54: Le coût du WGS, de 2001 à 2021  
(D'après *The Cost of Sequencing a Human Genome*, s. d.)

Par ailleurs, une étude récente a démontré que le WGS coûtait environ 1000 dollars de moins que les tests standards (qui comprenaient l'ACPA, la recherche de syndrome de l'X fragile, les tests ciblés sur un seul gène et les panels de gènes) et qu'il était plus de deux fois plus efficace (en termes de nombre de diagnostics moléculaires effectués) (Ontario Health

(Quality), 2020). La question de cet enjeu médico-économique à l'échelle nationale représente l'un des objectifs de l'étude française DEFIDIAG (Lejeune *et al.*, 2022).

### ***La non-détection des variants de transcrits***

Le WGS ne permet pas de détecter les variants affectant les transcrits, la machinerie de méthylation (méthylome) ou de régulation génique, pour lesquels il est possible de faire appel à d'autres techniques telles que le RNA-seq (Ozsolak et Milos, 2011; Courraud *et al.*, 2021), le Methyl-Seq (Brunner *et al.*, 2009) et le Chip-Seq (Di Vona *et al.*, 2015; Johnson *et al.*, 2007), respectivement.

### ***d) Le WGS en diagnostic : le Plan France Médecine génomique 2025***

Pour s'assurer que chaque personne atteinte d'une maladie rare d'origine génétique puisse accéder aux nouvelles technologies de la médecine génomique de manière équitable sur l'ensemble du territoire, la France a mis en place le Plan France médecine génomique 2025 (PFMG 2025) ([genomic\\_medicine\\_france\\_2025.pdf](https://www.sante.gouv.fr/IMG/pdf/genomic_medicine_france_2025.pdf) ([sante.gouv.fr](https://www.sante.gouv.fr))).

La DI s'inscrit dans un contexte particulier qui lui est propre car il s'agit de la 1<sup>ère</sup> cause de consultation dans les services de génétique médicale, d'une maladie rare « fréquente » (1% à 3% de la population générale), et d'une cause majeure d'errance diagnostique. Par ailleurs, elle est caractérisée par une très importante hétérogénéité clinique et moléculaire avec une myriade de syndromes rares et plus encore de gènes concernés (qui, du reste, ne sont pas tous identifiés à ce jour ce qui représente un enjeu majeur pour la recherche).

Dans ce contexte, le PFMG 2025, ainsi que le 3<sup>ème</sup> Plan National Maladies rares 2018-2022, ont permis le lancement de quatre projets pilotes dont un est dédié à la DI, le projet DEFIDIAG (DEFicience Intellectuelle DIAGnostic) (PI, Pr Hélène Dollfus) qui a pour objectif principal de démontrer la faisabilité du WGS en diagnostic pour les MM, ainsi que son efficacité, en première intention, dans la détermination des gènes impliqués dans la DI. Ce projet, dont les résultats ne sont pas encore publiés, est très prometteur et illustre la volonté nationale d'améliorer l'accès au diagnostic dans les maladies rares à fortiori avec DI (Binguet *et al.*, 2022; Lejeune *et al.*, 2022).

Toujours dans cette même dynamique, le PFMG 2025 a également permis la mise en place de deux laboratoires de biologie moléculaire dédiés au séquençage génomique (SeqOIA, Paris et AURAGEN, Lyon) permettant aux patients atteints de maladies rares d'avoir accès au WGS dans le cadre du diagnostic et cela sans frais. Actuellement, il est possible de prescrire un examen pangénomique pour plus de 70 préindications, dont 60 dans le champ des maladies rares, incluant la DI (Bilan d'activité 2022 – PFMG 2025 ([aviesan.fr](https://www.aviesan.fr))).

# Objectifs de la thèse

---

L'origine de ce projet de thèse est née d'un intérêt commun de la part du Docteur Amélie Piton et de moi-même autour de la DI d'une part et en particulier autour de la DILX, thème sur lequel nous travaillons depuis plusieurs années (Piton *et al.*, 2013, 2014; Redin *et al.*, 2014; El Chehadeh *et al.*, 2016; Quartier *et al.*, 2017; El Chehadeh *et al.*, 2017; Miguet *et al.*, 2018; Quartier *et al.*, 2019). Par ailleurs, de par ma pratique de la génétique clinique depuis bientôt 15 ans, je suis confrontée pluri-quotidiennement aux questionnements et problématiques qui résident autour du diagnostic non-résolu des familles concernées par la DI de cause rare, qui sont éprouvées par l'errance diagnostique et l'anxiété relative au risque de récurrence de la maladie familiale chez un futur enfant.

L'objectif de ce travail de thèse est multiple :

- Contribuer à la description phénotypique de maladies rares voire très rares avec DILX (Partie I - Caractérisation fine du syndrome de duplication du gène *MECP2*, Partie IV).
- Identifier de nouveaux gènes de DILX dans le but de mieux caractériser le spectre phénotypique des DILX, améliorer la compréhension physiopathologique voire identifier de possibles cibles thérapeutiques (Partie II – Les variations du gène *SLITRK2* dans les troubles du neurodéveloppement).
- Evaluer l'utilité et les difficultés d'interprétation du WGS dans la DILX après bilan étiologique de première intention négatif (Partie III - Poursuite des explorations par WGS chez des patients avec DI présumée liée à l'X après ACPA, panels et ES négatifs).

Par ailleurs, j'ai souhaité à titre personnel me former à l'interprétation de données de WGS, afin d'enrichir mes connaissances théoriques et pratiques dans le domaine de la biologie moléculaire et ainsi améliorer ma pratique de la génétique clinique (Partie IV – Contribution à l'interprétation de données issues d'ES et de WGS de patients atteints de maladies rares suivis dans le service de génétique médicale des HUS).

Ce travail de thèse se décompose donc en quatre parties.

**La première partie a consisté à rapporter les données cliniques, moléculaires et d'imagerie d'un syndrome rare de transmission liée au chromosome X, le syndrome de duplication du gène *MECP2* (région Xq28), avec pour principal objectif une meilleure caractérisation et description phénotypique sur le plan neurologique (clinique et d'imagerie)**



et morphologique en particulier, avec un focus sur le cas des filles atteintes. Il s'agit d'un travail collaboratif national débuté avant mon arrivée dans l'unité de recherche U1112 et qui a permis de rapporter en 2018 cinquante-neuf patients français atteints de ce syndrome, correspondant à l'unique cohorte française et la plus importante cohorte internationale recensées jusqu'alors. Ce premier travail illustre bien la complexité de la DILX, en particulier en ce qui concerne les filles porteuses de CNV pathogènes sur le chromosome X et symptomatiques, chez lesquelles la présence d'un phénotype (parfois sévère) n'est pas toujours bien élucidée. Ce travail a mené à la publication de 3 articles, cités dans la section « Résultats » (El Chehadeh *et al.*, 2016; El Chehadeh *et al.*, 2017; Miguet *et al.*, 2018).

**La seconde partie a consisté en la recherche de nouveaux gènes candidats de DILX :**

Suite à l'identification en ES, chez deux frères atteints de DI, d'un variant hémizygoté dans un nouveau gène candidat de DILX, le gène *SLITRK2* (qui code pour une protéine transmembranaire impliquée dans le contrôle de la croissance des neurites), j'ai initié plusieurs appels à collaboration au niveau national et international (réseaux clinico-moléculaires sur la DI, congrès nationaux et internationaux, GeneMatcher) dans le but de colliger d'autres observations de patients atteints de DI et porteur d'un variant candidat au sein de ce gène. Pour cela, j'ai réalisé un travail de recueil de données moléculaires, cliniques, et iconographiques afin, d'une part, d'analyser les variants identifiés (type de variant, effet protéique, caractère hérité ou non) et d'autre part, de comparer les phénotypes des patients et rechercher d'éventuels signes cliniques récurrents ainsi que discuter la présence de ces variants chez des individus de sexe féminin. Ceci m'a permis de colliger des informations clinico-moléculaires pour 15 individus, issus de 14 familles, porteurs de variants candidats dans le gène *SLITRK2*. Un travail de collaboration international s'est poursuivi dans le but de réaliser des études fonctionnelles poussées (Dr J Um et Pr J Ko – Daegon – Corée du Sud), y compris chez la souris, ayant permis de confirmer l'implication de ces variants dans les TND présentés par les patients. Nous poursuivons l'inclusion de nouveaux patients dans le but d'affiner la description phénotypique de cette pathologie rare et de confirmer le rôle du gène *SLITRK2* dans la DILX. Ce travail a fait l'objet d'une publication (El Chehadeh *et al.*, 2022) citée dans la section « Résultats » et en annexe 2.

**La troisième partie de ce travail a consisté à poursuivre les explorations moléculaires à visée étiologique, en recherche, par le biais d'un WGS, chez des patients avec DI supposée liée à X, chez qui l'ACPA, l'ES et/ou d'un panel de gènes de DI étaient négatifs.**

Pour cela, nous avons recruté quinze patients français atteints de DI non-résolue avec présomption d'une transmission liée au chromosome X, puis réalisé un recueil des données cliniques. Les critères de sélection de ces patients étaient les suivants : présence dans une même famille d'au minimum 3 individus atteints de sexe masculin, sur 2 générations et reliés

par une femme, symptomatique ou non. Les ADNs de quinze patients ont été adressés au Centre National de Recherche en Génomique Humaine (CNRGH - Evry) pour réalisation du WGS. Le travail sur les données de WGS a consisté :

- à rechercher des variants dans les régions codantes qui auraient pu être « manqués » par le séquençage d'exome (défaut de couverture, erreur d'interprétation, gène non-encore décrit) dans les gènes connus, les gènes OMIM, puis tous les gènes du chromosome X.

- à rechercher des variants dans les régions non-codantes: analyses des SNV dans les gènes connus, les gènes OMIM, puis tous les gènes du chromosome X (variations introniques profondes, promoteur, éléments régulateurs), analyse des variants de structure (déséquilibrées: CNV, équilibrées: inversions, translocations, éléments mobiles)...

Cette partie du travail a permis d'identifier un CNV pathogène sur l'X et plusieurs variants candidats dans des gènes de l'X chez différents patients. Une collaboration internationale est notamment en cours pour colliger les variants candidats identifiés dans le gène *BCORL1* (Xq26.1) chez des patients avec TND (coord. Dr N Miyake, Japon). Un autre variant candidat a été identifié dans le gène *MED14* (Xp11.4) chez deux frères et une poursuite des explorations est envisagée, notamment par la réalisation d'études fonctionnelles (collab. N Le May).

**La quatrième partie** traite de ce que j'ai pu effectuer en parallèle de ces précédentes étapes dans le but de me former à l'analyse des données de NGS et mieux comprendre les spécificités de l'analyse d'ES et de WGS. Il s'agissait d'interpréter des données d'ES de patients atteints de DI suivis dans le service de génétique médicale et de patients atteints de maladies dermatologiques rares dans le cadre du projet DERMASEQ.

J'ai ainsi pu contribuer à des travaux de description phénotypique et d'identification de nouveaux gènes concernant des syndromes rares avec DI/TND dont sont atteints plusieurs des patients que je suis et qui ont chacun abouti à une publication, citée en annexe (Ritter *et al.*, 2022 (abstract): annexe 3, El Chehadeh *et al.*, 2021: annexe 4, Durand *et al.*, 2020: annexe 5, Schuermans *et al.*, 2023 (sous presse) : annexe 6).



# Résultats

---

## I. Première partie : une forme rare de DILX, le syndrome de duplication du gène *MECP2*

De 2008 à 2016, une étude collaborative nationale française a été menée (PHRC RMLX, Pr V Des Portes) dont l'objectif principal était de caractériser le phénotype clinique de plusieurs maladies liées au chromosome X avec DI. Dans ce cadre, j'ai essayé de recenser et identifier tous les patients diagnostiqués avec un *MECP2DS* à ce moment précis en France. Pour chaque patient, j'ai recueilli des données médicales et moléculaires rétrospectivement à partir des dossiers médicaux et, le plus souvent possible, lors d'une consultation médicale dédiée. A ce moment, en 2016, 101 patients porteurs d'une duplication de *MECP2* avaient été identifiés, dont 91 garçons et 10 filles.

De l'étude de cette cohorte, ont abouti trois publications :

- La première, en 2016, consistait en la description des résultats d'IRM cérébrales de 30 patients atteints de *MECP2DS* dans le but de rechercher des anomalies communes et récurrentes ainsi qu'essayer d'établir une corrélation génotype-phénotype en fonction de ces dernières (hétérotopies nodulaires, dilatation ventriculaire ou anomalies de la substance blanche périventriculaire) contenu en gènes des duplications (notamment les gènes *FLNA*, *L1CAM* ou *IKBK*) (El Chehadeh *et al.*, 2016).

- La seconde, en 2017, a consisté à rapporter une série de six patientes symptomatiques porteuses d'une dup*MECP2* interstitielle *de novo* et passer en revue les 14 filles symptomatiques déjà publiées, dans le but de mieux définir leur phénotype et d'apporter des indications pour le conseil génétique des couples à risque (El Chehadeh *et al.*, 2017).

- La troisième, en 2018 a consisté, grâce au travail collaboratif débuté auparavant, à décrire le plus finement possible le phénotype, en particulier morphologique et neurologique, de 59 patients français de sexe masculin atteints de *MECP2DS* (Miguet *et al.*, 2018).

Ces travaux ont conduit à plusieurs communications orales et affichées lors de congrès nationaux et internationaux et ont contribué à l'élaboration d'un Protocole National de Diagnostic et de Soins (PNDS), publié en Décembre 2019 ([https://www.has-sante.fr/jcms/p\\_3148006/fr/syndrome-de-duplication-du-gene-mecp2](https://www.has-sante.fr/jcms/p_3148006/fr/syndrome-de-duplication-du-gene-mecp2)).

C'est avec l'étude de ce syndrome rare, qui n'est pas complètement élucidé aujourd'hui, que mon intérêt pour la DI liée à l'X a débuté et s'est poursuivi jusqu'à présent.

## 1. Syndrome de duplication du gène *MECP2* chez les filles : que pouvons-nous apprendre pour le diagnostic et le conseil génétique ?

Les mères porteuses (conductrices) sont généralement considérées comme étant non atteintes. En effet, la grande majorité des mères conductrices rapportées dans la littérature sont asymptomatiques et présentent un biais complet d'inactivation du chromosome X. Cependant, certaines d'entre elles peuvent manifester des symptômes, le plus souvent d'ordre psychologique ou psychiatrique tels qu'une anxiété généralisée, des symptômes dépressifs, un comportement compulsif. Dans certains cas très particuliers (duplication résultant d'un remaniement chromosomique tel qu'une translocation entre un chromosome X et un autosome), les filles peuvent développer une forme sévère de la maladie avec un tableau clinique semblable à celui observé chez les garçons atteints (associant DI sévère, infections récurrentes, constipation chronique, épilepsie, développement très pauvre du langage, et particularités morphologiques observées chez les garçons).

Comme nous l'avons démontré dans le travail ci-dessous, dans de très rares cas, les filles qui portent la duplication « classique » (interstitielle) sur un chromosome X à l'état hétérozygote peuvent présenter une DI non-spécifique et isolée, en général plus modérée que celle des garçons atteints. Dans ce cas précis, le plus souvent, les duplications sont survenues *de novo*, et les études d'inactivation du chromosome X ne montrent pas de biais complet d'inactivation. Néanmoins, dans l'état actuel des connaissances, il n'a pas été établi de corrélation claire entre le ratio d'inactivation du chromosome X, la taille de la duplication et le degré de sévérité de la DI chez ces filles. Il a également été rapporté de très rares cas (moins d'une dizaine) de filles avec DI sévère. Cependant, il n'y a pas eu d'analyse pangénomique réalisée chez l'ensemble de ces patientes pour rechercher une éventuelle autre cause de DI. En conséquence, ces observations sont délicates pour le conseil génétique, notamment pour les situations prénatales.

## Résumé

Nous rapportons une série de six nouvelles femmes symptomatiques porteuses d'un dupMECP2 interstitiel *de novo*, et passons en revue les 14 femmes symptomatiques signalées à ce jour, dans le but de mieux définir leur phénotype et de donner des indications pour le conseil génétique. Une patiente a été adoptée et parmi les 19 autres patients, sept (37%) avaient hérité de leur duplication de leur mère, dont la DI était légère chez 3 (XCI : 70/30, 63/37, 100/0 dans le sang et aléatoire dans la salive), modérée chez une (XCI : aléatoire) et sévère chez 3 (XCI : non informatif et 88/12). Après avoir combiné nos données avec les données de la littérature, nous n'avons pas pu mettre en évidence de corrélation entre l'XCI dans le sang, la taille de la duplication et la sévérité du phénotype, ni expliquer la présence d'un phénotype chez ces femmes. Ces résultats confirment qu'un phénotype anormal, même sévère, peut être un événement rare chez les femmes nées de mères porteuses asymptomatiques, ce qui rend le conseil génétique difficile chez les couples à risque en termes de pronostic en particulier en cas de grossesse.



## Original Article

# Xq28 duplication including *MECP2* in six unreported affected females: what can we learn for diagnosis and genetic counselling?

El Chehadeh S., Touraine R., Prieur F., Reardon W., Bienvenu T., Chantot-Bastaraud S., Doco-Fenzy M., Landais E., Philippe C., Marle N., Callier P., Mosca-Boidron A.-L., Mugneret F., Le Meur N., Goldenberg A., Guerrot A.-M., Chambon P., Satre V., Coutton C., Jouk P.-S., Devillard F., Dieterich K., Afenjar A., Burglen L., Moutard M.-L., Addor M.-C., Lebon S., Martinet D., Alessandri J.-L., Doray B., Miguet M., Devys D., Saugier-veber P., Drunat S., Aral B., Kremer V., Rondeau S., Tabet A.-C., Thevenon J., Thauvin-Robinet C., Perretton N., Des Portes V., Faivre L. Xq28 duplication including *MECP2* in six unreported affected females: what can we learn for diagnosis and genetic counselling?  
Clin Genet 2017; 91: 576–588. © John Wiley & Sons A/S. Published by John Wiley & Sons Ltd, 2016

Duplication of the Xq28 region, involving *MECP2* (dupMECP2), has been primarily described in males with severe developmental delay, spasticity, epilepsy, stereotyped movements and recurrent infections. Carrier mothers are usually asymptomatic with an extremely skewed X chromosome inactivation (XCI) pattern. We report a series of six novel symptomatic females carrying a *de novo* interstitial dupMECP2, and review the 14 symptomatic females reported to date, with the aim to further delineate their phenotype and give clues for genetic counselling. One patient was adopted and among the other 19 patients, seven (37%) had inherited their duplication from their mother, including three mildly (XCI: 70/30, 63/37, 100/0 in blood and random in saliva), one moderately (XCI: random) and three severely (XCI: uninformative and 88/12) affected patients. After combining our data with data from the literature, we could not show a correlation between XCI in the blood or duplication size and the severity of the phenotype, or explain the presence of a phenotype in these females. These findings confirm that an abnormal phenotype, even severe, can be a rare event in females born to asymptomatic carrier mothers, making genetic counselling difficult in couples at risk in terms of prognosis, in particular in prenatal cases.

### Conflict of interest

All authors declare that they have no conflict of interest.

S. El Chehadeh<sup>a,b</sup>, R. Touraine<sup>c</sup>,  
F. Prieur<sup>c</sup>, W. Reardon<sup>d</sup>,  
T. Bienvenu<sup>e</sup>,  
S. Chantot-Bastaraud<sup>f</sup>,  
M. Doco-Fenzy<sup>g</sup>, E. Landais<sup>h</sup>,  
C. Philippe<sup>i</sup>, N. Marle<sup>j</sup>,  
P. Callier<sup>k</sup>, A.-L. Mosca-Boidron<sup>l</sup>,  
F. Mugneret<sup>m</sup>, N. Le Meur<sup>k</sup>,  
A. Goldenberg<sup>l</sup>, A.-M. Guerrot<sup>l</sup>,  
P. Chambon<sup>m</sup>,  
V. Satre<sup>n</sup>, C. Coutton<sup>n</sup>,  
P.-S. Jouk<sup>n</sup>, F. Devillard<sup>n</sup>,  
K. Dieterich<sup>n</sup>, A. Afenjar<sup>o</sup>,  
L. Burglen<sup>o</sup>, M.-L. Moutard<sup>p</sup>,  
M.-C. Addor<sup>q</sup>, S. Lebon<sup>r</sup>,  
D. Martinet<sup>s</sup>, J.-L. Alessandri<sup>t</sup>,  
B. Doray<sup>u</sup>, M. Miguet<sup>b</sup>,  
D. Devys<sup>v</sup>, P. Saugier-veber<sup>w</sup>,  
S. Drunat<sup>x</sup>, B. Aral<sup>y</sup>, V. Kremer<sup>z</sup>,  
S. Rondeau<sup>aa</sup>, A.-C. Tabet<sup>ab</sup>,  
J. Thevenon<sup>a,ac</sup>,  
C. Thauvin-Robinet<sup>a,ac</sup>,  
N. Perretton<sup>ad</sup>, V. Des Portes<sup>ae</sup>  
and L. Faivre<sup>a,ac</sup>

<sup>a</sup>FHU TRANSLAD, Centre de Référence Maladies Rares «Anomalies du Développement et Syndromes Malformatifs» de l'Est, Centre de Génétique, CHU de Dijon, Dijon, France, <sup>b</sup>Service de Génétique Médicale, Institut de Génétique Médicale d'Alsace (IGMA), Centre de Référence Maladies Rares «Anomalies du Développement et Syndromes Malformatifs» de l'Est, Hôpitaux Universitaires de Strasbourg, Hôpital de Hautepierre, Strasbourg, France, <sup>c</sup>Service de Génétique Clinique Chromosomique et Moléculaire, CHU de Saint-Etienne, Saint-Etienne, France, <sup>d</sup>Clinical Genetics, Division National Centre for Medical Genetics, Our Lady's Children's Hospital, Dublin, Ireland, <sup>e</sup>AP-HP, Laboratoire de Génétique et Biologie Moléculaires, HJ Paris Centre, Site Cochin, France; Université Paris

## Xq28 duplication including *MECP2* in six unreported affected females

Descartes, Institut Cochin, INSERM U1016, Paris, France, <sup>1</sup>Service de Génétique et Embryologie Médicales, CHU Paris Est – Hôpital d'Enfants Armand-Trousseau, Paris, France, <sup>2</sup>Service de Génétique, EA3801, SFR-CAP Santé, CHU de Reims, Reims, France, <sup>3</sup>PRBI, Pôle de Biologie Médicale, CHU de Reims, Reims, France, <sup>4</sup>Laboratoire de Génétique Médicale, Hôpitaux de Brabois CHRU, Vandœuvre les Nancy, France, <sup>5</sup>Service de Cytogénétique, CHU de Dijon, Dijon, France, <sup>6</sup>Etablissement Français du Sang, CHU de Rouen, Rouen, France, <sup>7</sup>Service de Génétique et Inserm U1079, Centre Normand de Génomique Médicale et Médecine Personnalisée, CHU de Rouen, Inserm et Université de Rouen, Rouen, France, <sup>8</sup>Laboratoire D'histologie, Cytogénétique et Biologie de la Reproduction, CHU de Rouen, Rouen, France, <sup>9</sup>Département de Génétique et Procréation, CHU Grenoble Alpes, Université Grenoble Alpes, Grenoble, France, <sup>10</sup>Service de Génétique, CHU Paris Est – Hôpital d'Enfants Armand-Trousseau, Paris, France, <sup>11</sup>Unité de neuropédiatrie et pathologie du développement, CHU Paris Est – Hôpital d'Enfants Armand-Trousseau, Paris, France, <sup>12</sup>Service de Génétique Médicale, Centre Hospitalier Universitaire Vaudois CHUV, Lausanne, Switzerland, <sup>13</sup>Unité de Neuropédiatrie, Centre Hospitalier Universitaire Vaudois CHUV, Lausanne, Switzerland, <sup>14</sup>Laboratoire de Cytogénétique Constitutionnelle et Prénatale, Centre Hospitalier Universitaire Vaudois CHUV, Lausanne, Switzerland, <sup>15</sup>Pôle Enfants, CHU de la Réunion – Hôpital Félix Guyon, Saint-Denis, France, <sup>16</sup>Service de Génétique, CHU de la Réunion – Hôpital Félix Guyon, Saint-Denis, France, <sup>17</sup>Laboratoire de Diagnostic Génétique,

CHU de Strasbourg – Hôpital Civil, Strasbourg, France, <sup>18</sup>Laboratoire de Génétique Moléculaire, Faculté de Médecine et de Pharmacie, Rouen, France, <sup>19</sup>Laboratoire de Biologie Moléculaire, Hôpital Robert Debré, Paris, France, <sup>20</sup>Service de Biologie Moléculaire, CHU de Dijon, Dijon, France, <sup>21</sup>Laboratoire de Cytogénétique, Hôpitaux Universitaires de Strasbourg, Hôpital de Haute-pierre, Strasbourg, France, <sup>22</sup>Service de Pédiatrie Néonatale et Réanimation, CHU de Rouen, Rouen, France, <sup>23</sup>Laboratoire de Cytogénétique, Hôpital Robert Debré, Paris, France, <sup>24</sup>GAD, EA4271, Génétique et Anomalies du Développement, Université de Bourgogne, Dijon, France, <sup>25</sup>EPICIME-CIC 1407 de Lyon, Inserm, Service de Pharmacologie Clinique, CHU-Lyon, Bron, France, and <sup>26</sup>Service de Neurologie Pédiatrique, CHU de Lyon-GH Est, Bron, France

Key words: affected females – genetic counselling – *MECP2* duplication syndrome – *MECP2* gene – X chromosome inactivation – Xq28 duplication

Corresponding author: Salma El Chehadeh, MD, Service de Génétique Médicale, Hôpital de Haute-pierre, Avenue Molière, 67098 Strasbourg Cedex, France, Tel.: +33388 12 81 20; fax: +33388 12 81 25; e-mail: salma.elchehadeh@chru-strasbourg.fr

Received 1 August 2016, revised and accepted for publication 17 October 2016

Duplication of the Xq28 region, which includes the *MECP2* gene (dupMECP2) (MIM 300260), leads to increased levels of methyl-CpG-binding protein-2 (MeCP2) and has been primarily described in male patients with severe developmental delay, spasticity, epilepsy, stereotyped hand movements and recurrent infections (1–3). It is an X-linked condition in which carrier mothers are usually considered unaffected. Indeed, most of those that have been reported were asymptomatic with a severely to completely skewed X chromosome inactivation (XCI) pattern (1–4). However, some carrier mothers can manifest neuro-psychiatric symptoms such as generalized anxiety, depression and compulsion despite 100% skewing of X-inactivation and normal *MECP2* RNA levels in peripheral blood (5).

*MECP2* encodes for an essential epigenetic regulator in postnatal brain development that causes transcriptional activation as well as repression of many genes (6, 7). It has previously been shown that failure of homeostasis in *MECP2* expression, either by haploinsufficiency or overexpression, leads to a neurodevelopmental phenotype (8). Moreover, Nageshappa et al. recently showed that cortical neurons derived from different MECP2dup induced pluripotent stem cells (iPSC) lines display increased synaptogenesis and dendritic complexity, leading to an alteration of neuronal network synchronization, which was partially reversible using different compounds that act on epigenetic pathways, including histone deacetylase inhibitor (NCH-51) (9). These findings highlight the importance of correct *MECP2* levels for synaptic plasticity and homeostasis during brain development and function.

While dupMECP2 is thought to be at the origin of 1% of unexplained X-linked intellectual disability (ID) in males (4), it seems to be a very rare cause of ID in females. For example, Grasshoff et al. identified only two affected females in their cohort of 1000 unselected patients with ID (incidence of 0.002, sex ratio not specified) (10). However, during the last decade, 23 affected female patients with an Xq28 duplication involving *MECP2* have been reported. These duplications were of various size and resulted from a *de novo* or inherited intrachromosomal (intraC) Xq28 duplication ( $n=13$ ) (10–16), a *de novo* unbalanced X;autosome translocation or an insertion of the Xq28 duplicated segment into an autosome ( $n=9$ ) (12, 17–22), or a *de novo* intraC Xq28 triplication (one case) (23). It is now well-known that affected females carrying a large dupMECP2 resulting from an X;autosome translocation or an insertion of the duplicated Xq28 segment into an autosome can be severely affected with a phenotype very similar to that observed in male patients, due to the fact that they are functionally disomic for genes which are normally subject to inactivation and because of the larger size of the duplications (16, 18, 20, 22, 24). Conversely, the rare symptomatic female cases with an interstitial duplication usually present with non-specific mild to moderate intellectual deficiency (12, 16). However, several publications have reported interstitial dupMECP2, both *de novo* (14) and inherited (13, 15), in female patients with a severe phenotype. The XCI pattern was usually tested as the majority of female carriers had skewed X-inactivation, but this did not correlate with the presence of





Fig. 1. Photographs of the face and profile of affected females. (a) Patient 1, (b) Patient 2, (c) Patient 3, (d) Patient 5 at the age of 13 (left) and 22 (right) and (e) Patient 6. We can note that these patients are not or mildly dysmorphic. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

symptoms (11, 13). All these observations make genetic counselling difficult in terms of prognosis in female foetuses of female carriers, especially if the question of prenatal diagnosis is raised (16).

Here, we report on the clinical and cytogenetic features of six novel symptomatic female patients carrying a *de novo* intraC dupMECP2. By combining our data with data from a review of the literature of all symptomatic female patients reported to date, with a focus on the XCI pattern, we aimed to further delineate the clinical spectrum of females with dupMECP2 and give clues for genetic counselling.

#### Materials and methods

These female cases were recruited through a French study conducted from 2012 to 2015 to gather cases of Xq28 duplication involving *MECP2* (dupMECP2), for genotype–phenotype correlation purposes. A total of 99 cases (two Swiss patients) including 89 males and 10 females were identified and all were diagnosed in the context of their diagnostic workup for ID. The female cases included six patients with an interstitial (intraC) Xq28 duplication and four patients with an unbalanced X;autosome translocation. This study only reports on the six patients with intraC Xq28 duplications, all of which were detected using genomic microarrays. Array types included array comparative genomic hybridization (array-CGH) (Agilent 150 k or 180 k) (Patients 1, 2, 3, 4, and 6) or single nucleotide polymorphism (SNP) array study using the Illumina SNP-12 V2 (Patient 5) and were performed in the cytogenetic laboratories of the University Hospitals of Lyon-Saint-Etienne (France), Nancy (France), Robert Debré (Paris, France) and Lausanne (Switzerland). The duplications were confirmed by quantitative polymerase chain reaction (PCR) analyses. Parental analyses were performed using quantitative PCR.

578

The XCI pattern in lymphocytes was analysed by determining the methylation profile in the first exon of the androgen receptor gene (*HUMARA*) (25) and at the *FRAXA* locus. To assess differential methylation, samples were either treated with a methylation-sensitive endonuclease (HpaII) or with the EpiTect Bisulfite kit (Qiagen (Venlo, The Netherlands)) according to the manufacturer's recommendations. Semi-quantitative PCR amplification of the polymorphic repeats at both loci was performed with specific fluorescein amidite (FAM) - labelled primers, using an in-house protocol. Amplification products were migrated on an ABI PRISM 3,500 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) or on a CEQ-2000XLS (Beckman Coulter, Brea, CA, USA). The XCI ratios were calculated as previously described (26).

Clinical and molecular data were systematically gathered. All patients' families provided written informed consent and all procedures performed in the studies were in accordance with the ethical standards of the Institutional Research Committee and with the Declaration of Helsinki.

#### Results

##### Clinical and cytogenetic reports

The clinical features of all affected patients are presented in Fig. 1 and Table 1. The results of the XCI pattern studies are shown in Table 1 and molecular details are presented in Table 2. The clinical descriptions, histories and types of epilepsy are presented in Appendix S1 and Table S1, Supporting information. Positions were mapped according to the hg19 assembly of the UCSC genome browser. The rearrangements were *de novo* in all five patients in whom parents' samples were available (Table 1, Fig. 2) and the mean size was 408 kb (309–654 kb). The genes involved in the duplications are presented in Fig. 2.

Table 1. Cytogenetic and clinical characteristics of female patients with an intrachromosomal duplication involving the *MECP2* gene

Patients	Reardon et al. (11)	Grasshoff et al. (10)		Mayo et al. (23)	Bijlsma et al. (12)		Shimada et al. (13)
	Fam 1/Patient III4	Patient 1	Patient 2	Proband	Patient 4	Patient 5	Patient 4
<b>Cytogenetic characteristics</b>							
Cytogenetic nomenclature	arr[hg19] Xq28(152,788,477-153,706,446)x3 mat	arr[hg19] Xq28(153,239,587-153,505,448)x3 dn	arr[hg19] Xq28(153,272,227-153,749,960)x3 dn	arr[hg19] Xq28(153,277,239-153,576,940)x3 dn	arr[hg19] Xq28(152,706,073-153,455,489)x3 mat	arr[hg19] Xq28(153,261,054-153,368,556)x3 dn	arr[hg19] Xq28(152,916,694-153,576,940)x3 mat
Size of the duplication	918 kb	266 kb	478 kb	300 kb	700 kb	107.5 kb	660 kb
Inheritance	mat inherited	dn	dn	dn	mat inherited	dn	mat inherited
Parental origin	Maternal	Paternal	Paternal	Paternal	Maternal	ND	Maternal
X-inactivation (proband)	70/30	61/39	71/29	75/25	63/37	84/16	88/12
X-inactivation (mothers)	100/0				88/12		43/57
Second alteration	del 20q13.33 (900 kb, dn)	dup 2q37.3 (112 kb, dn)		2 Xq28 duplications (138 kb, 25 kb, dn)		dup 2q23.1-23.2 (824.5 kb, pat)	
<b>Clinical features</b>							
Term	ND	39 WG	40 WG	41 WG	39 WG	38 WG	38 WG
Birth W (g); L (cm); HC (cm)	ND	2900; 48; 34	3650; 52; 35.5	3055; 50; ND	2850; 48; 32	3430; 53.3; ND	2954; 46.6; 34.6
Age at examination	12 years	7 years 6 months	20 years	7 years	26 years	7 years	5 years
Height (SD)	ND	+0.3 SD	-1 SD	75-90th c	-0.3 SD	-0.5 SD	+0.9 SD
Microcephaly (<-2 SD)	ND	No (+0.7 SD)	No (0 SD)	No (0 SD)	No (+0.7 SD)	No (+1.1 SD)	No (+1.6 SD)
Hypotonia	No	Yes	No	No	No	Yes	Yes
Developmental delay	Mild	Moderate	Moderate	Severe	Mild	Moderate	Severe
Speech	Normal	First words 19 months	First words 3 years	Few words 7 years	Normal	Five-word sentences 7 years	Few words 5 years
Milestones	Walking at 28 months, special school	Walking at 19 months, not able to count or write	Walking < 18 months, not able to read or write at 10 years	Walking > 2 years	Walking < 18 months	Walking > 22 months	Walking at 30 months

Xq28 duplication including *MECP2* in six unreported affected females

579

580

Table 1. continued

Patients	Reardon et al. (11)	Grasshoff et al. (10)		Mayo et al. (23)	Bijlsma et al. (12)		Shimada et al. (13)
	Fam 1/Patient III4	Patient 1	Patient 2	Proband	Patient 4	Patient 5	Patient 4
Seizures	No	No	No	Yes <sup>a</sup>	No	No	No
Severe feeding problems	ND	No	No	ND	No	ND	ND
Constipation	ND	Yes	ND	ND	Yes	Yes	ND
Stereotyped movements	No	No	ND	Yes	No	ND	ND
Recurrent infections	No	Yes	No	ND	No	Yes (otitis media)	Yes (pneumonia and bronchitis)
Dysmorphic features	No	No	No	Yes	No	No	Yes
Other	Learning difficulties	Bilateral simian creases, sacral hemangioma, autistic features, learning difficulties	Lower limbs spasticity, vesico-ureteral reflux, autistic features	Occasionally no sphincter control, eats her own hair	Poor interactions, learning difficulties, psychiatric features (depression, compulsion), fifth finger clinodactyly	Eczema, learning difficulties, social interaction problem	Ataxia
Brain MRI	Unremarkable at 10 years of age	ND	Normal	Normal	NA	Normal	Posterior periventricular cystic lesions
Patients	Fieremans et al. (14)		Scott et al. (15)		Novara et al. (16)		
	Patient 1	Patient 2	Proposita	Proposita's sister	Patient 1	Patient 2	Patient 3
<b>Cytogenetic characteristics</b>							
Cytogenetic nomenclature	arr[hg19] Xq28(153,247,624-153,688,854)x3 dn	arr[hg19] Xq28(152,228,318-153,688,404)x3 dn	arr[hg19] Xq28(153,279,814-153,569,878)x3 mat	arr[hg19] Xq28(153,279,814-153,569,878)x3 mat	arr[hg19] Xq28(153,238,518-153,406,174)x3 mat	arr[hg19] Xq28(152,987,955-153,609,113)x3 mat	arr[hg19] Xq28(153,015,806-153,405,806)x3 dn
Size of the duplication	0.44 Mb	1.46 Mb	290 kb	290 kb	167 kb	621 kb	390 kb
Inheritance	dn	dn	mat inherited	mat inherited	mat inherited	mat inherited	dn
Parental origin	Paternal	Maternal	Maternal	Maternal	Maternal	Maternal	Paternal
X-inactivation (proband)	100/0	100/0	Uninformative	Uninformative	Random in blood and saliva	100/0 in blood, random in saliva	Skewed in blood and saliva
X-inactivation (mothers)	100/0	65/35	100/0	100/0	Random in blood and saliva	100/0 in blood, random in saliva	ND

El Chehadeh et al.

Table 1. continued

Patients	Fieremans et al. (14)		Scott et al. (15)		Novara et al. (16)		
	Patient 1	Patient 2	Proposita	Proposita's sister	Patient 1	Patient 2	Patient 3
Second alteration							
Clinical features							
Term	ND	ND	41 WG	41 WG	40 WG	39 WG	40 WG
Birth W (g); L (cm); HC (cm)	ND	ND	ND	ND	2400; 49; 36	3000; NA; 35	3780; 51; 36
Age at examination	13 years	12 years	25 years	25 years	School age	15 years	17 years
Height (SD)	90th c	0 SD	10th c	50th c	ND	ND	ND
Microcephaly (<-2 SD)	No (25th c)	No (30th c)	No (98th c)	No (50th c)	ND	ND	ND
Hypotonia	No	NA	Yes	Yes	Yes	Yes	Yes
Developmental delay	Mild (IQ: 63)	Severe	Severe	Profound	Moderate	Mild	Mild
Speech	Normal speech	Speaks a few words at 12 years	One word responses	No speech	Few words at 5 years	First words 4 years	First words 3 years
Milestones	Early developmental milestones were normal	Walking at 26 months	ND	Walking at 7 years	Walking at 18 months	Walking at 13 months	Walking at 15 months
Seizures	No	No	Yes <sup>a</sup>	Yes <sup>a</sup>	No	No	No
Severe feeding problems	No	Yes	ND	ND	No	No	Yes
Constipation	No	No	ND	ND	ND	ND	ND
Stereotyped movements	Yes	Yes	Yes	Yes	No	No	No
Recurrent infections	No	No	No	Yes (aspiration pneumonia)	No	No	No
Dysmorphic features	Yes	No	Yes	Yes	Yes	Yes	No
Other	Mild tremor of the hands, learning difficulties, moist hands	Persistent gastro-oesophageal reflux, sleep problems, autistic behaviour, learning difficulties, narrow feet, limited knee extension	Wide-based gait, spastic quadriparesis and hyperreflexia, scoliosis, uncontrollable bursts of laughter	Wide-based gait, mild spastic quadriparesis and hyperreflexia, scoliosis, uncontrollable bursts of laughter, skills estimated at 3-4 years old level, loss of the ability to walk or sit without support during teenage years	Motor and coordination difficulties, joint laxity, learning difficulties, hyperactivity with attention difficulties	Motor and coordination difficulties, joint laxity, learning difficulties, hyperactivity with attention difficulties, anxiety	Motor and coordination difficulties, joint laxity, learning difficulties, hyperactivity with attention difficulties, depressive mood

Xq28 duplication including MECP2 in six unreported affected females

581

Table 1. continued

Patients	Patient 1	Fieremans et al. (14)		Scott et al. (15)		Novara et al. (16)	
		Patient 2	Patient 2	Proposita	Proposita's sister	Patient 1	Patient 2
Brain MRI	Normal	Cavum septi pellucidi and vergae	Normal	Normal at 9 years, moderate cerebellar volume loss at 16 years	Normal	NA	Normal
Present study							
Patients	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6	
Cytogenetic characteristics							
Cytogenetic nomenclature	arr[hg19] Xq28(153,093,533-153,747,928)x3 dn	arr[hg19] Xq28(153,277,239-153,586,443)x3 dn	arr[hg19] Xq28(153,287,517-153,609,163)x3 dn	arr[hg19] Xq28(153,287,517-153,609,163)x3 dn	arr[hg19] Xq28(153,238,318-153,609,313)x3 dn	arr[hg19] Xq28(153,231,264-153,613,200)x3	
Size of the duplication	654 kb	309 kb	322 kb	322 kb	371 kb	382 kb	
Inheritance	dn	dn	dn	dn	dn	NA (adopted)	
Parental origin	ND	Paternal	ND	ND	ND	NA	
X-inactivation (proband)	93/7	95/5	76/24	76/24	56/44	81/19	
X-inactivation (mothers)	50/50	Random	Random	Random	70/30	NA	
Second alteration		dup Xp11.23 (396 Kb, dn)					
Clinical features							
Term	39 WG	40 WG	37 WG	37 WG	35 WG	39 WG	
Birth W (g); L (cm); HC (cm)	3250; 50; ND	3600; 50	2260; 46; 33	2490; 46; 32	2120; 45; 31	NA	
Age at examination	25 years	40 years	12 years	12 years	22 years	4 years 6 months	
Height (SD)	+1 SD	-0.5 SD	+0.5 SD	+0.5 SD	-1.5 SD	+1 SD	
Microcephaly (<-2 SD)	No (+1 SD)	No (+2 SD)	No (+1 SD)	No (+1 SD)	No (+1 SD)	No (+2 SD)	
Hypotonia	Yes	No	Yes	No	Yes	Yes	
Developmental delay	Moderate	Moderate	Moderate	Moderate	Severe	Moderate	
Speech	Few-word sentences	Speech delay, short sentences	Speech delay (4 years), short sentences	Speech delay (3.5 years), short sentences	Less than 10 words	Speech delay, 10 words barely intelligible	
Milestones	Walking at 3 years, sitting at 18 months	Walking at 18 months	Sitting at 9 months, walking 17 months	Sitting at 9 months, walking at 15 months	Walking at 22 months	Sitting at 20 months, walking at 26 months	
Seizures	Yes <sup>a</sup>	No	No	No	Yes <sup>a</sup>	No	

El Chehadeh et al.

582

Xq28 duplication including *MECP2* in six unreported affected females

Table 1, continued

Patients	Present study					
	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6
Severe feeding problems	No	No	No	No	ND	No
Constipation	Yes	No	No	No	ND	No
Stereotyped movements	Yes	No	Yes	Yes	Yes	Yes
Recurrent infections	Yes < 5 years	Yes (tonsillitis)	Yes (tonsillitis)	Yes (tonsillitis)	Yes (otitis media and tonsillitis)	No
Dysmorphic features	No	No	No	No	Yes	Yes
Other	Hypermetropia, flat feet, mild steppage and wide-based gait with flexed knees, obstructive apnoea, persistence of deciduous teeth, thick hair, obesity, slender fingers and toes	Pulmonary lymphangioleiomyomatosis, flat feet, 5th finger clinodactyly, sandal gap	Learning difficulties, 1 café au lait spot, persistent deciduous teeth, anxiety, sensitive to noise	Mild learning difficulties, persistent deciduous teeth, anxiety, sensitive to noise, finger pads	Tetralogy of Fallot, anxiety, motor restlessness, coordination difficulties, behavioural troubles (aggressiveness, bursts of laughter), wide-based gait with flexed knees, scoliosis	Wide-based and ataxic gait
Brain MRI	Atrophy of the frontal lobes, dilation of the LV	Not done (TDM normal)	Normal	Normal	Not done (TDM normal)	Normal

dn, *de novo*; Dup Xq, intrachromosomal Xq28 duplication involving the *MECP2* gene; HC, head circumference; IQ, intelligence quotient; L, length; LV, lateral ventricles; mat, maternally; NA, not available; ND, not determined; SD, standard deviation, W, weight; WG, weeks of gestation.

<sup>a</sup>Details about seizures are presented in Table S1.

Table 2. Molecular cytogenetic characteristics of dupMECP2 in female patients in this study

Patients	Chromosomal rearrangement	Cytogenetic nomenclature	Technique of detection	Array-CGH start and end positions (hg19)	Size of the Xq duplication (kb)	Gene content
Patient 1	Dup X	arr[hg19] Xq28(153,093,533–153,747,928)x3 dn	Array-CGH	153093533–153747928	654	PDZD4, L1CAM, IRAK1, MECP2, OPN1LW, OPN1MW, FLNA, RPL10, GDI1
Patient 2	Dup X	arr[hg19] Xq28(153,277,239–153,586,443)x3 dn	Array-CGH	153277239–153586443	309	IRAK1, MECP2, OPN1LW, OPN1MW, FLNA
Patient 3	Dup X	arr[hg19] Xq28(153,287,517–153,609,163)x3 dn	Array-CGH	153287517–153609163	322	MECP2, OPN1LW, OPN1MW, FLNA
Patient 4	Dup X	arr[hg19] Xq28(153,287,517–153,609,163)x3 dn	Array-CGH	153287517–153609163	322	MECP2, OPN1LW, OPN1MW, FLNA
Patient 5	Dup X	arr[hg19] Xq28(153,238,318–153,609,313)x3 dn	SNP array	153238318–153609313	371	IRAK1, MECP2, OPN1LW, OPN1MW, FLNA
Patient 6	Dup X	arr[hg19] Xq28(153,231,264–153,613,200)x3	Array-CGH	153231264–153613200	382	IRAK1, MECP2, OPN1LW, OPN1MW, FLNA

CGH, comparative genomic hybridization; dn, *de novo*; SNP, single nucleotide polymorphism.

**Discussion**

Here, we report on a new series of six sporadic females with interstitial dupMECP2 in order to further delineate the clinical spectrum of dupMECP2 in females and further discuss the difficulties of genetic counselling in such cases. These patients were identified among 99 patients with a dupMECP2 in France and Switzerland (two patients), and increased to 20 the number of females reported in the literature, compared with 140 male patients (16). A recruitment bias is probable, due to the fact that the affected females are probably more often reported because they are uncommon.

To date, 14 female patients with an intraC dupMECP2 of various sizes have been described in the literature (10–16, 23). Among these, seven (50%) had inherited the duplication from their mother (Table 1, Figs 2 and 3). Among these seven patients with an inherited intraC dupMECP2, including twins (11–13, 15, 16), four had an affected male in the family, and all of the carrier mothers were asymptomatic except one who had learning difficulties and depression (15) (Table 1 and Fig. 3). Moreover, among these seven patients, three were mildly affected, one was moderately affected, and three, including the twins, had severe to profound ID (11–13, 15, 16) (Table 1). The mean size of the duplicated segment in the reported patients with an intraC dupMECP2 was 506 kb (min. 107.5 kb, max. 1460 kb) (Table 1). It could be suggested that the size of the duplications could in part explain the differences in the phenotype among the affected females with intraC duplications. This, however, has not been proven (16), even in males (5, 27), as Shimada et al. (13) showed that the patient with the smallest duplication in their

study developed intractable epilepsy and developmental regression in early infancy (13).

It is now well-known that affected females carrying a large dupMECP2 resulting from an X;autosomal unbalanced translocation or an insertion of the duplicated Xq28 segment into an autosome can be severely affected with a phenotype as severe as that in affected males (12, 18–22), due to the fact that they are functionally disomic for genes that are normally subject to inactivation and because of the larger size of the duplications. Furthermore, some of these unbalanced translocations also result in deletions of other chromosomal material. In these cases, it has been suggested that local escape from inactivation (the X-translocated segment separated from its corresponding X-inactivation centre cannot be inactivated), expression of recessive genes from the active X, or disruption of a gene by the rearrangement could explain the abnormal phenotype (18, 24). We could also suppose that the second (autosomal) chromosomal rearrangement resulting from the imbalance could play a role in the phenotype.

In contrast, the female cases of this study, who carry an interstitial duplication, mostly presented a less severe phenotype with non-specific mild to moderate intellectual deficiency and no obvious dysmorphic features (Table 1 and Fig. 1). However, one of the six patients of the present series (Patient 5) has a moderate to severe phenotype including very poor language, drug-resistant epilepsy, chronic constipation, stereotyped movements, behavioural disorders, recurrent infections and facial dysmorphism and a random XCI (56/44) (Fig. 1, Table 1, Table S1). Other authors have also reported rare cases of severely affected females carrying an intraC dupMECP2 (13–15) or an Xq28 triplication (23). Figure 4

## Xq28 duplication including *MECP2* in six unreported affected females

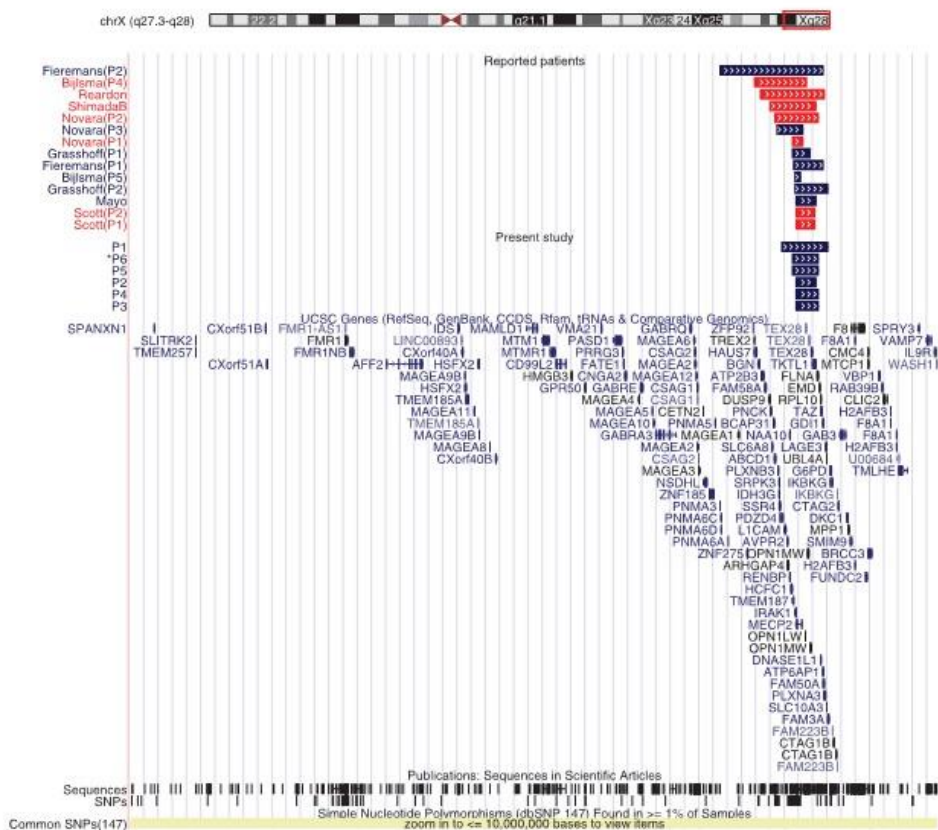


Fig. 2. Schematic representation of Xq28 region and all the affected female patients reported to date including the present series, for which the array Comparative Genomic Hybridization (array-CGH) breakpoints were available. The patients with a *de novo* duplication are represented in dark blue and patients with a maternally inherited duplication are shown in red. A star means that the parental origin is not available. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

summarizes the different clinical features in affected females with an interstitial duplication, including those in this study, as compared with the largest review of males (28, 29).

The major hypothesis to explain the symptoms in some females with dupMECP2 is abnormal XCI. It has been shown that carrier mothers, who are asymptomatic in the vast majority of cases, display extreme (>90/10) to complete (100/0) skewing of XCI with the duplication-bearing X chromosome being preferentially inactivated (1). In this study, among the six intraC dupMECP2 patients, two had an extremely skewed XCI with no information about which X chromosome was inactivated (Patient 1 and Patient 2, 93/7 and 95/5, respectively) and three patients, including twins, had moderately skewed XCI (Patient 3 and Patient 4: 76/24, Patient 6: 81/19) (Table 1, Fig. 5). It has been shown that XCI skewing can increase with age (30), which can

partly explain the absence of clear correlations between the XCI pattern and the severity of the phenotype in females. Interestingly, only the two oldest patients of the present series (Patient 2, aged 28 and 44 years, respectively) had an extremely skewed XCI (Table 1). It has also been hypothesized that the *de novo* occurrence might have an effect on the XCI, and thus affect the methylation process (10). In this series, all dupMECP2 in symptomatic females were *de novo*. The results of XCI studies are not always easy to correlate with the risk of symptoms as it has been reported that some carrier mothers with complete skewing of XCI could manifest neuro-psychiatric symptoms such as generalized anxiety/panic disorder, depression and compulsion, some of them having an intelligence quotient (IQ) in the low-to-average range (5, 12). Moreover, several mothers were found to have a random XCI pattern even though they were asymptomatic (13, 31)

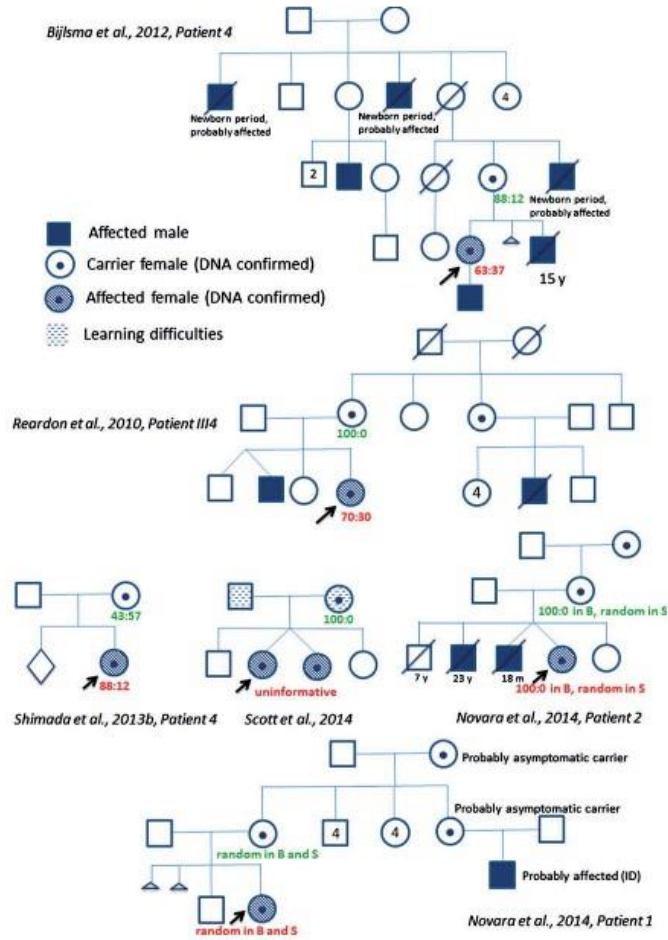


Fig. 3. Pedigrees of the seven families of affected female patients with inherited intrac dupMECP2 reported in the literature according to Reardon et al., (11) (Family 1, patient III4), Bijlsma et al., (12) (Patient 4), Shimada et al. (13) (Patient 4), Scott et al., and Novara et al. (16) (Patients 1 and 2). The X chromosome inactivation (XCI) ratios are represented in red below the probands (arrows) and in green below the carrier mothers. B, blood; S, saliva. [Color figure can be viewed at wileyonlinelibrary.com]

(Table 1 and Fig. 3). In the report of Novara et al. (Patients 1 and 2), no difference in the XCI pattern was observed between female probands and their unaffected carrier mothers (16) (Table 1). However, we believe that the XCI pattern, at least in blood, only offers incomplete information since for one mildly affected patient, XCI was 100/0 in the blood and random in saliva (16) (Table 1, Fig. 5). Another example of the fluctuation of XCI from one tissue to another is the mother of one male patient (Patient 1) studied by Kirk et al. (32) She had mild learning difficulties with synchronous strokes and 100% skewing of XCI in the blood vs a 74/26 ratio in hair roots (32). Because there is no agreement

among studies on the correlation between XCI patterns in different tissues (33), investigating X-inactivation in more than one tissue and concomitantly determining which of the two X chromosomes is inactivated could be more informative. Given these observations and as XCI is usually tested in blood, which does not reflect the cells of the entire body, we could assume the existence of mosaic XCI, with different degrees of skewing in the different tissues leading to tissue specific dosage alterations (12, 16, 32). This could raise the possibility that there is preferential activation of the abnormal X chromosome in the brain of affected females very early during embryogenesis, thus hampering normal brain development

### Xq28 duplication including *MECP2* in six unreported affected females

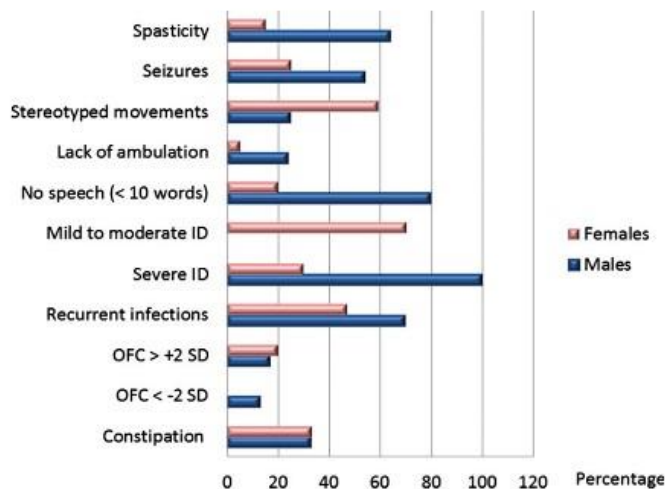


Fig. 4. Schematic representation of clinical features in affected females and males reported in the literature, including this study, carrying an interstitial dupMECP2. Proportions appear as dark blue bars in males and as pink bars in females. From Van Esch et al. (28) and El Chehadeh et al. (29). [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

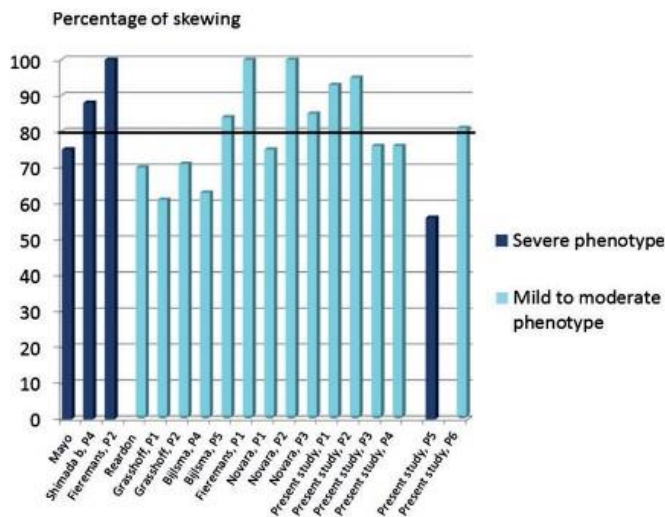


Fig. 5. Schematic representation of the X chromosome inactivation (XCI) pattern (in blood) of severely compared with mildly to moderately affected females with intraC dupMECP2, including this study. The bars correspond to XCI skewing (in %). Severely affected patients appear in dark blue and mildly to moderately affected patients appear in light blue. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

(16, 23, 32). In addition, the variability in the XCI level in the brain may explain why some carrier females are mildly affected (low-to-average cognitive scores) (5, 12), even more so since it has been suggested that increased expression of *MECP2*, in particular in the brain, could play a role in the severity of the phenotype in mice and in humans (12, 34, 35). This is supported by the description

of affected males and females carrying an *MECP2* triplication, who have a particularly severe phenotype (2, 23, 36). All these observations make genetic counselling complex in at-risk couples.

In conclusion, these findings are of importance for genetic counselling because they show that an abnormal phenotype, and even a severe one, is possible in females



born to asymptomatic carrier mothers, and that this phenotype does not necessarily correlate with the XCI pattern in the blood or the size of the duplication. This makes genetic counselling difficult in term of prognosis, in particular in prenatal cases. A large epidemiological study of the phenotypes of at-risk females in families with dupMECP2 is needed in order to determine the frequency of symptoms associated with a dupMECP2 in females and to give a more accurate risk assessment to families for genetic and prenatal counselling.

### Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web-site.

### Acknowledgements

We are deeply grateful to the patients and their parents for their participation in this study. We thank Philip Bastable from the 'Pôle de Recherche' of Dijon University Hospital for a helpful review of this article. We thank the Regional Council of Burgundy and the association Xtraordinaire for their financial support of the project, and the PHRC RMLX (Pr V. Deportes) for its support.

### References

- Van Esch H, Bauters M, Ignatiou J et al. Duplication of the MECP2 region is a frequent cause of severe mental retardation and progressive neurological symptoms in males. *Am J Hum Genet* 2005; 77: 442–453.
- Del Gaudio D, Fang P, Scaglia F et al. Increased MECP2 gene copy number as the result of genomic duplication in neurodevelopmentally delayed males. *Genet Med* 2006; 8: 784–792.
- Clayton-Smith J, Walters S, Hobson E et al. Xq28 duplication presenting with intestinal and bladder dysfunction and a distinctive facial appearance. *Eur J Hum Genet* 2009; 17: 434–443.
- Lugtenberg D, Kleefstra T, Oudakker AR, Nillesen WM, Yntema HG et al. Structural variation in Xq28: MECP2 duplications in 1% of patients with unexplained XLMR and in 2% of male patients with severe encephalopathy. *Eur J Hum Genet* 2009; 17: 444–453.
- Ramocki MB, Peters SU, Tavyev YJ et al. Autism and other neuropsychiatric symptoms are prevalent in individuals with MeCP2 duplication syndrome. *Ann Neurol* 2009; 66: 771–782.
- Adkins NL, Georget PT. MeCP2: structure and function. *Biochem Cell Biol* 2011; 89 (1): 1–11.
- Gonzales ML, LaSalle JM. The role of MeCP2 in brain development and neurodevelopmental disorders. *Curr Psychiatry Rep* 2010; 12 (2): 127–134.
- Chahrouh M, Jung SY, Shaw C et al. MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science* 2008; 320: 1224–1229.
- Nageshappa S, Carroumeu C, Trujillo CA et al. Altered neuronal network and rescue in a human MECP2 duplication model. *Mol Psychiatry* 2016; 21 (2): 178–188.
- Grasshoff U, Bonin M, Gochring I et al. De novo MECP2 duplication in two females with random X-inactivation and moderate mental retardation. *Eur J Hum Genet* 2011; 19 (5): 507–512.
- Reardon W, Donoghue V, Murphy AM et al. Progressive cerebellar degenerative changes in the severe mental retardation syndrome caused by duplication of MECP2 and adjacent loci on Xq28. *Eur J Pediatr* 2010; 169 (8): 941–949.
- Bijlsma EK, Collins A, Papa FT et al. Xq28 duplications including MECP2 in five females: expanding the phenotype to severe mental retardation. *Eur J Med Genet* 2012; 55 (6–7): 404–413.

- Shimada S, Okamoto N, Ito M et al. MECP2 duplication syndrome in both genders. *Brain Dev* 2013; 35 (5): 411–419.
- Fieremans N, Bauters M, Belet S et al. De novo MECP2 duplications in two females with intellectual disability and unfavourable complete skewed X-inactivation. *Hum Genet* 2014; 133 (11): 1359–1367.
- Scott Schwoerer J, Laffin J, Haun J, Raca G, Friez MJ, Giampietro PF. MECP2 duplication: possible cause of severe phenotype in females. *Am J Med Genet A* 2014; 164: 1029–1034.
- Novara F, Simonati A, Sicca F et al. MECP2 duplication phenotype in symptomatic females: report of three further cases. *Mol Cytogenet* 2014; 7 (1): 10.
- Bialer MG, Anguiano A, Taff I, Shanmughan A, LaGrave D, White BJ. De novo trisomy Xq28-qter detected by subtelomeric FISH screening. *Am J Hum Genet* 2003; 73: 30.
- Lachlan KL, Collinson MN, Sandford RO, van Zyl B, Jacobs PA, Thomas NS. Functional disomy resulting from duplications of distal Xq in four unrelated patients. *Hum Genet* 2004; 115: 399–408.
- Sanlaville D, Prieur M, de Blois MC et al. Functional disomy of the Xq28 chromosome region. *Eur J Hum Genet* 2005; 13: 579–585.
- Auber B, Burfeind P, Thiels C et al. An unbalanced translocation resulting in a duplication of Xq28 causes a Rett syndrome-like phenotype in a female patient. *Clin Genet* 2010; 77 (6): 593–597.
- Makrythanasis P, Moix I, Gimelli S et al. De novo duplication of MECP2 in a girl with mental retardation and no obvious dysmorphic features. *Clin Genet* 2010; 78 (2): 175–180.
- Shimada S, Okamoto N, Hirasawa K et al. Clinical manifestations of Xq28 functional disomy involving MECP2 in one female and two male patients. *Am J Med Genet A* 2013; 161A (7): 1779–1785.
- Mayo S, Monfort S, Roselló M et al. De novo interstitial triplication of MECP2 in a girl with neurodevelopmental disorder and random X chromosome inactivation. *Cytogenet Genome Res* 2011; 135 (2): 93–101.
- Sanlaville D, Schluth-Bolard C, Turleau C. Distal Xq duplication and functional Xq disomy. *Orphanet J Rare Dis* 2009; 4: 4.
- Kubota T. A new assay for the analysis of X-inactivation in carriers with an X-linked disease. *Brain Dev* 2001; 23 (Suppl. 1): S177–S181.
- Lau AW, Brown CJ, Peñaherrera M, Langlois S, Kalousek DK, Robinson WP. Skewed X-chromosome inactivation is common in fetuses or newborns associated with confined placental mosaicism. *Am J Hum Genet* 1997; 61: 1353–1361.
- Bauters M, Van Esch H, Friez MJ et al. Nonrecurrent MECP2 duplications mediated by genomic architecture-driven DNA breaks and break-induced replication repair. *Genome Res* 2008; 18 (6): 847–858.
- Van Esch H. MECP2 duplication syndrome. *Mol Syndromol* 2012; 2 (3–5): 128–136.
- El Chehadé S, Faivre L, Mosca-Boidron AL et al. Large national series of patients with Xq28 duplication involving MECP2: delineation of brain MRI abnormalities in 30 affected patients. *Am J Med Genet A* 2016; 170A (1): 116–129.
- Amos-Landgraf MJ, Cottle A, Plenge RM et al. X chromosome-inactivation patterns of 1,005 phenotypically unaffected females. *Am J Hum Genet* 2006; 79 (3): 493–499.
- Zhang Q, Zhao Y, Yang Y, Bao X. MECP2 duplication syndrome in a Chinese family. *BMC Med Genet* 2015; 16: 112.
- Kirk EP, Malaty-Brevaud V, Martini N et al. The clinical variability of the MECP2 duplication syndrome: description of two families with duplications excluding L1CAM and FLNA. *Clin Genet* 2009; 75: 301–303.
- Viggiano E, Ergoli M, Picillo E, Politano L. Determining the role of skewed X-chromosome inactivation in developing muscle symptoms in carriers of Duchenne muscular dystrophy. *Hum Genet* 2016; 135 (7): 685–698. Review.
- Collins AL, Levenson JM, Vilaythong AP et al. Mild overexpression of MECP2 causes a progressive neurological disorder in mice. *Hum Mol Genet* 2004; 13: 2679–2689.
- Na ES, Nelson ED, Adachi M et al. A mouse model for MeCP2 duplication syndrome: MeCP2 overexpression impairs learning and memory and synaptic transmission. *J Neurosci* 2012; 32: 3109–3117.
- Tang SS, Fernandez D, Lazarou LP, Singh R, Fallon P. MECP2 triplication in 3 brothers – a rarely described cause of familial neurological regression in boys. *Eur J Paediatr Neurol* 2012; 16 (2): 209–212.

## 2. Caractérisation des anomalies à l'IRM cérébrale chez 30 patients atteints de MECP2DS

### Résumé

Nous rapportons les données d'imagerie par résonance magnétique (IRM) cérébrale standardisée de 30 patients porteurs d'une duplication Xq28 impliquant le gène *MECP2* de différentes tailles (228 kb à 11,7 Mb). L'objectif de cette étude était de rechercher des malformations récurrentes et d'essayer de déterminer si les variations des caractéristiques d'imagerie pouvaient être expliquées par des différences dans la taille des duplications. Nous avons montré que 93% des patients présentaient des anomalies à l'IRM cérébrale telles que des anomalies du corps calleux (n = 20), une réduction du volume de la substance blanche (n = 12), une dilatation des ventricules latéraux (n = 9), la présence d'hypersignaux T2 périventriculaire anormaux (n = 6), et une hypoplasie du vermis (n = 5). La mesure du périmètre crânien variait considérablement entre > +2 DS chez cinq patients et < -2 DS chez quatre patients. Parmi les neuf patients présentant une dilatation des ventricules latéraux, six présentaient une duplication impliquant le gène *L1CAM* (impliqué dans l'hydrocéphalie liée à l'X avec pouces adductus). Le seul patient présentant des hétérotopies sous-épendymaires postérieures bilatérales était également porteur d'une duplication du gène *FLNA* (impliqué dans une forme liée à l'X d'hétérotopies nodulaires péri-ventriculaires). Nous n'avons pas pu démontrer de corrélation entre les hyperintensités périventriculaires de la substance blanche et le retard de myélinisation et la duplication du gène *IKBKG*. Nous concluons donc que les patients présentant une duplication du gène Xq28 impliquant *MECP2* partagent des anomalies à l'imagerie cérébrale similaires mais non spécifiques. Ces caractéristiques d'imagerie ne peuvent donc pas constituer un critère diagnostique. La corrélation génotype-phénotype n'a pas permis de mettre en évidence une relation entre la présence d'hétérotomie nodulaire, de dilatation ventriculaire, d'anomalies de la substance blanche et la présence de *FLNA*, *L1CAM* ou *IKBKG*, respectivement, dans le segment dupliqué.

## Large National Series of Patients With Xq28 Duplication Involving *MECP2*: Delineation of Brain MRI Abnormalities in 30 Affected Patients

Salima El Chehadeh,<sup>1,2\*</sup> Laurence Faivre,<sup>1,2</sup> Anne-Laure Mosca-Boidron,<sup>2,3</sup> Valérie Malan,<sup>4</sup> Jeanne Amiel,<sup>5</sup> Mathilde Nizon,<sup>5</sup> Renaud Touraine,<sup>6</sup> Fabienne Prieur,<sup>6</sup> Laurent Pasquier,<sup>7</sup> Patrick Callier,<sup>2,3</sup> Mathilde Lefebvre,<sup>1,2</sup> Nathalie Marle,<sup>2,3</sup> Christèle Dubourg,<sup>7</sup> Sophie Julia,<sup>8</sup> Catherine Sarret,<sup>9</sup> Christine Francannet,<sup>9</sup> Fanny Laffargue,<sup>9</sup> Odile Boespflug-Tanguy,<sup>10</sup> Albert David,<sup>11</sup> Bertrand Isidor,<sup>11</sup> Cédric Le Caignec,<sup>11</sup> Jacqueline Vigneron,<sup>12</sup> Bruno Leheup,<sup>12</sup> Laetitia Lambert,<sup>12</sup> Christophe Philippe,<sup>13</sup> Jean-Marie Cuisset,<sup>14</sup> Joris Andrieux,<sup>15</sup> Ghislaine Plessis,<sup>16</sup> Annick Toutain,<sup>17</sup> Alice Goldenberg,<sup>18</sup> Valérie Cormier-Daire,<sup>5</sup> Marlène Rio,<sup>5</sup> Jean-Paul Bonnefont,<sup>19</sup> Julien Thevenon,<sup>1,2</sup> Bernard Echenne,<sup>20</sup> Hubert Journal,<sup>21</sup> Alexandra Afenjar,<sup>22</sup> Lydie Burglen,<sup>22</sup> Thierry Bienvvenu,<sup>23</sup> Marie-Claude Addor,<sup>24</sup> Sébastien Lebon,<sup>25</sup> Danièle Martinet,<sup>26</sup> Clarisse Baumann,<sup>27</sup> Laurence Perrin,<sup>27</sup> Séverine Drunat,<sup>28</sup> Pierre-Simon Jouk,<sup>29</sup> Françoise Devillard,<sup>29</sup> Charles Coutton,<sup>29</sup> Didier Lacombe,<sup>30</sup> Marie-Ange Delrue,<sup>30</sup> Nicole Philip,<sup>31</sup> Anne Moncla,<sup>32</sup> Catherine Badens,<sup>33</sup> Nathalie Perreton,<sup>34</sup> Alice Masurel,<sup>1</sup> Christel Thauvin-Robinet,<sup>1,2</sup> Vincent Des Portes,<sup>35</sup> and Laurent Guibaud<sup>36</sup>

<sup>1</sup>FHU TRANSLAD, Centre de Référence Maladies Rares "Anomalies du Développement et Syndromes Malformatifs" de l'Est, Centre de Génétique, CHU de Dijon, France

<sup>2</sup>GAD, EA4271, Génétique et Anomalies du Développement, Université de Bourgogne, Dijon, France

<sup>3</sup>Service de Cytogénétique, CHU de Dijon, France

<sup>4</sup>Service de Cytogénétique, Hôpital Necker Enfants Malades, Paris, France

<sup>5</sup>Service de Génétique Clinique, Hôpital Necker Enfants Malades, Paris, France

<sup>6</sup>Service de Génétique Clinique Chromosomique et Moléculaire, CHU de Saint-Etienne, France

<sup>7</sup>Service de Génétique Clinique, CLAD Ouest, CHU de Rennes, France

<sup>8</sup>Service de Génétique Médicale, CHU de Toulouse, France

<sup>9</sup>Service de Génétique Médicale, CHU de Clermont-Ferrand, France

<sup>10</sup>Service de Neuropédiatrie et Maladies Métaboliques, Hôpital Robert Debré, Paris, France

<sup>11</sup>Service de Génétique Médicale, CHU de Nantes, France

<sup>12</sup>Service de Génétique Médicale, CHU de Nancy, France

<sup>13</sup>Laboratoire de Génétique Médicale, CHU de Nancy, France

<sup>14</sup>Service de Neuropédiatrie, CHRU de Lille, France

<sup>15</sup>Laboratoire de Génétique Médicale, Hôpital Jeanne de Flandre, CHRU de Lille, France

<sup>16</sup>Service de Génétique, CHU de Caen, France

The authors have declared no conflicts of interest.

Grant sponsor: Regional Council of Burgundy (PARI 2012); Grant sponsor: Association Xtraordinaire.

\*Correspondence to:

Salima El Chehadeh, M.D., Service de Génétique Médicale, Institut de Génétique Médicale d'Alsace (IGMA), Centre de Référence Maladies Rares "Anomalies du Développement et Syndromes Malformatifs" de l'Est, Hôpitaux Universitaires de Strasbourg, Hôpital de Hautepierre, Strasbourg, France.

E-mail: salima.elchehadeh@chru-strasbourg.fr

Article first published online in Wiley Online Library

(wileyonlinelibrary.com): 00 Month 2015

DOI 10.1002/ajmg.a.37384

- <sup>17</sup>Service de Génétique, CHRU de Tours, France
- <sup>18</sup>Service de Génétique Médicale, CHU de Rouen, France
- <sup>19</sup>Laboratoire de Biologie Moléculaire, Hôpital Necker Enfants Malades, Paris, France
- <sup>20</sup>Service de Neurologie Pédiatrique, CHU de Montpellier, France
- <sup>21</sup>Service de Génétique, Centre Hospitalier de Vannes, Vannes, France
- <sup>22</sup>Service de Génétique, Hôpital Armand Trousseau, Paris, France
- <sup>23</sup>Laboratoire de Biochimie et Génétique Moléculaire, GH Cochin-Broca-Hôtel Dieu, Paris, France
- <sup>24</sup>Service de Génétique Médicale, CHUV Lausanne, Lausanne, Suisse
- <sup>25</sup>Unité de Neuropédiatrie, CHUV de Lausanne, Lausanne, Suisse
- <sup>26</sup>Laboratoire de Cytogénétique Constitutionnelle et Prénatale, CHUV de Lausanne, Lausanne, Suisse
- <sup>27</sup>Service de Génétique Clinique, Hôpital Robert Debré, Paris, France
- <sup>28</sup>Laboratoire de Biologie Moléculaire, Hôpital Robert Debré, Paris, France
- <sup>29</sup>Département de Génétique et Procréation—UMR CNRS 5525 TIMC-IMAG équipe DYCTIM, CHU Grenoble, France
- <sup>30</sup>Service de Génétique Clinique, CHU de Bordeaux, France
- <sup>31</sup>Département de Génétique Médicale, Hôpital de la Timone, Marseille, France
- <sup>32</sup>Laboratoire de Génétique Chromosomique, Hôpital de la Timone, Marseille, France
- <sup>33</sup>Laboratoire de Biologie Moléculaire, Hôpital de la Timone, Marseille, France
- <sup>34</sup>EPICIME-CIC 1407 de Lyon, Inserm, Service de Pharmacologie Clinique, CHU de Lyon, Bron, France
- <sup>35</sup>Service de Neurologie Pédiatrique, CHU de Lyon-GH Est, Bron, France
- <sup>36</sup>Service de Radiologie, Hôpital Femme Mère Enfant, Bron, France

Manuscript Received: 19 August 2014; Manuscript Accepted: 7 September 2015

Xq28 duplications encompassing *MECP2* have been described in male patients with a severe neurodevelopmental disorder associated with hypotonia and spasticity, severe learning disability, stereotyped movements, and recurrent pulmonary infections. We report on standardized brain magnetic resonance imaging (MRI) data of 30 affected patients carrying an Xq28 duplication involving *MECP2* of various sizes (228 kb to 11.7 Mb). The aim of this study was to seek recurrent malformations and attempt to determine whether variations in imaging features could be explained by differences in the size of the duplications. We showed that 93% of patients had brain MRI abnormalities such as corpus callosum abnormalities ( $n = 20$ ), reduced volume of the white matter (WM) ( $n = 12$ ), ventricular dilatation ( $n = 9$ ), abnormal increased hyperintensities on T2-weighted images involving posterior periventricular WM ( $n = 6$ ), and vermis hypoplasia ( $n = 5$ ). The occipitofrontal circumference varied considerably between  $>+2SD$  in five patients and  $<-2SD$  in four patients. Among the nine patients with dilatation of the lateral ventricles, six had a duplication involving *LICAM*. The only patient harboring bilateral posterior subependymal nodular heterotopia also carried an *FLNA* gene duplication. We could not demonstrate a correlation between periventricular WM hyperintensities/delayed myelination and duplication of the *IKBK* gene. We thus conclude that patients with an Xq28 duplication involving *MECP2* share some similar but non-specific brain abnormalities. These imaging features, therefore, could not constitute a diagnostic clue. The genotype–phenotype correlation failed to demonstrate a relationship between the presence of nodular heterotopia, ventricular dilatation, WM

#### How to Cite this Article:

El Chehadeh S, Faivre L, Mosca-Boidron A-L, Malan V, Amiel J, Nizon M, Touraine R, Prieur F, Pasquier L, Callier P, Lefebvre M, Marle N, Dubourg C, Julia S, Sarret C, Francannet C, Laffargue F, Boespflug-Tanguy O, David A, Isidor B, Caignec CL, Vigneron J, Leheup B, Lambert L, Philippe C, Cuisset J-M, Andrieux J, Plessis G, Toutain A, Goldenberg A, Cormier-Daire V, Rio M, Bonnefont J-P, Thevenon J, Echenne B, Journel H, Afenjar A, Burglen L, Bienvenu T, Addor M-C, Lebon S, Martinet D, Baumann C, Perrin L, Drunat S, Jouk P-S, Devillard F, Coutton C, Lacombe D, Delrue M-A, Philip N, Moncla A, Badens C, Perretton N, Masurel A, Thauvin-Robinet C, Portes VD, Guibaud L. 2015. Large national series of patients with Xq28 duplication involving *MECP2*: Delineation of brain MRI abnormalities in 30 affected patients. *Am J Med Genet Part A* 9999A:1–14.

abnormalities, and the presence of *FLNA*, *LICAM*, or *IKBK*G, respectively, in the duplicated segment. © 2015 Wiley Periodicals, Inc.

**Key words:** Xq28 duplication; magnetic resonance imaging; genotype–phenotype correlation; *MECP2* gene

## INTRODUCTION

Duplications of Xq28 including the *MECP2* gene have been described primarily in male patients with severe developmental delay, progressive spasticity, epilepsy, stereotyped hand movements, and recurrent infections [Van Esch et al., 2005]. Most patients have inherited the duplication from their unaffected mothers and the size of the duplication varies from one patient to another. Female carriers of Xq28 duplication involving *MECP2* are usually asymptomatic, the main hypothesis to date being skewed X-chromosome inactivation (>80:20) with preferential inactivation of the rearranged chromosome. Few symptomatic females carriers have been described so far, with, in some cases, a severe phenotype, mostly secondary to unbalanced X-autosome translocation or de novo duplication with unbiased X inactivation [Bijlsma et al., 2012; Shimada et al., 2013]. It is well known from small duplications that high levels of MeCP2 are probably sufficient to cause the severe phenotype in these patients, despite the presence of several other genes, including known intellectual disability (ID)-related genes, in larger duplications. Such ID-related genes include *SLC6A8* (MIM 300036), *LICAM* (MIM 308840), *FLNA* (MIM 300017), and *GDI1* (MIM 300104), which are, respectively, responsible for X-linked creatine transporter deficiency, hydrocephalus with stenosis of the aqueduct of Sylvius, several phenotypes including isolated periventricular nodular heterotopia, and X-linked non-syndromic ID [Bauters et al., 2008; Van Esch, 2011].

Brain magnetic resonance imaging (MRI) abnormalities have rarely been described in patients with *MECP2* duplication syndrome [Echene et al., 2009; Reardon et al., 2010; Honda et al., 2012; Philippe et al., 2013]. The goal of this article is to describe brain abnormalities observed on MRI in 30 patients carrying an Xq28 duplication involving *MECP2*, and to seek a genotype–imaging phenotype correlation.

## MATERIALS AND METHODS

The patients were from an epidemiological study that identified 101 patients (92 males and 9 females ranging from 0 to 49 years) carrying an Xq28 duplication involving *MECP2* in France (98) and Switzerland (3), detected through targeted analysis (MLPA) (63% of patients) or array-CGH analysis (37%). Clinical and molecular data were systematically gathered.

For the purpose of the study, brain MRI data, when available for the patients of this series, were collected. The MRI were reinterpreted according to a standardized protocol by the same referent pediatric neuroradiologist (LG).

We also searched for correlations between the presence of various brain malformations, such as nodular heterotopia, ventricular dilatation, white matter (WM) abnormalities (WM hyper-

intensities/delayed myelination), and the presence of *FLNA*, *LICAM*, or *IKBK*G in the duplications.

## RESULTS

We were able to collect brain MRI scans from 30 patients (26 families), including five symptomatic females. The age of the patients at the time of the MRI varied from 0 months to 21 years and 8 months (means  $\pm$  SD for age: 5 years and 10 months  $\pm$  5 years and 3 months). We also collected antenatal brain MRI for two siblings (Tables I and II).

### Characteristics of Xq28 Duplications

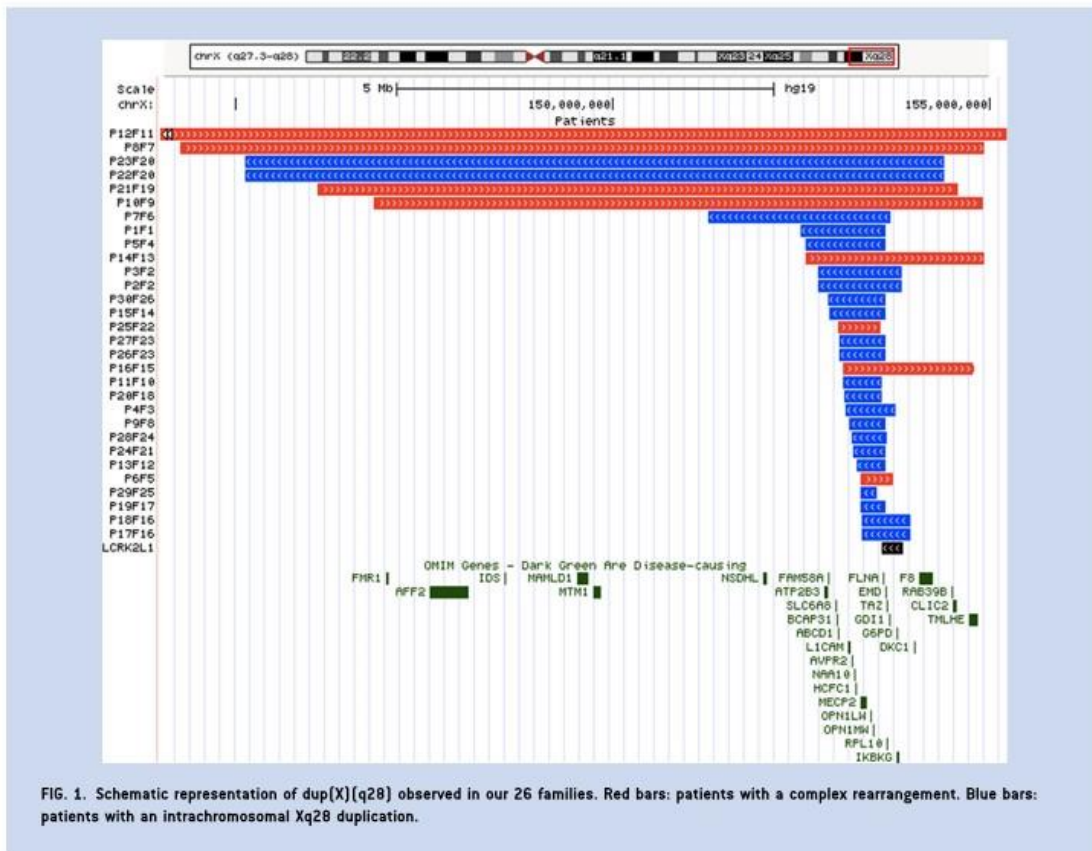
Intrachromosomal Xq28 duplications were found in 23 patients, and complex chromosome rearrangements were identified in the seven remaining patients including four patients with an unbalanced translocation. Among these, two patients had a de novo unbalanced X;9 and X;5 translocation resulting in a 9p, and a 5q microdeletion of 225 kb and 1.13 Mb, respectively. The size of the duplicated Xq28 segment in the 30 patients ranged from 228 kb to 11.7 Mb (Table III, Fig. 1).

### Occipitofrontal Circumference

The mean occipitofrontal circumference (OFC) was  $+0.15 \pm 1.78$  standard deviations (SD) (minimum =  $-3.5$  SD, maximum =  $+3.3$  SD). Four patients had microcephaly (OFC  $< -2$ SD) and five patients had macrocephaly (OFC  $> +2$ SD). It is interesting to note that among the seven patients with global lateral ventricle dilatation, only one had an OFC over  $+2$ SD (Tables I and II, Fig. 2).

### MRI Abnormalities

Only two patients, one severely affected male and one mildly affected female, had a normal brain MRI. Twenty-five of the thirty patients (83%) had  $\geq 2$  brain MRI abnormalities. Patients with Xq28 duplication involving *MECP2* shared several similar brain abnormalities. In the majority of cases, these involved the supratentorial structures and included corpus callosum dysgenesis, in 20/30 patients (67%), including hypoplasia in 12 patients (Fig. 3A), short but complete corpus callosum in eight patients (Fig. 3B), defective modeling of the genu in two patients, partial agenesis (non-visualization of the beak) in one case (Fig. 3C), and complete agenesis in one case (Fig. 3D) (Tables I and II). The other frequent abnormalities included reduced white matter volume in 12/30 (40%) patients (Fig. 4A), dilatation of the lateral ventricles in 9/30 (30%) patients (Fig. 4B), persistence of the cavum septum pellucidum in 12/30 (40%) patients (Fig. 4B), abnormal increased hyperintensities on T2-weighted images involving posterior periventricular WM in 6/30 (20%) patients and associated with dilatation of the Virchow Robin spaces in one patient (P28F24) (Fig. 5A), and physiological delay of WM myelination in 9/30 (30%) patients, which was particularly marked in two of them (Fig. 5B) (Tables I and II). Additional supratentorial malformations were also shown and included bilateral posterior asymmetric



subependymal nodular heterotopia (P5F4) (Fig. 6A), persistent subdural hygroma (P20F18) (Fig. 6B), and the lack of internal capsule myelination (P8F7) (Fig. 6C, Tables I and II).

Regarding infra-tentorial malformations, 11/30 (37%) patients had an increased fluid-filled retrocerebellar space (Fig. 7A), which was noted on prenatal brain MRI in one patient (P23F20) (Fig. 7B). Cerebellar abnormalities were also found in 10/30 patients (33%), including five patients with moderate cerebellar vermis hypoplasia (Fig. 7C), associated with moderate atrophy of the cerebellar hemispheres in three patients (Fig. 7D), and four patients with subnormal cerebellar vermis height.

Fetal cerebral MRI of patient P22F20, performed for a suspicion of midline abnormalities on prenatal routine ultrasonography, showed a large, triangular-shaped cavum septum pellucidum associated with a short corpus callosum (Fig. 4C). This patient died at the age of 19 months because of pulmonary arterial hypertension, before the diagnosis of *MECP2* duplication syndrome had been made. No brain MRI was done post-natally but transfontanellar ultrasonography showed corpus callosum hypo-

plasia associated with an unusual shape of the septum pellucidum (data not shown). Interestingly, a fetal cerebral MRI was also done for his sibling (P23F20), and showed the same large, triangular-shaped cavum septum pellucidum (Fig. 4C), discrete cerebellar vermis hypoplasia, and a moderately enlarged retrocerebellar cerebrospinal fluid space (Fig. 7B). Brain MRI done at the age of 10 years showed vermis hypoplasia (height = 33 mm), moderate cerebellar hemisphere atrophy, a moderately increased fluid-filled retrocerebellar space, a hypoplastic and short corpus callosum, physiological delay of posterior periventricular WM myelination, and persistence of the cavum septum pellucidum (Fig. 7C) (Table II).

### Epileptic and Drug Resistance Status

Different types of epilepsy were present in 15/30 (50%) patients. These included generalized tonic-clonic, absence, atonic and myoclonic seizures. Some patients had combinations of partial and generalized seizures. The mean age at onset was 4.6 years (mini-

TABLE I. Brain Malformations in Patients Carrying a Complex Xq Rearrangement Involving MECP2

Patients	Sex	Age at realization of MRI	OFC (SDS)	Sub-tentorial abnormalities				Supra-tentorial abnormalities										
				Cerebellar vermis abnormalities	Cerebellar hemispheres atrophy	Enlarged retro-cerebellar CFAS	Corpus callosum abnormalities	Reduced WM volume	T2 posterior periventricular WM hyperintense signal areas	Late myelinating areas	Lateral ventricles dilatation	Persistence of the CSP	Others					
P6F5	M	4 y 9 m	1.5															
P8F7	M	0 y 6 m	-3.4															
P10F9	M	2 y	-1.7															
P12F11	F	6 y 4 m	0.2	VH														
P14F13	F	6 y 10 m	-3.5	SNHV	Moderate atrophy	X												
P16F15	F	12 y	-0.7	VH														
P21F19	M	2 y 9 m	-2.3	VH														

OFC, occipitofrontal circumference; CFAS, cerebrospinal fluid appearing space; WM, white matter; VH, vermis hypoplasia; SNHV, subnormal height of the vermis; CSP, cavum septum pellucidum.

mum: 0.4 years, maximum: 18 years). Among these 15 patients, 8 (53%) had a reduced WM volume, 4 (27%) had delayed WM myelination, and 5 (33%) had dilated lateral ventricles, versus 4/15 (27%), 5/15 (33%), and 4/15 (27%) of the non-epileptic patients, respectively. The differences were not significant (Table IV). Regarding the corpus callosum abnormalities, they were present in 10/15 (67%) of both epileptic and non-epileptic patients (Table IV, Supplemental Fig. S1). Seven of the fifteen patients (47%) presented drug-resistant epilepsy, the youngest of these being 3 years old. Among these seven patients with drug-resistant epilepsy, four (57%) had a reduced white matter volume, one (14%) had delayed WM myelination, three (43%) had corpus callosum abnormalities, and four (57%) had dilated lateral ventricles. There were no significant differences compared with non-epileptic patients (Table IV). Surprisingly, two patients with drug-resistant epilepsy had none of these abnormalities (Supplemental Fig. S1).

Concerning age at the onset of epilepsy and brain malformations, we observed that patients whose age at onset of epilepsy was less than 3 years had only one brain abnormality among reduced WM volume, delayed WM myelination, and dilated lateral ventricles, versus one to three abnormalities when the age at onset was 3 years and more (Supplemental Fig. S1).

Concerning the comparison between brain malformations and aging, we observed that cerebellar vermis abnormalities, cerebellar hemisphere atrophy, and enlarged retrocerebellar CFAS were present in, respectively, 50%, 33%, and 50% of patients aged more than 10 years at the moment the MRI was done versus 20%, 0%, and 30% of patients aged less than 3 years (Supplemental Table S1).

### Genotype–Phenotype Correlation

Results of the genotype–phenotype correlation study are summarized in Table V. We found no correlation between the presence of nodular heterotopia, ventricular dilatation, WM abnormalities, and the presence of *FLNA*, *L1CAM*, or *IKBKKG*, respectively, in the duplicated segments.

### DISCUSSION

We collected and analyzed in a standardized manner the brain MRIs of 30 patients carrying an Xq28 duplication involving *MECP2*, the size of which ranged from 228 kb to 11.7 Mb and found that 93% had MRI abnormalities. No specific brain malformation pattern was identified but we found that these patients presented numerous recurrent non-specific abnormalities, the majority of which were supra-tentorial and included corpus callosum dysmorphic features, ventricular dilatation, reduced white matter volume, and abnormal increased hyperintensities on T2-weighted images involving posterior periventricular WM.

Other authors have reported such brain abnormalities in patients carrying an Xq28 duplication involving *MECP2*. Echenne et al. [2009] studied five affected male patients and found that two of them had an abnormal hyperintense signal of the posterior part of the cerebral hemispheres on brain MRI. Meins et al. [2005] reported one male patient whose MRI showed unspecific WM changes, including small cystic structures in the

TABLE II. Brain Malformations in Patients Carrying an Intrachromosomal Xq28 Duplication Involving MECP2

Patients	Sex	Age at realization of MRI	DFC (SDS)	Sub-tentorial abnormalities				Supra-tentorial abnormalities										
				Cerebellar vermis abnormalities	Cerebellar hemispheres atrophy	Enlarged retro-cerebellar CFAS	Corpus callosum abnormalities	Reduced WM volume	T2 posterior periventricular WM hyperintense signal areas	Late myelinating areas	Lateral ventricles dilatation	Persistence of the CSP	Others					
P1F1	M	22 m	-0.6				Brevilineal shaped and hypoplasia of the anterior part	X										
P2F2	M	8 y 6 m	-0.4	SNHV	Moderate atrophy										X			
P3F2	M	6 y 9 m	-0.2			X	Discrete anterior hypoplasia									X		
P4F3	F	14 y	0.7	SNHV		X	Brevilineal shaped										X	
P5F4	M	2 y 8 m	1.5				Discrete hypoplasia											
P7F6	M	21 y 8 m	0			X	Brevilineal shaped and hypoplasia of the anterior part									X		
P9F8	M	14 y 1 m	0.8				Complete, thin										X	
P11F10	M	2 y 2 m	0.2	VH			Discretely thin and hypoplastic											
P13F12	F	3 y 3 m	2.5				Thin											
P15F14	M	5 y 0 m	2				Discrete anterior hypoplasia											
P17F16	M	3 y	1.5															
P18F16	M	7 y 4 m	0			X												
P19F17	M	4 y	-0.8	SNHV		X												
P20F18	M	0 y 5 m	3.3															
P22F20	M	31.5A	3.1															
P23F20	M	10 y 3 m	-2.7	VH	Moderate atrophy	X	Brevilineal shaped and hypoplastic											
P24F21	M	18 m	1.3			X												
P25F22	M	4 y	-0.5			X	Discretely thin, according to the reduced WM volume											
P26F23	M	3 y 5 m	0				Brevilineal shaped in antero-posterior											
P27F23	M	5 y 1 m	0.5				Brevilineal shaped											
P28F24	M	3 y	-1.8															
P29F25	M	17 y	2.5		Moderate atrophy													
P30F26	M	2 y 9 m	1.2				Complete agenesis											

y, years; m, months; DFC, occipitofrontal circumference; CFAS, cerebrosplinal fluid appearing space; WM, white matter; VH, vermis hypoplasia; SNHV, subnormal height of the vermis; CSP, cavum septum pellucidum; VH, verchow Robin.



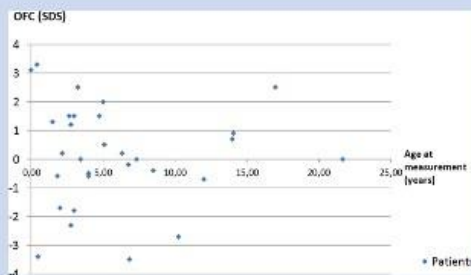


FIG. 2. Patients' OFC (occipitofrontal circumference). Note the high variability in patients' OFC.

supratentorial WM surrounding the trigonum. Friez et al. [2006] identified six families with multiple affected males with an Xq28 duplication involving *MECP2*, one of whom had agenesis of the corpus callosum, and one had gray matter heterotopia and Dandy-Walker variant. The largest series was gathered by Honda

et al. [2012], who reported brain MRI data of 12 Japanese male patients and 36 patients from the literature. They found that 54% of patients had hypoplasia of the corpus callosum, an apparent loss of cerebral volume in 74%, ventricular dilatation in 36%, an apparent loss of cerebellar volume in 27%, delayed myelination of the white cortex in 9%, and gray matter heterotopia in 9% [Honda et al., 2012].

Decobert and collaborators studied brain MRI in a total of 100 patients with unexplained non-syndromic non-progressive ID who were negative for fragile X syndrome. They showed a higher incidence of brain abnormalities, including mild corpus callosum abnormalities, signal abnormalities within the periventricular WM, lateral ventricular dilatation, and subtle cerebellar abnormalities in the ID group than in the control group (53 vs. 17, OR = 5.7) [Decobert et al., 2005]. We can thus note that the brain MRI abnormalities we observed in our series did not seem different from those observed in patients with non-syndromic non-progressive ID. Moreover, agenesis/hypoplasia of the corpus callosum, posterior fossa abnormalities, and hypomyelination can be observed in numerous malformative syndromes and thus do not provide any clues for the diagnosis of Xq28 duplication syndrome involving *MECP2* [Guibaud and des Portes, 2006; Schell-Apacik et al., 2008; Philippe et al., 2013].

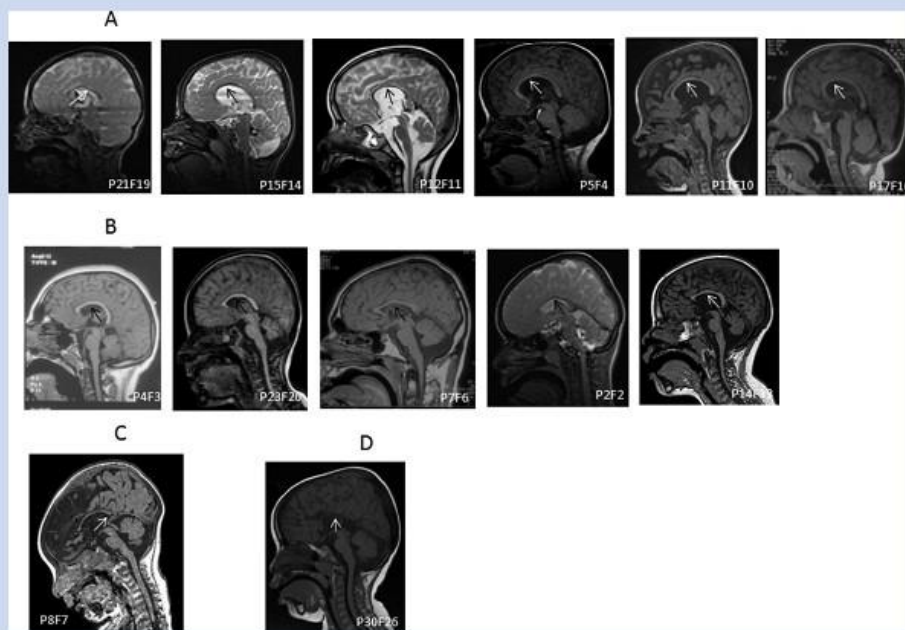


FIG. 3. Corpus callosum abnormalities. A: Hypoplasia of the corpus callosum [P21F19, P15F14, P12F11, P5F4, P11F10, P17F16] (arrow). B: Short and moderately hypoplastic corpus callosum [P4F3, P23F20, P7F6, P2F2, P14F13] (arrow). C: Hypoplastic, defective modeling of the genu and the splenium with non-identification of the beak, suggestive of partial agenesis [P8F7] (arrow). D: Complete agenesis of corpus callosum [P30F26] (arrow).

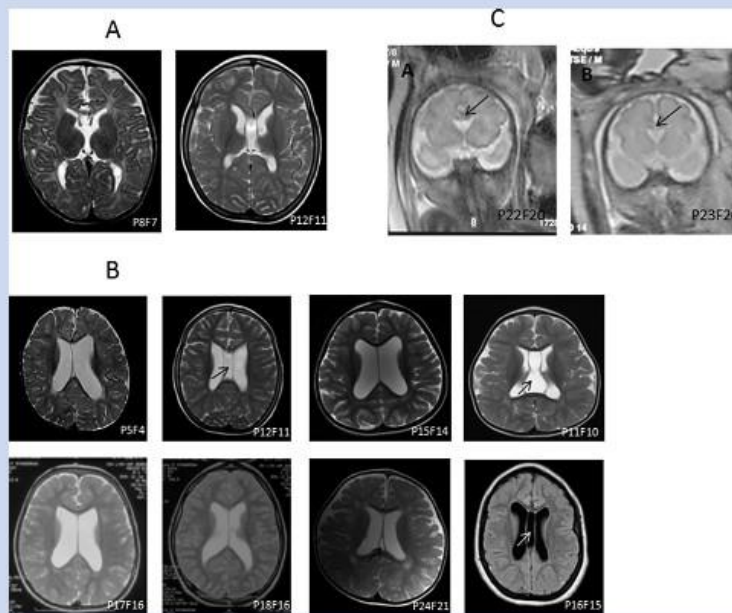


FIG. 4. A: Reduced white matter [P8F7, P12F11]. B: Lateral ventricle dilatation [P5F4, P11F10, P12F11, P15F14, P16F15, P17F16, P18F16, P24F21]. Note that P11F10, P12F11, P24F21, and P16F15 have persistence of the cavum septum pellucidum with or without cavum vergae [arrow]. C: Fetal cerebral MRI. Large, triangular-shaped cavum septum pellucidum on fetal cerebral MRI in two siblings: P22F20 [Left] and P23F20 [Right] [arrow].

In our study, among the 30 patients, five (17%) had macrocephaly and four microcephaly. These data are interesting given that such high variability in OFC is not frequent in syndromology. Among the five patients with macrocephaly, only one (P15F14) had lateral ventricle dilatation (Fig. 2, Tables I and II).

Concerning the epileptic aspect, in this study, we found that 50% (15/30) of the patients were epileptic, including 7/15 (47%) patients with drug-resistant epilepsy requiring multi-drug therapy. This result correlates with previously reported data, since Van Esch described in a review of reported *MECP2* duplication patients in 2011, that 54% were epileptic [Van Esch, 2011]. Vignoli and co-workers recently described a series of eight patients carrying an Xq28 duplication involving *MECP2*. The series included six epileptic patients, four of whom (67%) had drug-resistant epilepsy, with an age at onset ranging from 9 to 24 years. The authors reported that epilepsy occurred in 90% of patients with a long-term follow-up [Vignoli et al., 2012]. Since some of the brain malformations we observed were not congenital and given that non-controlled epilepsy can affect brain structures [Morgan et al., 2015], we raised the question of a possible link between the epileptic status, in particular drug-resistance, the age at onset of

epilepsy and malformations in patients with an Xq28 duplication involving *MECP2*. We found no significant differences between the epileptic and the non-epileptic patients concerning the reduced WM volume (53% vs. 27%), the delayed myelination of the cerebral WM (respectively, 27% vs. 33%), the dilatation of lateral ventricles (respectively, 33% vs. 27%), and the corpus callosum abnormalities (67% in both groups) (Table IV). We cannot therefore affirm that there is a link between the presence of epilepsy and some of the observed supra-tentorial brain abnormalities in patients with an Xq28 duplication involving *MECP2*.

Regarding the patients with drug-resistant epilepsy, 57% (4/7) of them had reduced WM volume and 57% had ventricular dilatation versus 27% (4/15) and 27%, respectively, of the non-epileptic patients (Table IV). Moreover, two patients with drug-resistant epilepsy had none of these abnormalities. We also observed that patients whose age at onset of epilepsy was less than 3 years had only one of these brain abnormalities versus one to three abnormalities when the age at onset was 3 years and more (Supplemental Fig. S1). These results do not allow us to conclude that severe uncontrolled seizures had an effect on brain structures in these patients.

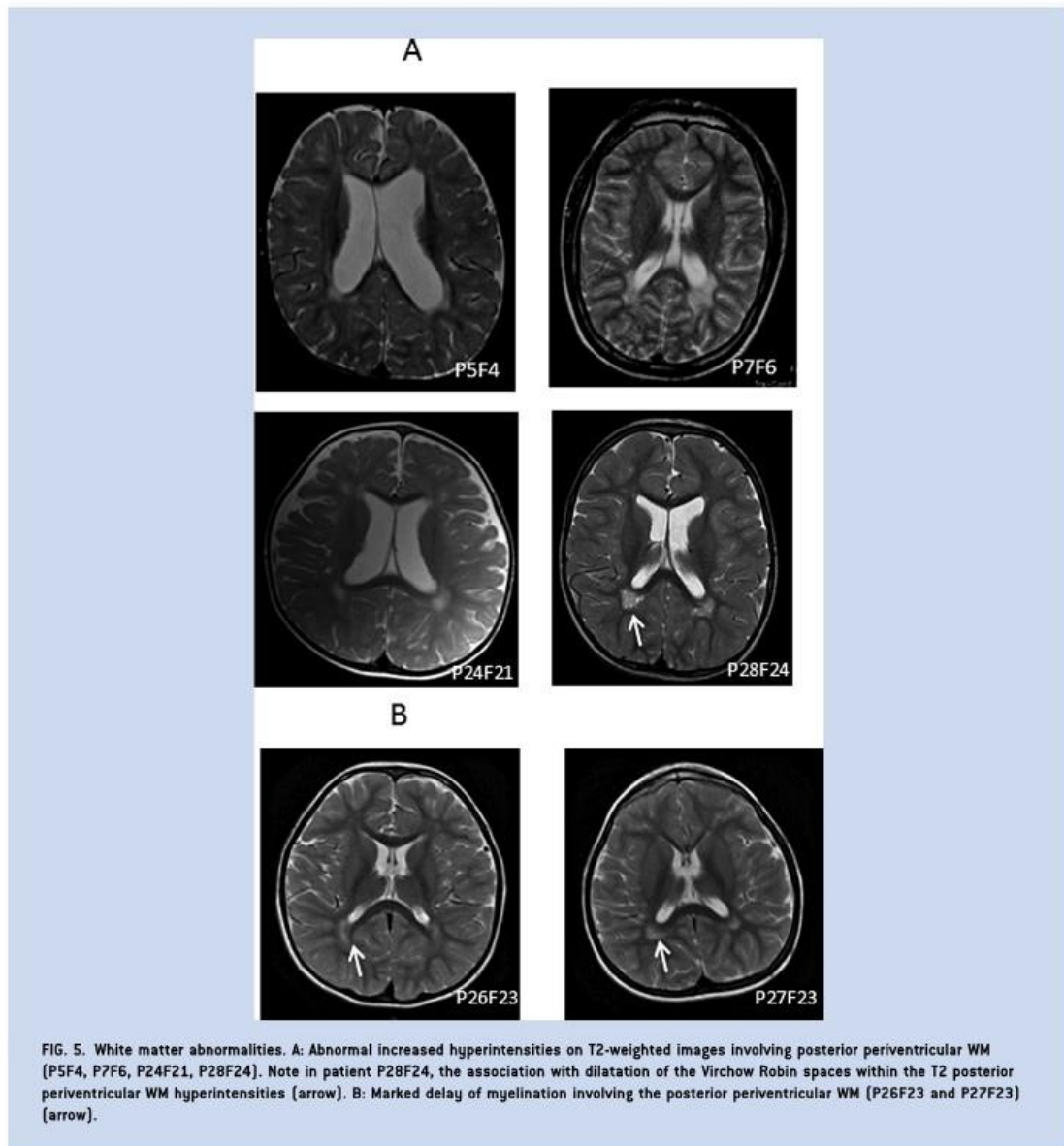


FIG. 5. White matter abnormalities. A: Abnormal increased hyperintensities on T2-weighted images involving posterior periventricular WM (P5F4, P7F6, P24F21, P28F24). Note in patient P28F24, the association with dilatation of the Virchow Robin spaces within the T2 posterior periventricular WM hyperintensities (arrow). B: Marked delay of myelination involving the posterior periventricular WM (P26F23 and P27F23) (arrow).

We observed that four patients harboured a de novo unbalanced translocation that involved the Xq28 region resulting in microdeletions next to the Xq28 duplication, including one 225 kb 9p24.3 deletion in P12F11 and one 1.13 Mb 5q35 deletion in P14F13 (Table III). We could hypothesize about a possible role of these deletions in the patients' brain phenotype. The

9p24.3 deletion encompassed the *DOCK8* gene, which has been reported to be involved in an autosomal dominant form of ID only once [Griggs et al., 2008]. The 5q35 deletion involved several genes, including *MAPK9* (OMIM 602896), which has been shown to be involved in brain formation in mice [Kuan et al., 1999]. We found no other patients carrying a de novo

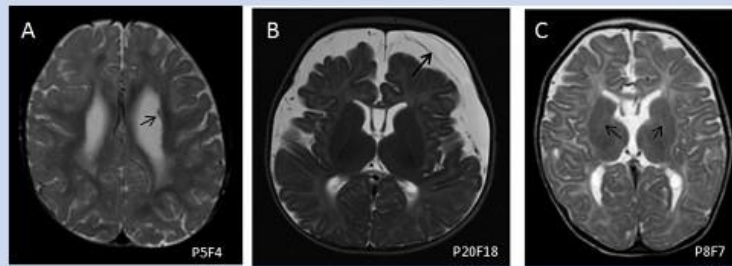


FIG. 6. Rare malformations (arrows). A: Bilateral asymmetric subependymal nodular heterotopia (P5F4). B: Persistent subdural hygroma (P20F18). C: Lack of internal capsule myelination (P8F7).

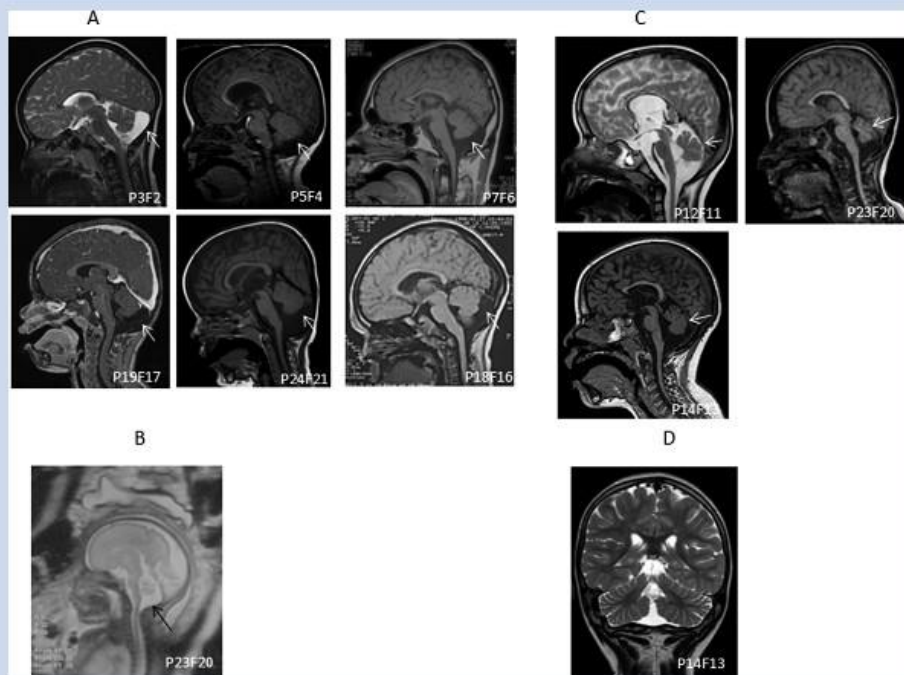


FIG. 7. Cerebellar malformations (arrows). A: Increased fluid-filled retrocerebellar space (P3F2, P5F4, P7F6, P19E17, P24F21, P18F16). B: Fetal cerebral MRI. Increased fluid-filled retrocerebellar space associated with vermian hypoplasia (P23F20). C: Moderate vermian hypoplasia [reduction of vermian height] (P12F11, P23F20, P14F13). Note the association with focal atrophy of the cranial part of the vermis in P12F11 and P14F13. D: Moderate atrophy of the cerebellar hemispheres (P14F13).

TABLE III. Cytogenetic Molecular Characteristics of Xq28 Duplications

Patients	Sex	Inheritance	Anomaly	CCH	Heterotopia	VD	Macrocephaly	DM	WHS	Proximal breakpoint	Distal breakpoint	Size (kb)	Genes
P1F1	M	I	dupXq						X	152485412	15309163	1124	SLC6A8, IDH3G, PZD4, LICAM, IRAK1, MECP2, OPN1LW, OPN1MW, FLNA
P2F2	M	I	dupXq	X						152720261	153632865	1100	SLC6A8, IDH3G, PZD4, LICAM, IRAK1, MECP2, OPN1LW, OPN1MW, FLNA, RPL10, GDI1, IKBK6
P3F2	M	I	dupXq	X						152720261	153632865	1100	SLC6A8, IDH3G, PZD4, LICAM, IRAK1, MECP2, OPN1LW, OPN1MW, FLNA, RPL10, GDI1, IKBK6
P4F3	F	DN	dupXq							153093533	153747928	654	PZD4, LICAM, IRAK1, MECP2, OPN1LW, OPN1MW, FLNA, RPL10, GDI1
P5F4	M	I	dupXq	X	X				X	152560105	153609163	1049	SLC6A8, IDH3G, PZD4, LICAM, IRAK1, MECP2, OPN1LW, OPN1MW, FLNA
P6F5	M	DN	dupXq + 2 dupXq							153272039	153706598	430	IRAK1, MECP2, OPN1LW, OPN1MW, FLNA, RPL10, GDI1
P7F6	M	I	dupXq	X				X		151261717	153675357	2400	SLC6A8, IDH3G, PZD4, LICAM, IRAK1, MECP2, OPN1LW, OPN1MW, FLNA, RPL10, GDI1
P8F7	M	DN	dupXq + delYp	X						144262624	154929420	10667	SLC6A8, IDH3G, PZD4, LICAM, IRAK1, MECP2, OPN1LW, OPN1MW, FLNA, RPL10, GDI1, IKBK6
P9F8	M	I	dupXq							153140483	153609163	467	LICAM, IRAK1, MECP2, OPN1LW, OPN1MW, FLNA
P10F9	M	DN	t(X;Y)*	X						148824301	154008471	8000	SLC6A8, IDH3G, PZD4, LICAM, IRAK1, MECP2, OPN1LW, OPN1MW, FLNA, RPL10, GDI1, IKBK6
P11F10	M	I	dupX			X				153058879	153558612	500	IDH3G, PZD4, LICAM, IRAK1, MECP2, OPN1LW, OPN1MW
P12F11	F	DN	t(X;9) <sup>a</sup>	X	X					143485415	155223860	11738	SLC6A8, IDH3G, PZD4, LICAM, IRAK1, MECP2, OPN1LW, OPN1MW, FLNA, RPL10, GDI1, IKBK6
P13F12	F	NA	dupXq				X			15321264	153613200	382	IRAK1, MECP2, OPN1LW, OPN1MW, FLNA
P14F13	F	DN	t(X;5) <sup>c</sup>	X			X			152561365	154929412	2368	SLC6A8, IDH3G, PZD4, LICAM, IRAK1, MECP2, OPN1LW, OPN1MW, FLNA, RPL10, GDI1, IKBK6
P15F14	M	I	dupXq	X	X		X			152867160	153609313	740	SLC6A8, IDH3G, PZD4, LICAM, IRAK1, MECP2, OPN1LW, OPN1MW, FLNA
P16F15	F	DN	t(X;17) <sup>d</sup>	X	X					153049024	154776918	1728	IDH3G, PZD4, LICAM, IRAK1, MECP2, OPN1LW, OPN1MW, FLNA, RPL10, GDI1, IKBK6
P17F16	M	I	dupXq			X		X		153300357	153949811	649	MECP2, OPN1LW, OPN1MW, FLNA, RPL10, GDI1, IKBK6
P18F16	M	I	dupXq			X				153300357	153949811	649	MECP2, OPN1LW, OPN1MW, FLNA, RPL10, GDI1, IKBK6
P19F17	M	I	dupXq							15327239	153609113	332	IRAK1, MECP2, OPN1LW, OPN1MW, FLNA
P20F18	M	I	dupXq	X			X			153558471	153558471	497	PZD4, LICAM, IRAK1, MECP2, OPN1LW, OPN1MW
P21F19	M	DN	t(X;8) <sup>e</sup>	X						146079000	154573000	8600	SLC6A8, IDH3G, PZD4, LICAM, IRAK1, MECP2, OPN1LW, OPN1MW, FLNA, RPL10, GDI1, IKBK6
P22F20	M	I	dupXq				X			145125374	154397191	9200	SLC6A8, IDH3G, PZD4, LICAM, IRAK1, MECP2, OPN1LW, OPN1MW, FLNA, RPL10, GDI1, IKBK6
P23F20	M	I	dupXq	X				X		145125374	154397191	9200	SLC6A8, IDH3G, PZD4, LICAM, IRAK1, MECP2, OPN1LW, OPN1MW, FLNA, RPL10, GDI1, IKBK6
P24F21	M	NA	dupXq			X			X	153190661	153609163	419	IRAK1, MECP2, OPN1LW, OPN1MW, FLNA

(Continued)

TABLE III. (Continued)

Patient	Sex	Inheritance	Anomaly	CCH	Heretopia	VD	Macrocephaly	DM	WMH	Proximal breakpoint	Distal breakpoint	Size (kb)	Genes
P25F22	M	I	2 dupXq					X		152987755	153541430	550	IDH3C, POZD4, L1CAM, IRAK1, MECP2, OPN1LW, OPN1MW
P28F23	M	I	dupXq		X			X		152996339	153609163	613	IDH3C, POZD4, L1CAM, IRAK1, MECP2, OPN1LW, OPN1MW, FLNA
P27F23	M	I	dupXq					X		152996339	153609163	613	IDH3C, POZD4, L1CAM, IRAK1, MECP2, OPN1LW, OPN1MW, FLNA
P28F24	M	I	dupXq					X	X	152171761	153636703	465	IRAK1, MECP2, OPN1LW, OPN1MW, FLNA, RPL10
P28F25	M	I	dupXq				X			153277239	153505485	228	IRAK1, MECP2, OPN1LW, OPN1MW
P30F26	M	I	dupXq			X		X		152857869	153609163	751	SLUG48, IDH36, POZD4, L1CAM, IRAK1, MECP2, OPN1LW, OPN1MW, FLNA

I, inherited; DM, de novo; M, male; F, female; CCH, corpus callosum hypoplasia; VD, ventricular dilatation; DM, delayed myelination; WMH, white matter hyperintensities.  
 \*46,XY,t(17)(X,Y)(q28,q11.23)[MECP2+RP-262K13+].  
 \*46,XX,t(17)(X,Y)(p24.3,q27.31)(225Kb).  
 \*46,XX,t(17)(X,Y)(q28,q35.31)(1.13 Mb).  
 \*46,XX,t(17)(X,Y)(q28,q35.31)(0.1752199,40X561+).  
 \*46,XY,t(18)(1)(q27.3q28qtel).

deletion with similar breakpoints in the DECIPHER database [https://decipher.sanger.ac.uk/]. Furthermore, patients P12F11 and P14F13 had a 11738 and 2368 kb Xq28 duplication, respectively, and their phenotype was very similar to that of the reported affected males and included severe ID, epilepsy, spasticity, constipation, stereotyped hand movements, and recurrent infections. In light of these data, we believe that these deletions do not play a major role in the brain phenotype of our patients.

We wondered whether there was an aging effect on the observed brain malformations, since Reardon and co-workers studied seven affected males from three families and found that patients harbored progressive cerebellar atrophy that was not present in early infancy [Reardon et al., 2010]. In our study, only 10/30 (33%) patients had cerebellar vermis abnormalities but the mean age at MRI examination was 5 years and 10 months and we found that cerebellar abnormalities seemed to be more frequent in patients aged more than 10 years than in patients less than 3 years (Supplemental Table S1). We can thus hypothesize that if we repeat brain MRI in a few years in the younger patients, other cerebellar abnormalities would be noticeable.

When the content of genes involved in the duplicated segment was studied, we first focused on mutations in the *L1CAM* gene, involved in overlapping conditions. These included X-linked hydrocephalus due to stenosis of the aqueduct of Sylvius (HSAS [MIM 307000]), ID, aphasia, shuffling gait, and adducted thumbs (MASA [MIM 303350]), and corpus callosum hypoplasia, retardation, adducted thumbs, spastic paraplegia, and hydrocephalus (CRASH [MIM 312900]) syndromes [Weller and Gärtner, 2001]. As mentioned by Van Esch and co-workers, no brain abnormalities that fitted with mutated *L1CAM* phenotypes were observed in the patients who carried a duplication that included *L1CAM* [Van Esch et al., 2005]. Velinov et al. [2009] and Kirk et al. [2009] described, respectively, one and two patients with an *MECP2* duplication excluding *L1CAM* who had no ventricular dilatation [Kirk et al., 2009; Velinov et al., 2009]. In the present series, among nine patients with ventricular dilatation, six (67%) had a duplication including *L1CAM* (Table V). Many children carrying an Xq28 duplication involving *MECP2* exhibit delayed myelination of the cerebral WM at the posterior part of the cerebral hemispheres, but these abnormalities are not constant features [Meins et al., 2005; Clayton-Smith et al., 2009; Kirk et al., 2009; Lugtenberg et al., 2009; Prescott et al., 2009; Velinov et al., 2009]. Philippe and co-workers recently studied skin fibroblasts derived from patients with WM abnormalities and an Xq28 duplication involving *MECP2* and the *IKBK* gene. They showed that *IKBK* overexpression caused impaired NF- $\kappa$ B signalling. They concluded that these data supported the role of NF- $\kappa$ B signalling in astroglial cells for normal myelin formation in the central nervous system [Philippe et al., 2013]. In our study, we found that 6/30 patients (20%) had T2 posterior periventricular WM hyperintensities and 9/30 (30%) showed a significant delay in myelination of the WM. Among these 15 patients with WM abnormalities, 3 had a duplication that included the *IKBK* gene (Table III, Table V). Interestingly, in the present series, one patient (P5F4) harbored bilateral posterior asymmetric subependymal

TABLE IV. Epileptic Status and Supra-Tentorial Brain Malformations

	Epileptic patients (%)			Non-epileptic patients (%)	Total
	Drug-susceptible	Drug-resistant	Total		
Reduced WM volume	4/8 (50%)	4/7 (57%)	8/15 (53%)	4/15 (27%)	12/30
Delayed WM myelination	3/8 (38%)	1/7 (14%)	4/15 (27%)	5/15 (33%)	9/30
Dilatation of lateral ventricles	1/8 (13%)	4/7 (57%)	5/15 (33%)	4/15 (27%)	9/30
Corpus callosum abnormalities	7/8 (88%)	3/7 (43%)	10/15 (67%)	10/15 (67%)	20/30
Total	8	7	15	15	30

TABLE V. Genotype–Phenotype Correlation: Genes Involved in Xq28 Duplications and Brain Malformations

	Gene included (+) or not (–) in the duplication (number of patients)								Total
	FLNA		L1CAM		IKBK		Xq28 duplications including LCRL1		
	+	–	+	–	+	–	Yes	No	
Peri-ventricular heterotopia	1	0							1
Ventricular dilatation			6	3					9
Macrocephaly			3	2					5
White matter hypersignals					0	6			6
Delayed myelination					3	6			9
Corpus callosum hypoplasia							8	4	12

nodular heterotopia and his duplication included the *FLNA* gene (Fig. 6A, Tables III and V) (MIM 300017), which is implicated in several phenotypes including X-linked periventricular heterotopia [Parrini et al., 2006].

Finally, Honda and co-workers suggested that copy-number gains including *MECP2* located between two low copy repeats (*LCRK2* and *LCRL1*) were associated with a higher incidence of hypoplasia of the corpus callosum. Our series showed similar results since we found that 67% (8/12) of patients with distal breakpoints near/on/over *LCRL1* and 22% (4/18) of patients with distal breakpoints proximal to *LCRL1* had hypoplasia of the corpus callosum [Honda et al., 2012] (Fig. 1, Table V).

In conclusion, this study showed that patients with Xq28 duplication involving *MECP2* shared several brain abnormalities at magnetic resonance imaging that were non-specific. However, some abnormalities, such as corpus callosum abnormalities, ventricular dilatation, reduced WM volume, vermis hypoplasia, and increased T2 hyperintensities in posterior periventricular WM, appeared to be rather recurrent among patients. These results did not correlate with the genes involved in the duplications. No link was found between the presence of epilepsy, the age at onset, the drug-resistant status, and the observed brain malformations. The absence of a correlation could, in part, be due to the small number of patients. These data showed that brain MRI could not be of help in the diagnostic work up of Xq28 duplication involving *MECP2*.

## ACKNOWLEDGMENTS

We are deeply grateful to the patients and their parents for their participation in this study. We thank Philip Bastable from the “Pôle de Recherche” of Dijon University Hospital for a helpful review of this article. We thank the Regional Council of Burgundy and the association Xtraordinaire for their financial support of the project, and the PHRC RMLX (Pr V. Des Portes) for its support.

## REFERENCES

- Bauters M, Van Esch H, Friez MJ, Boespflug-Tanguy O, Zenker M, Vianna-Morgante AM, Rosenberg C, Ignatius J, Raynaud M, Hollanders K, Govaerts K, Vandenreijt K, Niel F, Blanc P, Stevenson RE, Fryns JP, Marynen P, Schwartz CE, Froyen G. 2008. Nonrecurrent *MECP2* duplications mediated by genomic architecture-driven DNA breaks and break-induced replication repair. *Genome Res* 18:847–858.
- Bijlsma EK, Collins A, Papa FT, Tejada MI, Wheeler P, Peeters EA, Gijbsbers AC, van de Kamp JM, Kriek M, Losekoot M, Broekma AJ, Crolla JA, Pollazzon M, Mucciolo M, Katzaki E, Disciglio V, Ferreri MI, Marozza A, Mencarelli MA, Castagnini C, Dosa L, Ariani F, Mari F, Canitano R, Hayek G, Botella MP, Gener B, Minguez M, Renieri A, Ruivenkamp CAL. 2012. Xq28 duplications including *MECP2* in five females: Expanding the phenotype to severe mental retardation. *Eur J Med Genet* 55: 404–413.
- Clayton-Smith J, Walters S, Hobson E, Burkitt-Wright E, Smith R, Toutain A, Amiel J, Lyonnet S, Mansour S, Fitzpatrick D, Ciccone R, Ricca I,

- Zuffardi O, Donnai D. 2009. Xq28 duplication presenting with intestinal and bladder dysfunction and a distinctive facial appearance. *Eur J Hum Genet* 17:434–443.
- Decobert F, Grabar S, Merzoug V, Kalifa G, Ponsot G, Adamsbaum C, des Portes V. 2005. Unexplained mental retardation: Is brain MRI useful? *Pediatr Radiol* 35:587–596.
- Echenne B, Roubertie A, Lugtenberg D, Kleefstra T, Hamel BC, Van Bokhoven H, Lacombe D, Philippe C, Jonveaux P, de Brouwer APM. 2009. Neurologic aspects of *MECP2* gene duplication in male patients. *Pediatr Neurol* 41:187–191.
- Friez MJ, Jones JR, Clarkson K, Lubs H, Abuelo D, Blaymore Bier JA, Pai S, Simensen R, Williams C, Giampietro PF, Schwartz CE, Stevenson RE. 2006. Recurrent infections, hypotonia, and mental retardation caused by duplication of *MECP2* and adjacent region in Xq28. *Pediatrics* 118:e1687–e1695.
- Griggs BL, Ladd S, Saul RA, DuPont BR, Srivastava AK. 2008. Dedicator of cytokinesis 8 is disrupted in two patients with mental retardation and developmental disabilities. *Genomics* 91:195–202.
- Guilbaud L, des Portes V. 2006. Plea for an anatomical approach to abnormalities of the posterior fossa in prenatal diagnosis. *Ultrasound Obstet Gynecol* 27:477–481.
- Honda S, Hayashi S, Nakane T, Imoto I, Kurosawa K, Mizuno S, Okamoto N, Kato M, Yoshihashi H, Kubota T, Nakagawa E, Goto Y, Inazawa J. 2012. The incidence of hypoplasia of the corpus callosum in patients with dup (X)(q28) involving *MECP2* is associated with the location of distal breakpoints. *Am J Med Genet A* 158A:1292–1303.
- Kirk EP, Malaty-Brevaud V, Martini N, Lacoste C, Levy N, Maclean K, Davies L, Philip N, Badens C. 2009. The clinical variability of the *MECP2* duplication syndrome: Description of two families with duplications excluding *LICAM* and *FLNA*. *Clin Genet* 75:301–303.
- Kuan C-Y, Yang DD, Roy DRS, Davis RJ, Rakic P, Flavell RA. 1999. The Jnk1 and Jnk2 protein kinases are required for regional specific apoptosis during early brain development. *Neuron* 22:667–676.
- Lugtenberg D, Kleefstra T, Oudakker AR, Nillesen WM, Yntema HG, Tzschach A, Raynaud M, Rating D, Journal H, Chelly J, Goizet C, Lacombe D, Pedespan JM, Echenne B, Tariverdian G, O'Rourke D, King MD, Green A, van Kogelenberg M, Van Esch H, Geçz J, Hamel BCJ, van Bokhoven H, de Brouwer APM. 2009. Structural variation in Xq28: *MECP2* duplications in 1% of patients with unexplained XLMR and in 2% of male patients with severe encephalopathy. *Eur J Hum Genet* 17:444–453.
- Meins M, Lehmann J, Gerresheim F, Herchenbach J, Hagedorn M, Hameister K, Epplen JT. 2005. Submicroscopic duplications in Xq28 causes increased expression of the *MECP2* gene in a boy with severe mental retardation and features of Rett syndrome. *J Med Genet* 42:e12.
- Morgan VI, Conrad BN, Abou-Khalil B, Kang H. 2015. Increasing structural atrophy and functional isolation of the temporal lobe with duration of disease in temporal lobe epilepsy. *Epilepsy Res* 110:171–178.
- Parrini E, Ramazzotti A, Dobyns WB, Mei D, Moro F, Veggiotti P, Marini C, Brilstra EH, DallaBernardina B, Goodwin L, Bodell A, Jones MC, Nangeroni M, Palmeri S, Said E, Sander JW, Striano P, Takahashi Y, Van Maldergem L, Leonardi G, Wright M, Walsh CA, Guerrini R. 2006. Periventricular heterotopia: Phenotypic heterogeneity and correlation with Filamin A mutations. *Brain* 129:1892–1906.
- Philippe O, Rio M, Malan V, Van Esch H, Baujat G, Bahi-Buisson N, Valayannopoulos V, Gesny R, Bonnefont JP, Munnich A, Froyen G, Amiel J, Boddaert N, Colleaux L. 2013. NF- $\kappa$ B signalling requirement for brain myelin formation is shown by genotype/MRI phenotype correlations in patients with Xq28 duplications. *Eur J Hum Genet* 21:195–199.
- Prescott TE, Rodningen OK, Bjornstad A, Stray-Pedersen A. 2009. Two brothers with a microduplication including the *MECP2* gene: Rapid head growth in infancy and resolution of susceptibility to infection. *Clin Dysmorphol* 18:78–82.
- Reardon W, Donoghue V, Murphy AM, King MD, Mayne PD, Horn N, Birkmøller L. 2010. Progressive cerebellar degenerative changes in the severe mental retardation syndrome caused by duplication of *MECP2* and adjacent loci on Xq28. *Eur J Pediatr* 169:941–949.
- Schell-Apacik CC, Wagner K, Bihler M, Ertl-Wagner B, Heinrich U, Klopocki E, Kalscheuer VM, Muenke M, von Voss H. 2008. Agenesis and dysgenesis of the corpus callosum: Clinical, genetic and neuroimaging findings in a series of 41 patients. *Am J Med Genet Part A* 146A:2501–2511.
- Shimada S, Okamoto N, Ito M, Arai Y, Momosaki K, Togawa M, Maegaki Y, Sugawara M, Shimojima K, Osawa M, Yamamoto T. 2013. *MECP2* duplication syndrome in both genders. *Brain Dev* 35:411–419.
- Van Esch H, Bauters M, Ignatius J, Jansen M, Raynaud M, Hollanders K, Lugtenberg D, Bienvenu T, Jensen LR, Gécz J, Moraine C, Marynen P, Fryns JP, Froyen G. 2005. Duplication of the *MECP2* region is a frequent cause of severe mental retardation and progressive neurological symptoms in males. *Am J Hum Genet* 77:442–453.
- Van Esch H. 2011. *MECP2* Duplication Syndrome. *Mol Syndromol* 2:128–136.
- Velinov M, Novelli A, Gu H, Fenko M, Dolzhanskaya N, Bernardini L, Capalbo A, Dallapiccola B, Jenkins EC, Brown WT. 2009. De-novo 2.15Mb terminal Xq duplication involving *MECP2* but not *LICAM* gene in a male patient with mental retardation. *Clin Dysmorphol* 18:9–12.
- Vignoli A, Borgatti R, Peron A, Zucca C, Ballarati L, Bonaglia C, Bellini M, Giordano L, Romaniello R, Bedeschi MF, Epifanio R, Russo S, Caselli R, Giardino D, Darra F, La Briola F, Banderali G, Canevini MP. 2012. Electroclinical pattern in *MECP2* duplication syndrome: Eight new reported cases and review of literature. *Epilepsia* 53:1146–1155.
- Weller S, Gärtner J. 2001. Genetic and clinical aspects of X-linked hydrocephalus (L1 disease): Mutations in the *L1CAM* gene. *Hum Mutat* 18:1–12.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.



### 3. La cohorte française de 59 patients atteints de MECP2DS : description phénotypique fine avec un focus sur les caractéristiques morphologiques et neurologiques

La duplication Xq28 impliquant le gène *MECP2* est une pathologie rare principalement décrite chez des individus de sexe masculin présentant un retard de développement et une DI sévères, une spasticité des membres inférieurs, une épilepsie, des mouvements stéréotypés, une constipation chronique et des infections récurrentes, notamment respiratoires. Peu de grandes séries de patients ayant été publiées, j'ai souhaité par ce travail, décrire le plus finement possible le phénotype de cette affection, en mettant l'accent sur les caractéristiques morphologiques et neurologiques.

Par le biais d'une étude collaborative nationale (PHRC RMLX, Pr V Des Portes), 101 cas, dont 91 hommes et 10 femmes (y compris des jumeaux), porteurs d'une duplication du gène *MECP2*, identifiée dans le cadre du bilan étiologique de leur DI, ont été recensés en 2018 dans l'ensemble des 17 laboratoires français de cytogénétique et biologie moléculaire contactés. Parmi les 91 hommes, les duplications résultaient d'une duplication interstitielle de la région Xq28 dans 86 cas (54 familles) et d'une aberration chromosomique complexe (telle qu'une translocation X-autosome ou une insertion de la région Xq28 dans un autosome) dans cinq cas. Soixante-treize des 83 patients (88%) pour lesquels nous disposons de l'information, avaient hérité leur duplication de leur mère.

Vingt-trois des 86 patients (27%) sont décédés avant l'âge de 25 ans (âge moyen : 9,5 ans), ce qui rejoint les chiffres de la littérature. Parmi les 11 patients pour lesquels la cause du décès était connue, 8 (73%) sont décédés d'une infection respiratoire, 2 (18%) d'une hypertension pulmonaire et 1 (9%) d'un état de mal épileptique.

J'ai pu recueillir, auprès de 24 services de génétiques français où étaient suivis ces patients, des données cliniques pour 59 des 86 patients porteurs d'une duplication interstitielle et j'ai pu examiner 25 d'entre eux (dans le service hospitalier où ils étaient suivis ou à leur domicile), à travers 14 villes de France, ce qui m'a permis d'identifier des signes cliniques récurrents parmi les patients, certains d'entre eux n'ayant encore jamais été décrits.

## Résumé

Avec l'aide précieuse du Docteur Marguerite Miguet, que j'ai encadrée pour son travail de thèse de médecine, nous avons ainsi rapporté les données clinico-moléculaires de ces 59 patients que l'on peut résumer ainsi : la grande majorité des patients partageaient des particularités morphologiques faciales similaires, qui évoluaient avec l'âge (hypoplasie malaire, arête nasale étroite et proéminente, lèvre inférieure épaisse, oreilles proéminentes), des cheveux épais, un livedo des membres et du tronc, des doigts effilés, des petits pieds et troubles vasomoteurs. L'hypotonie précoce, le retard de développement global et la DI sévère étaient constants, 21% des patients n'avaient pas acquis la marche. Chez les patients capables de tenir debout, la faiblesse musculaire et la spasticité des membres inférieurs étaient responsables d'un habitus particulier avec flexion des genoux et des hanches, un élargissement du polygone de sustentation et une démarche pseudo-ataxique. D'autres signes étaient fréquemment observés tels qu'une scoliose (53%), un strabisme divergent (76%), une hypermétropie (54%), une diminution de la sensibilité à la douleur (78 %) et des mouvements stéréotypés (89%), sans retrait social évident. La plupart des patients n'ont pas développé de langage expressif, 35% ne prononçant que quelques mots. L'épilepsie était fréquente (59%), avec un âge moyen de début de 7,4 ans, et était souvent pharmacorésistante (62%). D'autres problèmes médicaux étaient fréquents comme la constipation chronique (78%) et les infections récurrentes (89%). L'hypertension artérielle pulmonaire est apparue comme une cause de décès précoce chez ces patients, soulignant l'intérêt d'un dépistage précoce dans la vie.

## REVIEW

## Further delineation of the *MECP2* duplication syndrome phenotype in 59 French male patients, with a particular focus on morphological and neurological features

Marguerite Miguet,<sup>1</sup> Laurence Faivre,<sup>2</sup> Jeanne Amiel,<sup>3</sup> Mathilde Nizon,<sup>3</sup> Renaud Touraine,<sup>4</sup> Fabienne Prieur,<sup>4</sup> Laurent Pasquier,<sup>5</sup> Mathilde Lefebvre,<sup>2</sup> Julien Thevenon,<sup>2</sup> Christèle Dubourg,<sup>6</sup> Sophie Julia,<sup>7</sup> Catherine Sarret,<sup>8</sup> Ganaëlle Remerand,<sup>8</sup> Christine Francannet,<sup>9</sup> Fanny Laffargue,<sup>9</sup> Odile Boespflug-Tanguy,<sup>10</sup> Albert David,<sup>11</sup> Bertrand Isidor,<sup>11</sup> Jacqueline Vigneron,<sup>12</sup> Bruno Leheup,<sup>12</sup> Laetitia Lambert,<sup>12</sup> Christophe Philippe,<sup>13</sup> Mylène Béri-Dexheimer,<sup>13</sup> Jean-Marie Cuisset,<sup>14</sup> Joris Andrieux,<sup>15</sup> Ghislaine Plessis,<sup>16</sup> Annick Toutain,<sup>17</sup> Laurent Guibaud,<sup>18</sup> Valérie Cormier-Daire,<sup>3</sup> Marlene Rio,<sup>3</sup> Jean-Paul Bonnefont,<sup>19</sup> Bernard Echenne,<sup>20</sup> Hubert Journal,<sup>21</sup> Lydie Burglen,<sup>22</sup> Sandrine Chantot-Bastaraud,<sup>22</sup> Thierry Bienvenu,<sup>23</sup> Clarisse Baumann,<sup>24</sup> Laurence Perrin,<sup>24</sup> Séverine Drunat,<sup>25</sup> Pierre-Simon Jouk,<sup>26</sup> Klaus Dieterich,<sup>26</sup> Françoise Devillard,<sup>26</sup> Didier Lacombe,<sup>27</sup> Nicole Philip,<sup>28</sup> Sabine Sigaudy,<sup>28</sup> Anne Moncla,<sup>29</sup> Chantal Missirian,<sup>29</sup> Catherine Badens,<sup>30</sup> Nathalie Perretton,<sup>31</sup> Christel Thauvin-Robinet,<sup>2</sup> Réseau AChro-Puce,<sup>32</sup> Jean-Michel Pedespan,<sup>33</sup> Caroline Rooryck,<sup>27</sup> Cyril Goizet,<sup>27</sup> Catherine Vincent-Delorme,<sup>34</sup> Bénédicte Duban-Bedu,<sup>35</sup> Nadia Bahi-Buisson,<sup>36</sup> Alexandra Afenjar,<sup>37</sup> Kim Maincent,<sup>37</sup> Delphine Héron,<sup>38</sup> Jean-Luc Alessandri,<sup>39</sup> Dominique Martin-Coignard,<sup>40</sup> Gaëtan Lesca,<sup>41,42</sup> Massimiliano Rossi,<sup>41,42</sup> Martine Raynaud,<sup>43</sup> Patrick Callier,<sup>44</sup> Anne-Laure Mosca-Boidron,<sup>44</sup> Nathalie Marle,<sup>44</sup> Charles Coutton,<sup>45</sup> Véronique Satre,<sup>45</sup> Cédric Le Caignec,<sup>46,47</sup> Valérie Malan,<sup>48</sup> Serge Romana,<sup>48</sup> Boris Keren,<sup>49</sup> Anne-Claude Tabet,<sup>50</sup> Valérie Kremer,<sup>51</sup> Sophie Scheidecker,<sup>51</sup> Adeline Vigouroux,<sup>52</sup> Marilyn Lackmy-Port-Lis,<sup>53</sup> Damien Sanlaville,<sup>54</sup> Marianne Till,<sup>54</sup> Maryline Carneiro,<sup>55</sup> Brigitte Gilbert-Dussardier,<sup>56</sup> Marjolaine Willems,<sup>57</sup> Hilde Van Esch,<sup>58</sup> Vincent Des Portes,<sup>59,60</sup> Salima El Chehadeh<sup>1,2</sup>

► Additional material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/jmedgenet-2017-104956>).

For numbered affiliations see end of article.

### Correspondence to

Dr Salima El Chehadeh, Service de génétique médicale, Hôpitaux Universitaires de Strasbourg, Hôpital de Hautepierre, Strasbourg 67098, France; [salima.elchehadeh@chru-strasbourg.fr](mailto:salima.elchehadeh@chru-strasbourg.fr)

Received 6 September 2017

Revised 4 February 2018

Accepted 15 February 2018

Published Online First

4 April 2018



**To cite:** Miguet M, Faivre L, Amiel J, et al. *J Med Genet* 2018;**55**:359–371.

### ABSTRACT

The Xq28 duplication involving the *MECP2* gene (*MECP2* duplication) has been mainly described in male patients with severe developmental delay (DD) associated with spasticity, stereotypic movements and recurrent infections. Nevertheless, only a few series have been published. We aimed to better describe the phenotype of this condition, with a focus on morphological and neurological features. Through a national collaborative study, we report a large French series of 59 affected males with interstitial *MECP2* duplication. Most of the patients (93%) shared similar facial features, which evolved with age (midface hypoplasia, narrow and prominent nasal bridge, thick lower lip, large prominent ears), thick hair, livedo of the limbs, tapered fingers, small feet and vasomotor troubles. Early hypotonia and global DD were constant, with 21% of patients unable to walk. In patients able to stand, lower limbs weakness and spasticity led to a singular standing habitus: flexion of the knees, broad-based stance with pseudo-ataxic gait. Scoliosis was frequent (53%), such as divergent strabismus (76%)

and hypermetropia (54%), stereotypic movements (89%), without obvious social withdrawal and decreased pain sensitivity (78%). Most of the patients did not develop expressive language, 35% saying few words. Epilepsy was frequent (59%), with a mean onset around 7.4 years of age, and often (62%) drug-resistant. Other medical issues were frequent: constipation (78%), and recurrent infections (89%), mainly lung. We delineate the clinical phenotype of *MECP2* duplication syndrome in a large series of 59 males. Pulmonary hypertension appeared as a cause of early death in these patients, advocating its screening early in life.

### INTRODUCTION

The *MECP2* gene encodes an essential epigenetic regulator in postnatal brain development involved in transcriptional activation as well as repression of many genes.<sup>1–3</sup> Loss of function mutations of *MECP2* in females result in classic Rett syndrome,

## Phenotypes

a progressive childhood neurodevelopmental disorder, while hemizygous mutations found in males are often lethal.<sup>4</sup> Different *MECP2* pathogenic variations have also been described in male patients with a phenotype ranging from severe neonatal encephalopathy to non-specific moderate intellectual disability (ID).<sup>5,6</sup> An increased dosage of methyl-CpG binding protein-2 (MeCP2) resulting from a duplication of the Xq28 region including *MECP2* (*MECP2* duplication) (MIM 300260), leads to a severe neurodevelopmental disorder in males.<sup>7–9</sup> This disorder generally includes early hypotonia, progressive lower limbs spasticity, very poor to absent speech, severe epilepsy, stereotypic movements and recurrent infections. Since the initial report, around 160 male patients have been reported. This has led to the description of a clinically recognisable syndromic ID.<sup>10–12</sup> It has been shown that a small proportion of carrier females could be symptomatic, with a non-specific and mild developmental delay (mostly) or a severe phenotype.<sup>13–17</sup> In a previous study, we reported a series of 20 symptomatic females carrying a *de novo* interstitial *MECP2* duplication (6 novel and 14 reported patients). Thirty-seven per cent had inherited their duplication from their mother, including three mildly, one moderately and three severely affected patients. No correlation was shown between X chromosome inactivation (XCI) in the blood or duplication size and the severity of the phenotype. These findings confirmed that an abnormal phenotype, even severe, can be a rare event in females born to asymptomatic carrier mothers, making genetic counselling difficult in couples at risk.<sup>17</sup>

Previous studies in mice also showed that *MECP2* dosage plays a major role in postnatal neurodevelopment. Mice with twice the endogenous level of MeCP2 appear normal until 10–12 weeks of age, after which they display forepaw stereotypies, impaired coordination, seizures, hypoactivity and spasticity.<sup>18</sup> It has also been shown that increased MeCP2 dosage in mice leads to increased synaptic numbers, having therefore a major impact on the formation of excitatory synapses.<sup>19</sup> In the same way, it has been elegantly demonstrated that cortical neurons derived from different induced pluripotent stem cells (iPSC) lines of patients with *MECP2* duplication syndrome have increased synaptogenesis and dendritic complexity and that their neuronal network synchronisation was altered.<sup>20</sup> These findings contribute to the understanding of the underlying pathways that lead to the very severe neurodevelopmental phenotype of this disease. Furthermore, a more recent study showed that antisense oligonucleotides treatment could rescue many features in adult symptomatic transgenic *MECP2* duplication mice, which appears to be a very promising therapeutic approach.<sup>21</sup>

Although the size of the duplications differs from one patient to another, and despite the involvement of several other brain-expressed genes than *MECP2*, such as *SLC6A8* (MIM 300036), *L1CAM* (MIM 308840), *FLNA* (MIM 300017) and *GDI1* (MIM 300104)—which are responsible for various neurodevelopmental disorders—it has been demonstrated that the increased dosage of *MECP2* is sufficient to explain the core phenotype of *MECP2* duplication syndrome.<sup>10</sup> Indeed, several affected patients harbouring the smallest duplications, only including *IRAK1* and *MECP2*, presented with developmental delay, epilepsy, spasticity and recurrent lung infections.<sup>8,22,23</sup>

The exact prevalence of *MECP2* duplication syndrome remains unknown, although different screening attempts have estimated that it may explain about 1% of severe X-linked intellectual disability (XLID) cases.<sup>24</sup> Very few large series have been reported to date, in particular those regarding the morphological phenotype. Following active data collection from 24 different French clinical genetic departments and 17 cytogenetic and

molecular laboratories, we now report a large French series of affected males with an intrachromosomal Xq28 duplication involving *MECP2*. We describe the clinical phenotype with a focus on the morphological and neurological features of this disorder. We also review the patients reported to date in the literature and report more clinical features.

## PATIENTS AND METHODS

### Data collecting

Patients carrying an Xq28 duplication involving *MECP2* (*MECP2* duplication) were recruited through a French collaborative study conducted from 2012 to 2016, for phenotypical delineation and genotype-phenotype correlation purposes (XLID Research Project 2008–2016, French Ministry of Health). For each patient included, medical data were taken both retrospectively from medical records and as far as possible, cross-sectionally during a dedicated medical consultation. During the latter, the physician asked the parents about the medical history (including the motor skills, behaviour, way of life, kinds of re-education, etc), performed a clinical examination, filled out a detailed questionnaire and took morphological pictures. Since most of the patients have a poor intellectual functioning, no standardised psychometric test could be performed. The degree of ID was determined on the basis of adaptive skills (ambulation, motor development, communication, sociability, daily life autonomy), in accordance with the Diagnostic and Statistical Manual of mental Disorders, fifth edition.<sup>25</sup> Additional investigations were only performed if beneficial for the patient, accounting for missing clinical and radiological data.

### Cytogenetic and molecular analyses

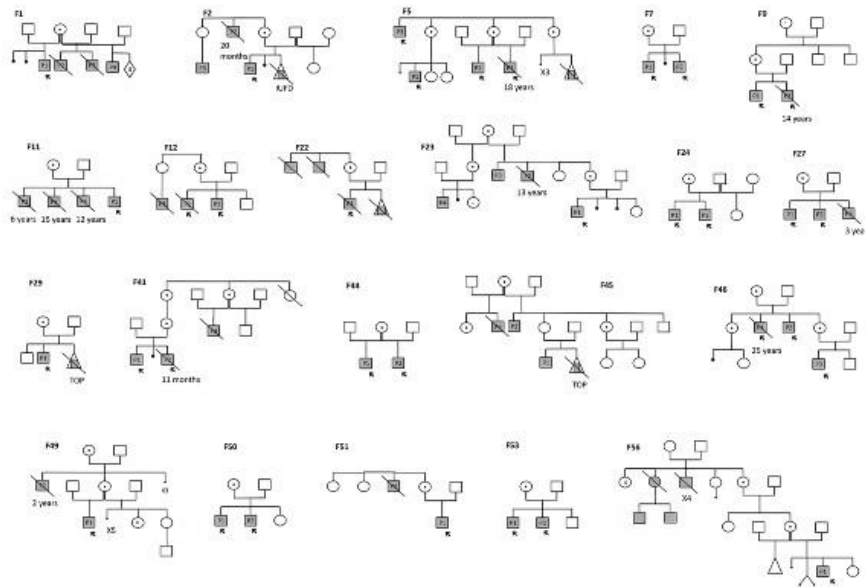
The duplications were detected using a pangenomic (Array Comparative Genomic Hybridization (Array-CGH)) or a targeted approach (Multiplex Ligation-dependent Probe Amplification (MLPA) or quantitative PCR (qPCR) analyses), which were performed in 16 French cytogenetic and molecular laboratories and one Belgian laboratory (Leuven). They were performed on genomic DNA isolated from peripheral blood samples. Regarding Array-CGH, the platform used for most of the patients was the Human Genome CGH Microarray 150K or 180K from Agilent according to the manufacturer's protocol (Agilent Technologies, Santa Clara, California, USA). MLPA was performed using the method described by Schouten *et al* in case the clinicians suspected the *MECP2* duplication syndrome.<sup>26</sup> The duplications were confirmed by fluorescence in situ hybridization (FISH) with various specific probes on chromosome preparations from leucocyte cultures or by qPCR using standard protocols. Parental analyses were performed using FISH or qPCR. Genomic positions were relative to human genome GRCh37/hg19.

All patients' families provided written informed consent and all procedures performed in the studies were done in accordance with the ethical standards of the institutional research committee and with the Declaration of Helsinki.

## RESULTS

### Study population

One hundred and one cases, including 91 males and 10 females (including twins), were identified with a *MECP2* duplication in the context of their diagnostic workup for ID. Among the 91 males, the duplications resulted from interstitial (intraC) Xq28 duplication in 86 cases (54 families) and from a complex chromosomal aberration, such as X-autosome translocation or insertion of the Xq28 region into an autosome, in five cases. Twenty-three out of



**Figure 1** Family trees. The full squares represent the affected males

the 86 patients (27%) died before the age of 25 years (11 months to 24 years, mean age: 9.5 years). Among the 11 patients for whom the cause of the death was known, 8 (73%) died from respiratory infections, 2 (18%) from a pulmonary hypertension and 1 (9%) from status epilepticus. Seventy-three out of the 83 for whom we had the information (88%), had inherited their duplication from their mothers. The pedigrees of the 19 families with an X linked inheritance are illustrated in figure 1. The present study reports only on the 86 males with intraC *MECP2* duplication, with a focus on the 59 patients for whom we were able to collect clinical and morphological data, including photographs. Twenty-five were examined by the same physician (SEC), which could help highlight recurrent clinical features that were often overlooked.

#### Cytogenetic and molecular data

Duplications were detected using genomic microarrays in 25 out of the 54 families (46%) and targeted analysis in 29/54 families (54%). The analysis of the 46 families in which the duplication breakpoints were available showed that the Xq28 duplications were different in size and breakpoint between patients as illustrated in online supplementary figure 1. We observed a recurrent distal breakpoint in nine patients (153609163 bp) suggesting a susceptibility region in Xq28 to such a rearrangement, and 85% (39/46) of the duplications shared a common region of 1.5 Mb located between position 152 485 412 and position 153 949 811 (hg19) (online supplementary figure 1 and table 1). The mean size of the duplications was 1856 kb (range: 102–16 704 kb), and the main genes included in the duplications were *MECP2*, *IRAK1*, *GDI1*, *FLNA* and *LICAM*. Since the patients with the smallest duplications did not have a milder neurological phenotype, we could not show an obvious correlation between the size of the duplications and the severity of the phenotype. The *FMR1* gene was involved only in the three patients harbouring the largest duplications (online supplementary figure 1 and table 1). XCI

studies showed that all of the carrier mothers for whom data were available had completely skewed XCI (XCI ratio >80:20) (data not shown).

#### Clinical data

We were able to collect clinical data for 59 out of the 86 male patients harbouring an intraC *MECP2* duplication. Among them, nine (15%) died before the age of 25 years. The mean age at examination was 11.7 years (range: 2 months to 48 years), with 13 adult patients (age  $\geq 18$  years). The clinical data are summarised in tables 1 and 2 and online supplementary table 2.

#### Perinatal period

The mean term of birth was 39 weeks of gestation (WG) (range: 32–41). Four patients were born prematurely (<37 WG). Birth parameters were within the normal range with a mean weight at  $-0.1$  SD score (SDS) ( $-2.6$  SDS;  $+3.5$  SDS), mean height at  $-0.7$  SDS ( $-4.5$  SDS;  $+1.9$  SDS) and mean head circumference at  $+0.1$  SDS ( $-2.5$  SDS;  $+4.5$  SDS). The Apgar score was abnormal (<8) for 10% (5/50) of the newborns and 20% (10/50) were admitted to either a neonatology department or an intensive care unit. Later, 55% (32/58) of the patients had complications during the neonatal period, including early hypotonia (11/58), feeding difficulties (15/58) or infection (4/58) (online supplementary table 2).

#### Growth parameters and craniofacial features

The growth parameters were mainly in the normal range. Mean occipitofrontal circumference (OFC) was  $+0.2$  SDS ( $-3.4$  SDS;  $+4.8$  SDS). Four out of 53 (8%) patients had microcephaly (OFC  $< -2$  SDS) and 9/53 (17%) were macrocephaly (OFC  $> +2$  SDS) (table 1). The mean height was  $-0.4$  SDS ( $-5.2$  SDS;  $+2.8$  SDS) and the mean weight was  $+0.4$  SDS ( $-4.7$

## Phenotypes

**Table 1** Summary of the main facial and limb morphological features from the present cohort

Main facial and limb features	Present	Percentage (%)
Midface hypoplasia	42/45	93
Microcephaly (OFC <-2 SDS)	4/53	8
Macrocephaly (OFC >+2 SDS)	9/53	17
Hair		
Sparse anteriorly	29/42	69
Thick and dense	39/43	91
Eyes		
Deep set eyes	25/43	58
Strabismus	41/54	76
Sparse eyebrows	23/43	54
Synophrys	11/43	26
Epicanthus	18/44	41
Ptosis	13/43	30
Hypertelorism	20/44	45
Mouth and teeth		
Open mouth appearance	44/45	98
Drooling	40/50	80
Small mouth	32/46	70
Thick lower lip	36/46	78
Prominent central incisors	16/32	50
Teeth anomalies	32/34	94
Persistence of deciduous teeth	23/26	88
Nose		
Anteverted nostrils	23/44	52
Prominent nasal tip	29/43	67
Prominent nasal bridge	32/42	76
Narrow nasal bridge	28/44	64
Ears		
Large prominent ears	39/46	85
Limbs		
Small hands and feet	21/34	62
Tapered fingers	36/39	92
Vasomotor troubles	15/31	48
Livedo of the limbs	34/39	87
Valgus flat feet	26/35	74

OFC, occipitofrontal circumference; SDS, SD score.

SDS; +5.9SDS) (table 2). Most of the patients had particular facial features (51/55), summarised in table 1 and illustrated in figures 2 and 3. We noted that most of the affected males had a similar facial gestalt, characterised by a midface hypoplasia (93%), deep set eyes (58%), hypertelorism (45%), a small (70%) open mouth (98%) with a thick lower lip (78%) and intermittent drooling (80%), large prominent ears (85%), narrow (64%) and prominent (76%) nasal bridge with anteverted nostrils (52%). Hair was often thick and dense (91%) with a sparse area on the lateral part of the forehead (69%) (figure 2, patients F53P2, F5P1, F55P1), which led to a temporal spike aspect, and the eyebrows were often sparse (54%) (figure 2, patients F29P1, F40P1, F5P4) with a synophrys sometimes apparent (26%). A bilateral epicanthus was noted in 41% of the patients. We noted that patients with a light skin phototype had particularly pale and bright blue eyes. Other recurrent features were dental shape and/or eruption abnormalities in 94% of the series, including prominent central incisors (50%), and persistence of the deciduous teeth (88%), often requiring multiple tooth extractions (table 1, figures 2 and 3).

We found that 62% of the patients had abnormally small feet which were often flat with a valgus deformation (74%). The hands showed tapered fingers in 36/39 (92%) patients and a camptodactyly of several fingers was noticed in few cases (figure 4B). Frequent vasomotor problems were noted (48%) leading to moist and red hands and/or feet (figure 4C). We also noted an apparent venous network (livedo) of the limbs in 34/39 (87%) of the children, which tended to be less evident after adolescence (figure 4D) (table 1).

Interestingly, in the 45 patients for whom photographs were available from early infancy (<3 years) to school age (6–14 years) or adulthood, we were able to observe, as often in syndromology, an evolution of the facial features with ageing. In early infancy, the face looked mostly round with a small chin and full cheeks and the mouth already appeared small, with downturned corners, a convex upper lip and open appearance with frequent drooling visible (figures 3 and 5). With time, the face became elongated, with a prominent chin and a larger mouth with sometimes prominent central incisors. The nose was initially particularly small with anteverted nostrils and changed with frequent prominent nasal bridge and tip. Facial features tended to become coarser and eyes appeared deeper after adolescence and during adulthood, although facial dysmorphism was very mild in some patients (figure 5).

### Hearing and vision

Seven out of 23 (30%) patients who were assessed had mild-to-moderate hearing loss (table 2). The type of the deafness was specified in three cases only (two patients had conductive deafness in a context of numerous otitis and one had perceptive deafness). A divergent strabismus was frequent (41/54, 76%) and could be related to visual impairment since 54% (21/39) of the patients had hypermetropia. Three patients had corneal anomalies (scars and sclerocornea). A ptosis was noted in 30% of the patients.

### Neurodevelopmental and orthopaedic features

During childhood, almost all patients had early hypotonia (57/58) (online supplementary figure 2) and all had global DD (59/59). Motor skills were severely affected as 12/56 (21%) were unable to achieve ambulation without support. For the remaining patients, the mean age of independent walking was delayed (3.6 years (range: 1.8; 5.4)). Among them, seven lost the ability to walk. Patients developed orthopaedic symptoms that could be, at least partially, the consequences of lower limbs spasticity and hypotonia. One of the main features was kyphoscoliosis, present in 53% of them (23/43). Another striking feature was a lower limbs spasticity leading to a singular habitus with a flexion and sometimes contractures of the ankles, the knees and the trunk. This spasticity was not often associated with pyramidal signs as Babinski sign or hyper-reflexia but led to an abnormal pseudoataxic gait with a degree of instability and a broad-based stance (figure 4A, table 1). We hypothesise that this unstable gait is partly due to an abnormal development of the proprioceptive pathways. This leads to a shift of the centre of gravity of the body which is partially compensated by the flexion attitude.

### Epilepsy

Seizures were seen in 59% (35/59) of individuals with a mean onset around 7.4 years of age (0.4–35). Seizure types were generalised tonic-clonic seizures (19/35), absences (5/35), partial complex (6/35), atonic (12/35) and myoclonic seizures (8/35). Sixty-two per cent (21/34) had drug-resistant epilepsy.





**Figure 2** Facial phenotype of the patients with an interstitial Xq28 duplication involving *MECP2*. Patients share common facial features including an open mouth with a thick lower lip and intermittent drooling, large ears, midface hypoplasia, deep set eyes with frequent divergent strabismus, sparse eyebrows, narrow and prominent nasal bridge, dense and thick hair with an anterior sparse area.

In many patients, active epilepsy did not stop in late adolescence, since 12 of the 14 oldest living patients of our cohort (16–49 years), still experience seizures, including 10 who were

drug-resistant. Among 31 patients for whom data were available, a neurological regression was observed after seizure onset in 12 patients (39%), including 11 patients with drug-resistant



**Figure 3** Focus on the lower part of the face in patients with an interstitial Xq28 duplication involving *MECP2*. Note the small mouth and the open mouth appearance with a convex upper lip during early infancy (A, B), and the prominent central incisors after childhood (C).



## Phenotypes



**Figure 4** Common habitus with a flexion of the trunk and the knees (A), slender hands with tapering fingers and camptodactyly in few cases (F9P1, F44P1) (B), vasomotor troubles in the hands which appear red and wet (C), apparent venous network (livedo) of the limbs (D).

epilepsy. Among the 42/54 (78%) patients with the most severe developmental delay (5 youngest patients excluded), 14 (33%) did not develop epilepsy.

### Communication and behaviour

All of the patients had ID, which was severe in most of them. A majority of patients were dependent for daily life needs (dressing, bathing, toileting, etc). Only five children became toilet trained but two of them lost this skill. Among the patients for whom data were available, the majority did not develop speech (32/57, 56%) or had just babbling (4/57, 7%) or few words (20/57, 35%). However, we observed that for a few patients who had highly active familial and paramedical stimulations and re-education, a communicative strategy could be elaborated, including the use of pictograms. Many patients had stereotypies (49/55, 89%), including flapping, midline hands movements, head

movements and 33/46 (72%) had bruxism. These stereotypic movements tended to appear during school age and persisted, in a variable way, throughout life. Regarding the autistic features, we could observe that, even though most of the patients had stereotypies and poor to absent language, with regression of skills for some, a considerable proportion demonstrated several interactions during the consultation, without social withdrawal, they smiled frequently and were interested in trying to carry out simple tasks.

In addition, a majority of parents reported to us decreased pain sensitivity (29/37, 78%) (table 2). The parents of two children who had broken their leg mentioned that they had not shown any pain.

### Other medical issues

Not surprisingly, one of the most frequent features was chronic constipation, which was found in 78% of the



**Figure 5** Evolution of the phenotype with time. Evolution of the facial features from early infancy to school age or adulthood (from left to right). Note the changing facial phenotype from round face with small mouth and nose to elongated coarser face with prominent nose and ears and thick lower lip.

patients (43/55), including in some of them, subobstruction episodes or even chronic intestinal pseudo-obstruction (CIPO) syndrome. When rectal biopsies were performed they were normal. Most of the patients had early feeding difficulties (34/56, 61%). They were partially due to an important gastro-oesophageal reflux (34/51, 67%) and to swallowing difficulties. These difficulties often disappeared with ageing and most of them were later described as having a very good appetite (table 2).

The majority of patients (49/55, 89%) had recurrent and severe respiratory infections requiring numerous hospitalisations, and recurrent ENT (pharyngitis, otitis, sinusitis) and urinary tract infections. These infections and their complications were the primary cause of premature death (table 2). We also noted significant obstructive sleep apnoea (16/49,

33%), leading to adenoidectomy or tonsillectomy and sometimes to non-invasive mechanical ventilation.

Three patients (F41P2, F27P1, F27P3) suffered from pulmonary hypertension, which led to premature death, unrelated to a severe pulmonary infection, in two of them at the age of 11 months and 3 years, respectively. Another (F12P1), with right ventricular hypertrophy, died at the age of 19 years. One patient had been operated on for an atrial septum defect, one had a situs inversus and another had Bouveret tachycardia.

Urogenital anomalies were present in 43% of the patients (23/53) and included unilateral or bilateral cryptorchidism in 26% (14/53), micropenis in 9% (5/53), bilateral ureteral dilation in 9% (5/53) and bladder dilation or hypertrophy in three patients (table 2).

Gynecomastia was observed in three patients.

## DISCUSSION

Two large series of patients harbouring a *MECP2* duplication have been reported to date,<sup>10,12</sup> including 129 and 56 affected patients, respectively. In addition, we previously reported the MRI data of 30 affected males and Bauer *et al* reported the immunological phenotype of 27 affected patients.<sup>27</sup> Our present series describes the demographic and clinical features of 59 French male patients with a focus on the morphological and neurodevelopmental features in order to improve the phenotypical characterisation of this syndrome. However, given that this disorder can be easily diagnosed using Array-CGH, which is commonly requested in the diagnosis of ID, the results presented in this study will be of value for the follow-up of patients.

There was a wide age range, from 2 to 48 years, the latter being the oldest male patient diagnosed to date with *MECP2* duplication syndrome to our knowledge, Friez *et al* having reported a man aged 46 years.<sup>28</sup> Among the whole French series of 86 male patients with an intraC *MECP2* duplication, 27% died before the age of 25 years (11 months to 24 years, mean age: 9.5 years), from respiratory infections for 73% of them—which is similar to what has already been reported<sup>10,12</sup>—pulmonary hypertension for 18% and epilepsy for 9%. Like other authors, we were not able to demonstrate a clear correlation between the size of the duplication and the severity of the neurodevelopmental phenotype as the patients in our series with the smallest duplication did not have a milder neurological phenotype.<sup>15,22,29</sup>

We observed that the large majority of the patients were born at term and with normal birth parameters (online supplementary table 2). Few fetal anomalies were noted during pregnancies, including brain malformations, intrauterine growth retardation in two cases and bilateral hydronephrosis in two cases. In the same way, Fu *et al* reported four fetuses harbouring a *MECP2* duplication and indicated that ventriculomegaly, hydrocephalus, agenesis of the corpus callosum, choroid plexus cysts, fetal growth retardation and hydronephrosis might be recurrent obstetrical ultrasound findings in fetuses with this disorder.<sup>30</sup> We observed obstetrical complications in 12.5% of pregnancies, including polyhydramnios (7.8%), premature delivery threats (5.9%) or toxæmia (3.9%). Lim *et al* observed the same proportion of polyhydramnios and more hypertensive disorders (19%).<sup>12</sup>

We had previously described the OFC values in a smaller series of 30 patients and shown that there were nearly as many with microcephaly (13%) as with macrocephaly (17%).<sup>31</sup> In the present series, we found that 7.5% of the patients were microcephalic and 17% were macrocephalic, which confirms that the OFC is not a key clinical feature in this syndrome (table 1). *MECP2* duplication syndrome is classically described by a core phenotype that includes severe developmental delay with absent or very poor speech, epilepsy, respiratory infections and chronic constipation.<sup>7,9</sup> Additionally, through a detailed analysis of the patients' phenotype during consultations and photographs taken at different ages, our study showed other similarities emerging between the patients affected by this condition including distinctive facial features and dental anomalies including shape and eruption anomalies (table 1, figures 2 and 3). However, these signs alone are not sufficient to raise the diagnosis of *MECP2* duplication clinically. We studied the evolution of the facial features with ageing and noted that they became coarser and that the syndrome was more difficult to recognise in adulthood, as usual in syndromology.<sup>32</sup>

In the literature, these patients are frequently reported to be ataxic.<sup>7,10</sup> We noted in the present study that they mainly had a flexion attitude of the knees with a broad-based stance, mainly due to both hypotonia and spasticity, which led to a peculiar and unsteady gait (figure 4A). Another particularity was a decreased sensitivity to pain (29/37), which was also reported by Lim *et al* in 50% of their series.<sup>12</sup>

We also noticed frequent vasomotor problems, and a livedo of the limbs. Some of these signs had already been reported individually, like in the patient reported by Clayton-Smith *et al* who was described with a thin appearing skin and cold red feet with poor circulation.<sup>7</sup> The association of decreased sensitivity to pain, chronic constipation, vasomotor troubles and unsteady gait with an abnormal posture could possibly result from an abnormal development of the proprioceptive and nociceptive pathways. Most of the patients developed kyphoscoliosis, which is important to take into account in their clinical care in order to start physiotherapy early in life and prescribe an orthopaedic corset when necessary. It seems important to note that some of these clinical features, including vasomotor troubles, small cold hands and feet, decreased response to pain, the stereotypic movements, bruxism and gait abnormalities are also supportive criteria of Rett syndrome. However, in *MECP2* duplication syndrome, the development is abnormal from the outset and we do not observe frequently a deceleration of the OFC or breathing disturbances when awake.<sup>33</sup>

In our series, among the 32/59 patients with epilepsy, 62% were drug-resistant, compared with 76% in the series of Lim *et al*.<sup>12</sup> Multiple different seizure types (table 2) and different EEG patterns (data not shown) were observed, which is consistent with other reports.<sup>24,34,35</sup> Peters *et al* previously described that among 17 males with *MECP2* duplication syndrome, 7 (41%) exhibited regression in motor skills, which coincided with seizure onset in 6 of them, compared with 12/31 (39%) patients in the present series.<sup>36</sup> We observed very frequent bruxism and hand stereotypies as did other authors.<sup>12,24,29</sup> Compared with typical Rett syndrome, in which the onset of stereotypies mostly precedes or coincides with the loss of purposeful hand movements between ages 1 and 4 years (stage II) and tends to decrease after age 10 years,<sup>37,38</sup> stereotypies in *MECP2* duplication patients tend to appear during school age and will persist, in a variable way, throughout life.

Although some *MECP2* duplication patients have been reported as having autistic behaviour,<sup>8,29</sup> we observed that most of them had good social interactions, although we could not confirm this clinical impression with standardised diagnostic tool since our series was partly retrospective. In the same way, Xi *et al* did not find any *MECP2* duplication in a series of 82 autistic males, suggesting that this disorder is not frequent in patients with autism.<sup>39</sup> Conversely, hearing loss is not often reported in the literature in *MECP2* duplication patients. Since we found hearing impairment in 7/23 patients in our cohort, this sensory impairment must be systematically assessed to avoid worsening of communication skills.

Digestive problems, ranging from chronic constipation to CIPO, are well-known features in *MECP2* duplication patients. Lim *et al* noted that chronic constipation was present in 80% of the subjects in their series.<sup>12</sup> This is consistent with our findings (78%), but seems to be more frequent than in previous reports (table 3).<sup>10,11</sup> Several hypotheses were raised on the origin of the chronic constipation. Clayton-Smith *et al* attributed this to the *FLNA* gene, often involved in the duplication, and supported by the fact that point mutations in this gene were found in X linked families with pseudo-obstruction.<sup>7,40,41</sup> Fernandes *et al*, on the

**Table 3** Comparison of the clinical data from the present study and those reported by Van Esch *et al* and Lim *et al*<sup>10–12</sup>

	Van Esch <i>et al</i> (%)	Lim <i>et al</i> (%)	Present study (%)
Patient number	129	49	59
Gender	M	M	M
Age at diagnosis (years)	NR	3	10
Age at examination (years)	NR	7.9	12
Neonatal complications	NR	33*	32/58 (55)
Early death (<25 years)	34/88 (39)	NR	9/59 (15)
Facial dysmorphism	65/86 (76)	NR	51/55 (93)
<b>Growth</b>			
Head circumference (SDS)	Normal range	NR	0.2
Height (SDS)	Normal range	NR	-0.4
Weight (SDS)	Normal range	NR	0.4
<b>Neurodevelopment</b>			
Developmental delay	129/129 (100)	41/49 (84)	59/59 (100)
Early hypotonia	110/117 (94)	31/48 (65)	57/58 (98)
Seizures	66/123 (54)	21/49 (43)	35/59 (59)
Drug resistance	NR	17/21 (81)	21/34 (62)
Age at onset of seizures (years)	6	1 to 20	7.4
Type of seizures	TC, At, Ab, M	NR	TC, At, Ab, M, CP
Walking age (years)	18 months to 4 years	4	3.6
Walking without support	83/109 (76)	13/49 (27)	37/56 (66)
Absent speech	86/109 (79)	23/46 (50)	32/57 (56)
Stereotyped movements	20/78 (26)	55% at 3 years	49/55 (89)
<b>Gastrointestinal anomalies</b>			
Chronic constipation	29/88 (33)	37/45 (82)	43/55 (78)
Gastroesophageal reflux	NR	26/49 (53)	34/51 (67)
Recurrent pulmonary infections	87/121 (72)	38/49 (78)	49/55 (89)
<b>Skeletal anomalies</b>			
Flexion of the knees and the trunk	57/89 (64)	NR	42/49 (86)
Scoliosis and/or kyphosis	NR	10/46 (22)	23/43 (53)
Decreased pain sensitivity	NR	25/50 (50)	29/37 (78)

\*Includes both males and females.

At, atonic; M, myoclonic; Ab, absence; CP, complex partial; TC, tonic-clonic; NR, not reported; SDS, SD score.

other hand, hypothesised that this feature may be linked to the duplication of *LICAM*, which is expressed in the enteric nervous system, plays a key role in its development during embryogenesis and has been found to be mutated in some patients with either X linked hydrocephalus or acrocallosal syndrome and Hirschsprung disease.<sup>42–44</sup> In our series, among 39 patients with chronic constipation and for whom molecular data were available, we found that the duplication included *FLNA* and *LICAM* in 25 cases, only *LICAM* or only *FLNA* in 9 and 4 cases respectively, and none of them in 1 case. Therefore, in addition to the frequent constipation rate described in patients with severe neurological handicap, we may speculate whether increased dosage of both *FLNA* and *LICAM* genes could be responsible for the increased risk to develop a severe constipation in these patients.

Recurrent and severe infections were a significant health issue for 89% of *MECP2* duplication patients, as in the study by Lim *et al*, who found recurrent respiratory infections in 78% of affected males, three-quarters of them requiring hospital admissions within the first 2 years of life for supportive care and interventional

procedures (mechanical ventilation, intravenous antibiotics, tracheostomy).<sup>12</sup> Van Esch, in a review of 129 affected males, reported 70% of recurrent infections, especially of the respiratory tract.<sup>10</sup> These results raise again the question of the use of a long-term prophylactic antibiotic therapy in these males. It has been shown that increased susceptibility to infections in *MECP2* duplication syndrome is associated with defects in the interferon- $\gamma$ -mediated TH1 response,<sup>45</sup> IgA/IgG2 deficiency, low antibody titres against pneumococci and elevated acute-phase responses. Some authors suggest prophylactic IgG substitution in *MECP2* duplication patients with low IgA/IgG2 rates.<sup>27</sup>

Three patients in our series suffered from pulmonary hypertension (PH), being lethal for two of them. Two other unrelated *MECP2* duplication patients who had lethal pulmonary hypertension, have been reported (rarechromosome.org—Xq28 duplication; <http://www.goldcoastbulletin.com.au/news/gold-coast/gold-coast-family-mourns-11-month-old-after-mystery-heart-attack/news-story/4d1f28630e5b6780359721cad0e41988>). Interestingly, Cronk *et al* showed that Influenza A infected *MeCP2*<sup>Tg3</sup> mice, overexpressing *MeCP2* at levels threefold to fivefold higher than normal, experienced PH and arterial narrowing, which contributed to infection-related mortality.<sup>46</sup> In our series, we speculate that PH might be the cause of the sudden and unexplained death of some patients, especially as *MECP2* duplication patients do not usually undergo regular cardiac evaluation. We believe that these findings are important enough to introduce a cardiac assessment in these patients from the first year of life and then regularly, especially in patients who experienced several pulmonary infections.

Urogenital anomalies, in particular cryptorchidism and micropenis, were more frequent (26%) than in the literature,<sup>8 23 44 47 48</sup> and gynecomastia was present in three patients, suggesting the possible existence of hypogonadism in this syndrome.

In conclusion, we report on the clinical and morphological features of 59 affected males with interstitial *MECP2* duplication in order to expand the clinical phenotype of *MECP2* duplication syndrome. Most of the clinical findings were consistent with previous data and we demonstrated that these patients shared common facial features that tend to evolve with age. In addition, we described commonly associated morphological and neuro-orthopaedic features not reported to date, such as a trunk and knees-flexion habitus leading to an ataxic-like gait. Pulmonary hypertension appeared as a cause of early death in these patients, which suggests that patients may benefit from regular cardiac assessment.

#### URLs

USCS, <https://genome.usc.edu/>

#### Author affiliations

<sup>1</sup>Service de génétique médicale, Institut de Génétique Médicale d'Alsace (IGMA), Centre de Référence Maladies Rares "Anomalies du développement et syndromes malformatifs", Centre de Référence Maladies Rares "Des déficiences intellectuelles de causes rares", Hôpitaux Universitaires de Strasbourg, Hôpital de Hautepierre, Strasbourg, France

<sup>2</sup>FHU TRANSLAD, Centre de Référence Maladies Rares «Anomalies du développement et syndromes malformatifs», Centre de Génétique, CHU de Dijon, Dijon, France

<sup>3</sup>Service de Génétique Clinique, Hôpital Necker Enfants Malades, APHP, Paris, France

<sup>4</sup>Service de Génétique Clinique, CHU de Saint-Etienne, Saint-Etienne, France

<sup>5</sup>Service de Génétique Clinique, CLAD Ouest, CHU de Rennes, Rennes, France

<sup>6</sup>Laboratoire de Génétique Moléculaire, CHU de Rennes, Rennes, France

<sup>7</sup>Service de Génétique Médicale, CHU de Toulouse, Toulouse, France

<sup>8</sup>Service de Neuropédiatrie, CHU de Clermont-Ferrand, Clermont-Ferrand, France

<sup>9</sup>Service de Génétique Médicale, CHU de Clermont-Ferrand, Clermont-Ferrand, France

<sup>10</sup>Service de Neuropédiatrie et Maladies Métaboliques, Hôpital Robert Debré, APHP, Paris, France

<sup>11</sup>Service de Génétique Médicale, CHU de Nantes, Nantes, France

## Phenotypes

- <sup>17</sup>Service de Génétique Médicale, CHU de Nancy, Nancy, France  
<sup>13</sup>Laboratoire de Génétique Médicale, CHU de Nancy, Nancy, France  
<sup>14</sup>Service de Neuropédiatrie, CHRU de Lille, Lille, France  
<sup>15</sup>Laboratoire de Génétique Médicale, Hôpital Jeanne de Flandre, CHRU de Lille, Lille, France  
<sup>16</sup>Service de Génétique, CHU de Caen, Caen, France  
<sup>17</sup>Service de Génétique, CHRU de Tours, Tours, France  
<sup>18</sup>Service de Radiologie, Hôpital Femme Mère Enfant, Bron, France  
<sup>19</sup>Laboratoire de Biologie Moléculaire, Hôpital Necker Enfants Malades, APHP, Paris, France  
<sup>20</sup>Service de Neurologie pédiatrique, CHU de Montpellier, Montpellier, France  
<sup>21</sup>Service de Génétique, Centre Hospitalier de Vannes, Vannes, France  
<sup>22</sup>Service de Génétique, Hôpital Armand Trousseau, APHP, Paris, France  
<sup>23</sup>Laboratoire de Génétique Moléculaire, GH Cochin-Broca Hôtel Dieu, APHP, Paris, France  
<sup>24</sup>Service de Génétique Clinique, Hôpital Robert Debré, APHP, Paris, France  
<sup>25</sup>Laboratoire de Biologie Moléculaire, Hôpital Robert Debré, APHP, Paris, France  
<sup>26</sup>Département de Génétique et Procréation - UMR CNRS 5525 TIMC-IMAG - équipe DYCTIM, CHU Grenoble, Grenoble, France  
<sup>27</sup>Université de Bordeaux, Laboratoire MRGM, INSERM U1211 and Service de Génétique Médicale, CHU de Bordeaux, Bordeaux, France  
<sup>28</sup>Département de Génétique Médicale, Hôpital de la Timone, Marseille, France  
<sup>29</sup>Laboratoire de Génétique Chromosomique, Hôpital de la Timone, Marseille, France  
<sup>30</sup>Laboratoire de Biologie Moléculaire, Hôpital de la Timone, Marseille, France  
<sup>31</sup>CIC1407 Inserm, CHU de Lyon, Lyon, France  
<sup>32</sup>Réseau d'Analyse Chromosomique sur Pucés à ADN (AChro-Puce)  
<sup>33</sup>Unité de Neuropédiatrie, CHU Pellegrin, Bordeaux, France  
<sup>34</sup>Clinique de Génétique Guy Fontaine, CHRU de Lille, Hôpital Jeanne de Flandre, Lille, France  
<sup>35</sup>Centre de Génétique Chromosomique, GH de l'Institut Catholique de Lille, Hôpital Saint-Vincent-de-Paul, Lille, France  
<sup>36</sup>Service de Neuropédiatrie, Hôpital Necker Enfants Malades, APHP, Paris, France  
<sup>37</sup>Département de Génétique Médicale, Centre de Référence "Malformations et maladies congénitales du cerveau", APHP, Hôpital Armand Trousseau, APHP, Paris, France  
<sup>38</sup>Service de Génétique Clinique, Hôpital Pitié-Salpêtrière, APHP, Paris, France  
<sup>39</sup>Service de Pédiatrie, CHU Félix Guyon, Saint-Denis, France  
<sup>40</sup>UF génétique médicale, Centre Hospitalier du Mans, Le Mans, France  
<sup>41</sup>Service de génétique, Hospices Civils de Lyon, Lyon, France  
<sup>42</sup>INSERM U1028, CNRS UMR5292, Centre de Recherche en Neurosciences de Lyon, GENDEV Team, Université Claude Bernard Lyon 1, Lyon, France  
<sup>43</sup>Laboratoire de Génétique Moléculaire, CHRU de Tours, Tours, France  
<sup>44</sup>Laboratoire de Cytogénétique, CHU de Dijon, Dijon, France  
<sup>45</sup>Laboratoire de Cytogénétique, CHU de Grenoble, Grenoble, France  
<sup>46</sup>Laboratoire de Cytogénétique, CHU de Nantes, Nantes, France  
<sup>47</sup>Sarcomes osseux et remodelage des tissus calcifiés, Université Bretagne Loire, INSERM, UMR1238, Nantes, France  
<sup>48</sup>Laboratoire de Cytogénétique, Hôpital Necker Enfants Malades, APHP, Paris, France  
<sup>49</sup>Laboratoire de Cytogénétique, Hôpital Pitié-Salpêtrière, APHP, Paris, France  
<sup>50</sup>Laboratoire de Cytogénétique, Hôpital Robert Debré, APHP, Paris, France  
<sup>51</sup>Laboratoire de Cytogénétique, CHU de Strasbourg, Hôpital de Haute-pierre, Strasbourg, France  
<sup>52</sup>Laboratoire de Cytogénétique, CHU de Toulouse, Toulouse, France  
<sup>53</sup>Service de Génétique Médicale, CHU de Pointe à Pitre, Pointe à Pitre, France  
<sup>54</sup>Laboratoire de Cytogénétique, CHU de Lyon, Lyon, France  
<sup>55</sup>Service de Neuropédiatrie, CHU de Lyon, Hôpital Femme-Mère-Enfant, Lyon, France  
<sup>56</sup>Service de Génétique Médicale, CHU de Poitiers, Poitiers, France  
<sup>57</sup>Service de Génétique Médicale, CHU de Montpellier, Montpellier, France  
<sup>58</sup>Laboratory for Genetics of Cognition, Center for Human Genetics, University Hospitals Leuven, Leuven, Belgium  
<sup>59</sup>Centre de Référence Maladies Rares «Des déficiences intellectuelles de causes rares», HFME, Hospices Civils de Lyon et Université de Lyon, Lyon, France  
<sup>60</sup>Institut des Sciences Cognitives, CNRS UMR 5304, Bron, France

**Acknowledgements** The authors are deeply grateful to the patients and their parents for their participation in this study. The authors would like to thank the association Xtraordinaire for their financial support of the project, and the French Ministry of Health (PHRC RMLX 2008, Pr V. Des Portes) for its support. The authors would like to thank the AChro-Puce network for its help in collecting cytogenetic data. The authors would also like to thank Mr James Garfield for his helpful review of this article.

**Contributors** MM wrote the paper, MN, CD, CP, JA, JPB, SCB, TB, MT, SDS, SD, AM, CB, CRT, GL, MR, CM, JT, MBD, DS, PC, ALM, NM, CC, VS, VM, SR, BK, ACT, VK, SS, AV and RAP analysed the cytogenetic and molecular data. LF, JA, RT, FP, LP, ML, SJ, CS, CE, FL, OBT, AD, BI, JV, BL, LL, JMC, GP, AJ, AG, VCD, MR, BE, GR, HJ, LB, CB, LP, PSJ, KD, FD, DL, NP, SS, CTR, JMP, CG, CVD, BDB, NBB, AA, KM, CLC, DH, JLA, GL, MLPL, MR, BGD, MW, MC, HVE and DMC are the referring clinicians. LG interpreted brain

MRI of many patients of the cohort, VD and NP respectively directed and organised the PHRC RMLX, SEC collected the clinical and molecular data and the photographs during consultations (for 25 patients) or by contacting the clinicians and the labs, and supervised this work.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient consent** Detail has been removed from this case description/these case descriptions to ensure anonymity. The editors and reviewers have seen the detailed information available and are satisfied that the information backs up the case the authors are making.

**Ethics approval** The ethics committee "Comité de Protection des Personnes SUD-EST II" delivered a favourable opinion (22 September 2010) with regard to the realisation of the XLID Research Project 2008-2016 (RMLX).

**Provenance and peer review** Not commissioned; externally peer reviewed.

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

## REFERENCES

- Adkins NL, Georgel FT. MeCP2: structure and function. *Biochem Cell Biol* 2011;89:1–11.
- Chahrouh M, Jung SY, Shaw C, Zhou X, Wong ST, Qin J, Zoghbi HY. MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science* 2008;320:1224–9.
- Gonzales ML, LaSalle JM. The role of MeCP2 in brain development and neurodevelopmental disorders. *Curr Psychiatry Rep* 2010;12:127–34.
- Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* 1999;23:185–8.
- Villard L. MECP2 mutations in males. *J Med Genet* 2007;44:417–23.
- Lambert S, Maystadt I, Boulanger S, Vrielynck P, Destreé A, Lederer D, Moortgat S. Expanding phenotype of p.Ala140Val mutation in MECP2 in a 4 generation family with X-linked intellectual disability and spasticity. *Eur J Med Genet* 2016;59:522–5.
- Clayton-Smith J, Walters S, Hobson E, Burkitt-Wright E, Smith R, Toutain A, Amiel J, Lyonnet S, Mansour S, Fitzpatrick D, Ciccone R, Ricca I, Zuffardi O, Donnai D. Xq28 duplication presenting with intestinal and bladder dysfunction and a distinctive facial appearance. *Eur J Hum Genet* 2009;17:434–43.
- del Gaudio D, Fang P, Scaglia F, Ward PA, Craigen WJ, Glaze DG, Neul JL, Patel A, Lee JA, Irons M, Berry SA, Pursley AA, Grebe TA, Freedenberg D, Martin RA, Hsieh GE, Khara JR, Friedman NR, Zoghbi HY, Eng CM, Lupski JR, Beaudet AL, Cheung SW, Roa BB. Increased MECP2 gene copy number as the result of genomic duplication in neurodevelopmentally delayed males. *Genet Med* 2006;8:784–92.
- van Esch H, Bauters M, Ignatius J, Jansen M, Raynaud M, Hollanders K, Lugtenberg D, Biemvenu T, Jensen LR, Geetz J, Moraine C, Marynen P, Fyns JP, Froyen G. Duplication of the MECP2 region is a frequent cause of severe mental retardation and progressive neurological symptoms in males. *Am J Hum Genet* 2005;77:442–53.
- van Esch H. MECP2 Duplication Syndrome. *Mol Syndromol* 2012;2:128–36.
- van Esch H. MECP2 Duplication Syndrome. *GeneReviews(R)*. Seattle (WA): University of Washington, Seattle, 1993.
- Lim Z, Downs J, Wong K, Ellaway C, Leonard H. Expanding the clinical picture of the MECP2 duplication syndrome. *Clin Genet* 2017;91:557–63.
- Grasshoff U, Bonin M, Goehring I, Ekioglu A, Dufke A, Cremer K, Wagner N, Rossier E, Jauch A, Walter M, Bauer C, Bauer P, Horber K, Beck-Woedl S, Wiczorek D. De novo MECP2 duplication in two females with random X-inactivation and moderate mental retardation. *Eur J Hum Genet* 2011;19:507–12.
- Bijlsma EK, Collins A, Papa FT, Tejada MI, Wheeler P, Peeters EA, Gijbbers AC, van de Kamp JM, Kriek M, Losekoot M, Broekma AJ, Crolla JA, Pollazzon M, Mucciolo M, Katzaki E, Disciglio V, Ferrieri MI, Marozza A, Mencarelli MA, Castagnini C, Dosa L, Ariani F, Mari F, Canitano R, Hayek G, Botella MP, Gener B, Minguez M, Renieri A, Riuvenkamp CA. Xq28 duplications including MECP2 in five females: Expanding the phenotype to severe mental retardation. *Eur J Med Genet* 2012;55(6–7):404–13.
- Shimada S, Okamoto N, Ito M, Arai Y, Momosaki K, Togawa M, Maegaki Y, Sugawara M, Shimojima K, Osawa M, Yamamoto T. MECP2 duplication syndrome in both genders. *Brain Dev* 2013;35:411–9.
- Scott Schwoerer J, Laffin J, Haun J, Raca G, Friez MJ, Giampietro PF. MECP2 duplication: possible cause of severe phenotype in females. *Am J Med Genet A* 2014;164A:1029–34.
- El Chehadeh S, Touraine R, Prieur F, Reardon W, Biemvenu T, Chantot-Bastaraud S, Doco-Fenzy M, Landais E, Philippe C, Marle N, Callier P, Mosca-Boidron AL, Mugneret F, Le Meur N, Goldenberg A, Gerrot AM, Chambon P, Satre V, Coutton C, Jouk PS, Devillard F, Dieterich K, Aftenjar A, Burglen L, Moutard ML, Addor MC, Lebon S, Martinet D, Alessandri JL, Doray B, Miguet M, Devys D, Saugier-Veber P, Drunat S, Aral B, Kremer V, Rondeau S, Tabet AC, Thevenon J, Thauvin-Robinet C, Perretton N,

- Portes DV, Faivre L. Xq28 duplication including MECP2 in six unreported affected females: what can we learn for diagnosis and genetic counselling? *Clin Genet* 2017;91:576–88.
- 18 Collins AL, Levenson JM, Vilaythong AP, Richman R, Armstrong DL, Noebels JL, David Sweatt J, Zoghbi HY. Mild overexpression of MeCP2 causes a progressive neurological disorder in mice. *Hum Mol Genet* 2004;13:2679–89.
- 19 Chao HT, Zoghbi HY, Rosenmund C. MeCP2 controls excitatory synaptic strength by regulating glutamatergic synapse number. *Neuron* 2007;56:58–65.
- 20 Nageshappa S, Carroue C, Trujillo CA, Mesci P, Espuny-Camacho I, Pasciuto E, Vanderhaeghen P, Verfaillie CM, Raitano S, Kumar A, Carvalho CM, Bagni C, Ramocki MB, Araujo BH, Torres LB, Lupski JR, van Esch H, Muotri AR. Altered neuronal network and rescue in a human MECP2 duplication model. *Mol Psychiatry* 2016;21:178–88.
- 21 Sztainberg Y, Chen HM, Swann JW, Hao S, Tang B, Wu Z, Tang J, Wan YW, Liu Z, Rigo F, Zoghbi HY. Reversal of phenotypes in MECP2 duplication mice using genetic rescue or antisense oligonucleotides. *Nature* 2015;528:123–6.
- 22 Bauters M, van Esch H, Friez MJ, Boespflug-Tanguy O, Zenker M, Vianna-Morgante AM, Rosenberg C, Ignatius J, Raynaud M, Hollanders K, Govaerts K, Vandenreijt K, Niel F, Blanc P, Stevenson RE, Frys JP, Marynen P, Schwartz CE, Froyen G. Nonrecurrent MECP2 duplications mediated by genomic architecture-driven DNA breaks and break-induced replication repair. *Genome Res* 2008;18:847–58.
- 23 Lugtenberg D, Kleefstra T, Oudakker AR, Nillesen WM, Yntema HG, Tzschach A, Raynaud M, Rating D, Journal H, Chelly J, Goizet C, Lacombe D, Pedespan JM, Echenne B, Tarverdián G, O'Rourke D, King MD, Green A, van Kogelenberg M, van Esch H, Geetz J, Hamel BC, van Bokhoven H, de Brouwer AP. Structural variation in Xq28: MECP2 duplications in 1% of patients with unexplained XLMR and in 2% of male patients with severe encephalopathy. *Eur J Hum Genet* 2009;17:444–53.
- 24 Ramocki MB, Tavayev YI, Peters SU. The MECP2 duplication syndrome. *Am J Med Genet A* 2010;152A:1079–88.
- 25 Association AP. *Diagnostic and statistical manual of mental disorders: DSM-5*. 5th ed. United States: Washington, D.C: ed American Psychiatric Association, 2013.
- 26 Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res* 2002;30:57e–57.
- 27 Bauer M, Kölsch U, Krüger R, Untenwälder N, Hameister K, Kaiser FM, Vignoli A, Rossi R, Botella MP, Budisteanu M, Rosello M, Orellana C, Tejada MI, Paguc SM, Patat O, Julia S, Touraine R, Gomes T, Wenner K, Xu X, Aftenjar A, Toutain A, Philip N, Jezela-Stanek A, Görtner L, Martínez F, Echenne B, Wahn V, Meisel C, Wiczorek D, El-Chehadeh S, Van Esch H, von Bernuth H. Infectious and immunologic phenotype of MECP2 duplication syndrome. *J Clin Immunol* 2015;35:168–81.
- 28 Friez MJ, Jones JR, Clarkson K, Lubs H, Abuelo D, Bier JA, Pai S, Simensen R, Williams C, Giampietro PF, Schwartz CE, Stevenson RE. Recurrent infections, hypotonia, and mental retardation caused by duplication of MECP2 and adjacent region in Xq28. *Pediatrics* 2006;118:e1687–e1695.
- 29 Ramocki MB, Peters SU, Tavayev YI, Zhang F, Carvalho CM, Schaaf CP, Richman R, Fang P, Giaze DG, Lupski JR, Zoghbi HY. Autism and other neuropsychiatric symptoms are prevalent in individuals with MECP2 duplication syndrome. *Ann Neurol* 2009;66:771–82.
- 30 Fu F, Liu HL, Li R, Han J, Yang X, Min P, Zhen L, Zhang YL, Xie GE, Lei TY, Li Y, Li J, Li DZ, Liao C. Prenatal diagnosis of fetuses with congenital abnormalities and duplication of the MECP2 region. *Gene* 2014;546:222–5.
- 31 El Chehadeh S, Faivre L, Mosca-Boidron AL, Malan V, Amiel J, Nizon M, Touraine R, Prieur F, Pasquier L, Callier P, Lefebvre M, Marie N, Dubourg C, Julia S, Sarret C, Francannet C, Laffargue F, Boespflug-Tanguy O, David A, Isidor B, Le Caignec C, Vigneron J, Leheup B, Lambert L, Philippe C, Cuisset JM, Andrieux J, Plessis G, Toutain A, Goldenberg A, Cormier-Daire V, Rio M, Bonnefont JP, Thevenon J, Echenne B, Journal H, Aftenjar A, Burglen L, Bienvu T, Ador MC, Lebon S, Martinet D, Baumann C, Perrin L, Drunat S, Jouk PS, Devillard F, Coutton C, Lacombe D, Delne MA, Philip N, Moncla A, Badens C, Perret N, Masurel A, Thauvin-Robinet C, Des Portes V, Guibaud L. Large national series of patients with Xq28 duplication involving MECP2: Delineation of brain MRI abnormalities in 30 affected patients. *Am J Med Genet A* 2016;170A:116–29.
- 32 Van Buggenhout GJ, Pijckels E, Holvoet M, Schaap C, Hamel BC, Frys JP. Cri du chat syndrome: changing phenotype in older patients. *Am J Med Genet* 2000;90:203–15.
- 33 Christodoulou J, Ho G. *MECP2-Related Disorders. GeneReviews(R)*. Seattle (WA: University of Washington, Seattle, 1993.
- 34 Vignoli A, Borgatti R, Peron A, Zucca C, Ballarati L, Bonaglia C, Bellini M, Giordano L, Romaniello R, Bedeschi MF, Epifanio R, Russo S, Caselli R, Giardino D, Darra F, La Briola F, Banderali G, Canevini MP. Electroclinical pattern in MECP2 duplication syndrome: eight new reported cases and review of literature. *Epilepsia* 2012;53:1146–55.
- 35 Echenne B, Roubertie A, Lugtenberg D, Kleefstra T, Hamel BC, Van Bokhoven H, Lacombe D, Philippe C, Jonveaux P, de Brouwer AP. Neurologic aspects of MECP2 gene duplication in male patients. *Pediatr Neurol* 2009;41:187–91.
- 36 Peters SU, Hundley RJ, Wilson AK, Carvalho CM, Lupski JR, Ramocki MB. Brief report: regression timing and associated features in MECP2 duplication syndrome. *J Autism Dev Disord* 2013;43:2484–90.
- 37 Temudo L, Oliveira P, Santos M, Dias K, Vieira J, Moreira A, Calado E, Carrilho I, Oliveira G, Levy A, Barbot C, Fonseca M, Cabral A, Dias A, Cabral P, Monteiro J, Borges L, Gomes R, Barbosa C, Mira G, Eusebio F, Santos M, Sequeiros J, Maciel P. Stereotypes in Rett syndrome: analysis of 83 patients with and without detected MECP2 mutations. *Neurology* 2007;68:1183–7.
- 38 Chin Wong L, Hung PL, Jan TY, Lee WT, Taiwan Rett Syndrome Association. Variations of stereotypes in individuals with Rett syndrome: a nationwide cross-sectional study in Taiwan. *Autism Res* 2017;10:1204–14.
- 39 Xi CY, Lu Y, Tan YH, Hua TY, Zhao YJ, Liu XM, Gao H. Analysis of MECP2 gene copy number in boys with autism. *J Child Neurol* 2011;26:570–3.
- 40 Hehr U, Hehr A, Uyanik G, Phelan E, Winkler J, Reardon W. A filamin A splice mutation resulting in a syndrome of facial dysmorphism, periventricular nodular heterotopia, and severe constipation reminiscent of cerebro-fronto-facial syndrome. *J Med Genet* 2006;43:541–4.
- 41 Gargiulo A, Auricchio R, Barone MV, Cotugno G, Reardon W, Milla PJ, Ballabio A, Ciccodicola A, Auricchio A. Filamin A is mutated in X-linked chronic idiopathic intestinal pseudo-obstruction with central nervous system involvement. *Am J Hum Genet* 2007;80:751–8.
- 42 Heanue TA, Pachnis V. Enteric nervous system development and Hirschsprung's disease: advances in genetic and stem cell studies. *Nat Rev Neurosci* 2007;8:466–79.
- 43 Okamoto N, Del Maestro R, Valero R, Monros E, Poo P, Kanemura Y, Yamasaki M, Hydrocephalus YM. Hydrocephalus and Hirschsprung's disease with a mutation of L1CAM. *J Hum Genet* 2004;49:334–7.
- 44 Fernández RM, Núñez-Torres R, González-Meneses A, Antiñolo G, Borrego S. Novel association of severe neonatal encephalopathy and Hirschsprung disease in a male with a duplication at the Xq28 region. *BMC Med Genet* 2010;11:137.
- 45 Yang T, Ramocki MB, Neul JL, Lu W, Roberts L, Knight J, Ward CS, Zoghbi HY, Kheradmand F, Corry DB. Overexpression of methyl-CpG binding protein 2 impairs TH1 responses. *Sci Transl Med* 2012;4:163ra158.
- 46 Cronk JC, Herz J, Kim TS, Louveau A, Moser EK, Sharma AK, Smimov I, Tung KS, Braciale TJ, Kipnis J. Influenza A induces dysfunctional immunity and death in MECP2-overexpressing mice. *JCI Insight* 2017;2:e88257.
- 47 Bartsch O, Gebauer K, Lechno S, van Esch H, Froyen G, Bonin M, Seidel J, Thamm-Mücke B, Horn D, Klopocki E, Hertzberg C, Zechner U, Haaf T. Four unrelated patients with Lubs X-linked mental retardation syndrome and different Xq28 duplications. *Am J Med Genet A* 2010;152A:305–12.
- 48 Smyk M, Obersztyn E, Nowakowska B, Nawara M, Cheung SW, Mazurczak T, Stankiewicz P, Bocian E. Different-sized duplications of Xq28, including MECP2, in three males with mental retardation, absent or delayed speech, and recurrent infections. *Am J Med Genet B Neuropsychiatr Genet* 2008;147B:799–806.

## II. Deuxième partie : la recherche de nouveaux gènes candidats de DILX : les variations du gène *SLITRK2* dans les troubles du neurodéveloppement

Suite à l'identification de variants par séquençage d'exome dans un nouveau gène candidat de DI liée à l'X, le gène *SLITRK2* (qui code pour une protéine transmembranaire contrôlant la croissance des neurites), j'ai lancé plusieurs appels à collaboration en France et sur le plan international dans le but de colliger d'autres observations de patients atteints de DI et porteurs d'un variant candidat au sein de ce gène. Pour cela j'ai réalisé un travail de recueil de données moléculaires, cliniques, et iconographiques afin de comparer les phénotypes des patients et rechercher des signes cliniques et d'imagerie récurrents parmi eux. Dans ce but, j'ai présenté une communication écrite (poster) au congrès annuel de l'European Society of Human Genetics (Juin 2020). J'ai par la suite analysé l'ensemble des variants identifiés chez ces patients et tenté de les prioriser sur la base de leurs caractéristiques moléculaires, des différentes prédictions des logiciels *in silico*, du phénotype morphologique et neurocognitif des patients et du caractère hérité ou *de novo* des variants. S'agissant d'un gène encore non décrit en pathologie humaine avec une très grande majorité de variants faux-sens identifiés dans notre cohorte, j'ai pris contact avec une équipe de recherche coréenne ayant déjà travaillé sur le gène *SLITRK2* et d'autres gènes de la famille *SLITRK* (Dr Ji Won Um, Daegu, Corée du Sud) afin d'initier des études fonctionnelles dans le but de valider certains des variants candidats identifiés. Ce travail a fait l'objet d'une publication (El Chehadeh *et al.*, 2022) (*Supplemental data* en annexe 2).

## Résumé

SLITRK2 est une protéine transmembranaire exprimée dans les neurones postsynaptiques qui régule la croissance des neurites et le maintien des synapses excitatrices. Dans la présente étude, nous rapportons des variants rares (un variant non-sens et six variants faux-sens) dans le gène *SLITRK2*, situé sur le chromosome X, identifiés par ES chez des individus atteints de troubles du neurodéveloppement. Des études fonctionnelles ont montré que certains variants de *SLITRK2* entraînaient une altération du transport membranaire et des capacités d'activation et de maintien des synapses excitatrices. Ces variations abolissaient également la capacité de la protéine SLITRK2 sauvage à réduire les niveaux du récepteur tyrosine kinase TrkB dans les neurones. De plus, les souris knock-out conditionnelles (cKO) pour *Slitrk2* au niveau des neurones hippocampiques présentaient des troubles de la mémoire à long terme et une démarche anormale, récapitulant un sous-ensemble de caractéristiques cliniques des patients présentant des variants de *SLITRK2*. En outre, l'altération du maintien des synapses excitatrices induite par le cKO de *Slitrk2* entraînaient des anomalies de la mémoire de référence spatiale chez les souris. En résumé, ces données suggèrent que des variants de *SLITRK2* sont impliqués dans une forme de troubles du neurodéveloppement liés au chromosome X qui sont causés par la perturbation de diverses facettes de la fonction de la protéine SLITRK2.



# SLITRK2 variants associated with neurodevelopmental disorders impair excitatory synaptic function and cognition in mice

Salima El Chehadeh<sup>1,2,3,33</sup>, Kyung Ah Han<sup>4,33</sup>, Dongwook Kim<sup>4,33</sup>, Gyubin Jang<sup>4,33</sup>, Somayeh Bakhtiari<sup>5,6</sup>, Dongseok Lim<sup>4</sup>, Hee Young Kim<sup>4</sup>, Jinhui Kim<sup>4</sup>, Hyeonho Kim<sup>4</sup>, Julia Wynn<sup>7</sup>, Wendy K. Chung<sup>7,8</sup>, Giuseppina Vitiello<sup>9</sup>, Ioana Cutcutache<sup>10</sup>, Matthew Page<sup>10</sup>, Jozef Gecz<sup>11,12,13</sup>, Kelly Harper<sup>11,12</sup>, Ah-reum Han<sup>14</sup>, Ho Min Kim<sup>14,15</sup>, Marja Wessels<sup>16</sup>, Allan Bayat<sup>17,18</sup>, Alberto Fernández Jaén<sup>19</sup>, Angelo Selicorni<sup>20</sup>, Silvia Maitz<sup>21</sup>, Arjan P. M. de Brouwer<sup>22</sup>, Anneke Vulto-van Silfhout<sup>22,23</sup>, Martin Armstrong<sup>24</sup>, Joseph Symonds<sup>25</sup>, Sébastien Küry<sup>26,27</sup>, Bertrand Isidor<sup>26,27</sup>, Benjamin Cogné<sup>26,27</sup>, Mathilde Nizon<sup>26,27</sup>, Claire Feger<sup>28</sup>, Jean Muller<sup>3,28</sup>, Erin Torti<sup>29</sup>, Dorothy K. Grange<sup>30</sup>, Marjolaine Willems<sup>31</sup>, Michael C. Kruer<sup>5,6</sup>, Jaewon Ko<sup>4,33</sup>, Amélie Piton<sup>2,28,32,33</sup> & Ji Won Um<sup>4,33</sup>

SLITRK2 is a single-pass transmembrane protein expressed at postsynaptic neurons that regulates neurite outgrowth and excitatory synapse maintenance. In the present study, we report on rare variants (one nonsense and six missense variants) in *SLITRK2* on the X chromosome identified by exome sequencing in individuals with neurodevelopmental disorders. Functional studies showed that some variants displayed impaired membrane transport and impaired excitatory synapse-promoting effects. Strikingly, these variations abolished the ability of SLITRK2 wild-type to reduce the levels of the receptor tyrosine kinase TrkB in neurons. Moreover, *Slitrk2* conditional knockout mice exhibited impaired long-term memory and abnormal gait, recapitulating a subset of clinical features of patients with SLITRK2 variants. Furthermore, impaired excitatory synapse maintenance induced by hippocampal CA1-specific cKO of *Slitrk2* caused abnormalities in spatial reference memory. Collectively, these data suggest that SLITRK2 is involved in X-linked neurodevelopmental disorders that are caused by perturbation of diverse facets of SLITRK2 function.

A full list of author affiliations appears at the end of the paper.

Intellectual disability (ID) affects about 3% of the general population and is the most common reason for referral to genetic centers. It is estimated that at least half of cases have a monogenic cause, with more than 1000 genes corresponding to an equivalent number of rare diseases currently identified. The most frequently identified genes are responsible for at most 0.5% of cases. All modes of transmission—autosomal dominant, recessive, and X-linked (dominant or recessive)—are observed. Because of the disproportionate number of males among individuals with ID and the possibility of performing genetic studies of X-linked patterns of inheritance in large families, special attention has been paid to identifying the genetic basis of X-linked ID (XLID). Prior to the advent of next-generation sequencing (NGS), major efforts by the EuroMRX consortium, the Sanger Center, and other international groups encompassing thousands of families<sup>1–4</sup>, led to the identification of a large number of XLID genes. The generalization of NGS approaches, including exome sequencing and whole-genome sequencing has greatly accelerated the discovery of autosomal and XLID genes and increased the diagnostic yield of ID, resulting in a shortened diagnostic odyssey for families and expanded access to genetic counseling, including prenatal and pre-implantation genetic diagnosis. Thus, there are now about ~130 identified XLID genes, with an additional ~50 potential candidate XLID genes<sup>5–7</sup>. These include *SLITRK2*, a gene encoding a transmembrane protein that is highly expressed in the brain.

Slit- and Trk-like family proteins, composed of six members (*Slitrk1–6*), are type I transmembrane proteins that are strongly expressed in the central nervous system<sup>8,9</sup>. Intriguingly, *Slitrs* positively regulate excitatory and inhibitory synapse development in cultured hippocampal neurons in a paralogue-dependent manner<sup>10,11</sup>. Extensive analyses have shown that *Slitrs* interact with distinct members of the LAR-RPTP (leukocyte common antigen-related receptor protein tyrosine phosphatase) family to promote presynaptic assembly<sup>10,12</sup>. In addition, *Slitrk5* interacts with receptor tyrosine kinase *TrkB* in *cis* to regulate their interactions with LAR-RPTPs<sup>13</sup>. Structural analyses have further demonstrated that the N-terminal leucine-rich repeat 1 (LRR1) of *Slitrs* interacts with splicing variants of LAR-RPTPs that contain an insert at the MeB site<sup>14,15</sup>. Divergent intracellular mechanisms of *Slitrk2* and *Slitrk3* also appear to be involved in their synaptic functions<sup>16,17</sup>. Furthermore, systematic analyses to investigate the effects of previously reported *SLITRK2* missense variants revealed that a subset of *SLITRK2* variants associated with neuropsychiatric disorders<sup>18</sup> disrupt the trafficking of *Slitrs* to the cell surface and impair their synapse-promoting function<sup>19</sup>. Collectively, these prior studies have established *Slitrs* as important synapse organizers.

The precise role of the *Slitrk2* protein has not yet been clearly defined. The human *SLITRK2* gene is located on the X chromosome, and two missense variants, p.S549F and p.V89M, were reported to be associated with schizophrenia in two female patients who inherited the missense variants from their asymptomatic heterozygous mothers<sup>18</sup>. Furthermore, *SLITRK2* interacts with the PDZ domain-containing excitatory scaffold *Shank3*<sup>17</sup>, a gene associated with autism spectrum disorders (ASDs)<sup>20</sup>, suggesting its potential involvement in the pathogenesis of neurodevelopmental disorders (NDD).

Here, we report that eight individuals from seven unrelated families with moderate to severe ID and/or NDDs, including neuropsychiatric and behavioral problems, harbor rare potential disease-causing variants of the X-linked gene, *SLITRK2*. We systemically investigated the effects of *SLITRK2* variants on biochemical properties, surface transport, ligand-binding activity, and synaptogenic activities in cultured neurons. In addition, we analyzed the synaptic and behavioral effects of a *Slitrk2* deficiency using *Slitrk2* conditional knockout (cKO) mice.

## Results

### Identification of individuals with rare variants in *SLITRK2*.

Exome sequencing, used as a diagnostic approach in individual P1, identified a nonsense c.1381G>T p.E461\* variant in the *SLITRK2* gene (Table 1 and Fig. S1a). Sanger sequencing confirmed that this variant was present in his affected brother (P2), but absent from mother's blood, suggesting a de novo occurrence in the maternal germ line (Fig. 1 and Fig. S1a). Because the *SLITRK2* gene contains only one coding exon, we suspected that the mutant transcript does not undergo nonsense mRNA-mediated decay. Through an international data-sharing arrangement initiated through Genematcher<sup>21</sup>, we identified six additional individuals with NDDs carrying rare nonsynonymous *SLITRK2* variants predicted to be damaging. Among those were six missense variants—c.221T>C p.L74S (P7), c.628G>A p.E210K (P10), c.934A>G p.T312A (P3), c.1121C>G p.P374R (P8), c.1276C>T p.R426C (P4) and c.1531G>A p.V511M (P5), all found in male individuals (except p.R426C)—that were not previously reported in hemizygous state in gnomAD (Table S1). In all but one of these individuals, no other diagnostic candidate variants were identified during exome sequencing analyses. In P7, a likely pathogenic variant in the homeobox transcription factor *ARX* (c.1109C>T p.A370V) was identified but this did not seem likely to fully explain the patient's phenotype, suggesting that the p.L74S variant identified in *SLITRK2* contributes to the observed symptoms. One variant (p.R426C) occurred de novo, and extended segregation analyses showed that one maternally inherited variant (p.P374R) occurred de novo in the unaffected mother. The remaining variants were maternally inherited. All of these missense variants were predicted to be damaging, based on analyses using SIFT and Polyphen2 (PPH2), with CADD scores between 24.1 and 28.2 (Table S1). The amino acid residues affected by the missense variants are all highly conserved across species and are located in the extracellular part of the *SLITRK2* protein (Fig. 1 and Fig. S2).

Seven additional missense *SLITRK2* variants were identified in seven unrelated male patients with NDDs. Among them: (i) a rare missense variant (p.E555D) was identified in a male (P14) who also carried a pathogenic variant in *CUL4B*, encoding a ubiquitin E3 ligase subunit. This *CUL4B* variant appeared sufficient to explain the majority of his clinical manifestations and the *SLITRK2* p.E555D variant was therefore not considered as potentially disease-causing; (ii) four other missense variants—c.601G>A p.V201I (P13), c.931C>G p.P311A (P11), c.1451G>A p.R484Q (P12), and c.2374C>T p.R792C (P9)—were also reported in the hemizygous state in gnomAD (Table S1 and Fig. S1b). They were not therefore likely to be disease-causing, particularly p.V201I variant, which was detected in 19 hemizygous males from gnomAD; and (iii) two additional missense variants were reported in the Decipher database (<https://www.deciphergenomics.org/>)—c.26G>T, p.S9I and c.44G>A, p.G15E—in individuals with nervous system abnormalities.

### Clinical manifestations in individuals with rare variants in *SLITRK2*.

We collected clinical information for the eight individuals with likely pathogenic variants, P1–5, P7–8, and P10 (Table 1 and Supplementary information). The mean age at examination was 16.9 years (range: 8–30 years). All but one patient had mild to severe intellectual disability and speech delay (mostly severe). Neurological regression was observed in one patient after the age of 6. Patients mainly presented with neuropsychiatric and behavioral issues, including major anxiety (7/8), ASD (4/8), and aggressiveness (2/8). Two patients had attention-deficit hyperactivity disorder (ADHD) with executive impairment, and two were diagnosed with hyperactivity or agitation.

**Table 1 Clinical manifestations identified in individuals harboring potential disease-causing *SLITRK2* variants.**

	P1	P2	P3	P4	P7	P8	P5	P10
SLITRK2 variant	p.E461*	p.E461*	p.T312A	p.R426C	p.L74S (+ARX variant)	p.P374R	p.V511M	p.E210K
Sex, age	M, 30 yrs	M, 28 yrs	M, 21 yrs	F, 13 yrs	M, 12 yrs	M, 11 yrs	M, 12 yrs	M, 8 yrs
OFC (SDS)	>+2 SDS	>+2 SDS	N.A.	<-2 SDS	+1.5 SDS	NA	N.A.	+0.8 SDS
Height (SDS)	0 SDS	0 SDS	<-5 SDS	-2.5 SDS	+1 SDS	+0.5 SDS	Short stature <-2 DS	+0.5 SDS
Feeding difficulties	-	-	+ GT	+GOR	-	-	+ GOR, PEG	+GOR
Kyphoscoliosis	-	-	+	+	+	-	+	-
Developmental delay	+	+	Normal before regression	+	+	+	+	+
Speech	Delay	Delay	Regression	Absent	Delay	Delay	Delay	Language impairment
ID	Mild	Moderate to severe	Severe	Severe	Moderate to severe	Mild	Moderate to severe	Border line
Seizure (age at onset)	-	-	+ (10 yrs)	+ (11 mths)	-	+ (8 yrs)	-	-
Seizures type	-	-	Multifocal, prominent in left central region	Generalized seizures	-	Focal seizures	-	-
Spasticity/Dystonia	-	-	Spasticity	Diplegic cerebral palsy + c	-	-	Dystonic diplegia + g	-
Unsteady gait	-	-	+ a	+ c	+	-	+	-
Neuropsychiatric manifestations	Anxiety	Major anxiety, ASD, mutism, aggressiveness	-	ASD, anxiety, mutism, aggressiveness	ASD, anxiety, hyperactivity	Anxiety, ADHD, very sensitive, easily frustrated	Anxiety, ADHD	ASD, severe anxiety, ADHD, OCD, vocal tics
Other								
Brain MRI anomalies	Normal	Normal	+ b	+ d	Not done	+ f	+ e	Not done

a: Confined to a wheelchair at 21 yrs; b: severe cerebral and cerebellar volume loss, ventricle dilation, atrophy of corpus callosum and brainstem, bilateral hippocampal atrophy with increased FLAIR signal; c: walking with support; d: thin corpus callosum, white matter diffuse reduction, and leukomalacia; e: paucity of white matter, bilateral periventricular lesions, lateral ventricles dilation; f: unspecific minor white substance punctate changes on the right side around the trigonum; g: uses a walker and wheelchair for getting around.

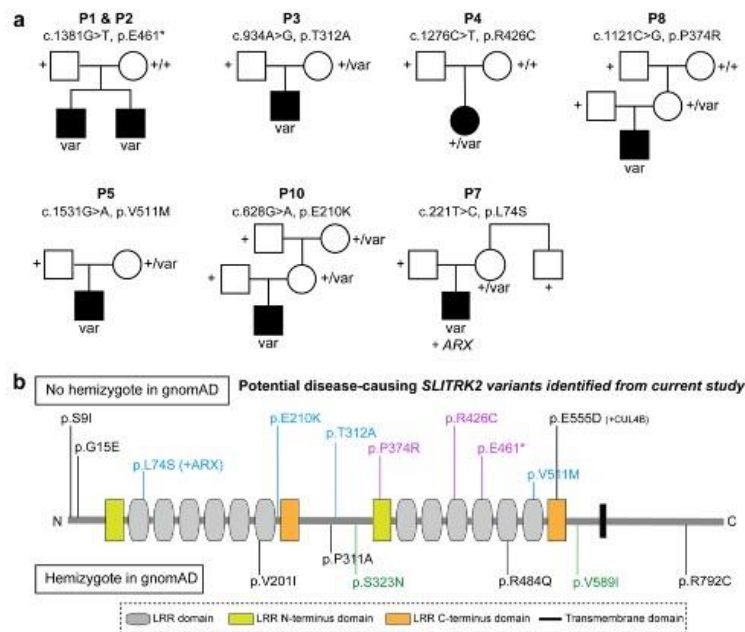
ADHD attention-deficit hyperactivity disorder, ASD autism spectrum disorder, GOR gastroesophageal reflux, GT gastric tube feeding, mths months, MRI magnetic resonance imaging, OCD obsessive-compulsive disorder, OFC occipitofrontal circumference, PEG percutaneous endoscopic gastrostomy, SDS standard deviation score, yrs years.

One patient had obsessive-compulsive behavior, vocal tics, “meltdowns” and tantrums. We also noted an unsteady gait (4/6), spasticity and/or dystonia (3/8), and seizures (3/8). Several patients (4/8) had kyphoscoliosis. Three patients had short stature ( $\leq 2$  SDS). The patients for whom we were able to get photographs were not clearly dysmorphic except for one patient (Supplementary information). Occipitofrontal circumference measurements were highly variable, including two patients with macrocephaly and one patient with microcephaly. Brain magnetic resonance imaging (MRI) of most patients was normal or revealed nonspecific anomalies (4/8), including white matter reduction, ventricular dilation, and thin corpus callosum; three patients had white matter anomalies (Table 1).

**Biochemical, structural, and ligand-binding phenotypes of disease-associated *SLITRK2* variants.** To investigate the structural effect of patient variants on human *SLITRK2*, we predicted the structure of *SLITRK2* using AlphaFold<sup>22,23</sup>. Three variants (L74S, V201I, and E210K) are located in the LRR1 domain (aa C33-D270) and six variants (P374R, R426C, E461\*, R484Q, V511M, and E555D) are located in the LRR2 domain (Fig. 2, top). Free energy calculations indicated that three variants (L74S, P374R, and V511M) are predicted to affect the structural stability of the protein (Table S2), possibly causing misfolding and aberrant trafficking. In particular, a point mutation in the nonpolar leucine, L74, to a polar serine residue in the consensus

LxxLxLxxNxL motif of LRR1-1 is expected to disrupt hydrophobic interactions with neighboring nonpolar residues (V54, L69, L77, L91, and L93). Similarly, mutation of P374 to arginine is expected to disrupt the LRRNT hydrophobic core of LRR2, comprising with V357 and L380. A point mutation, resulting in the replacement of a small, nonpolar valine by a long methionine (V511M), may collapse LRRCT folding of the LRR2 domain. The crystal structure of the human *SLITRK1* LRR1 domain in complex with PTP $\delta$  Ig1-3 has shown that E206 on the concave surface of *SLITRK1* makes ionic interactions with the side chain of K135 and main chain of V136<sup>14</sup>. Mutation of *SLITRK2* E210 corresponding to *SLITRK1* E206 to positively charged lysine would be expected cause charge repulsion, thus interrupting the interaction of *SLITRK2* LRR1 with LAR-RPTs. Because cysteine residues are usually undesirable in exposed regions of a protein unless they acquire functions such as metal binding<sup>24</sup>, the R426C variant on the concave surface may adversely affect protein stability. However, other variants exposed on the surface (E461\*, R484Q, and E555D in LRR2) and V201I, which form hydrophobic networks in LRR1, are predicted to have little effect on the three-dimensional (3D) structure of *SLITRK2*.

To further investigate the impact of amino acid substitutions on the structure and/or synaptic function of *SLITRK2*, we generated mammalian expression vectors encoding *SLITRK2* variants identified in individuals with NDD, both those considered potentially disease-causing and those identified in gnomAD (Table S1 and Fig. 3a). As controls, in addition to the



**Fig. 1** Identification of *SLITRK2* variants in individuals with NDD. **a** Pedigree of families with rare *SLITRK2* variants. **b** Representation of the *SLITRK2* protein, including the different variants identified in individuals with NDD and not previously reported in the hemizygous state in the gnomAD database (Purple, variants occurring de novo in the proband or his mother; blue, variants inherited from unaffected mothers without additional information; green, variants identified in gnomAD in at least two hemizygous males).

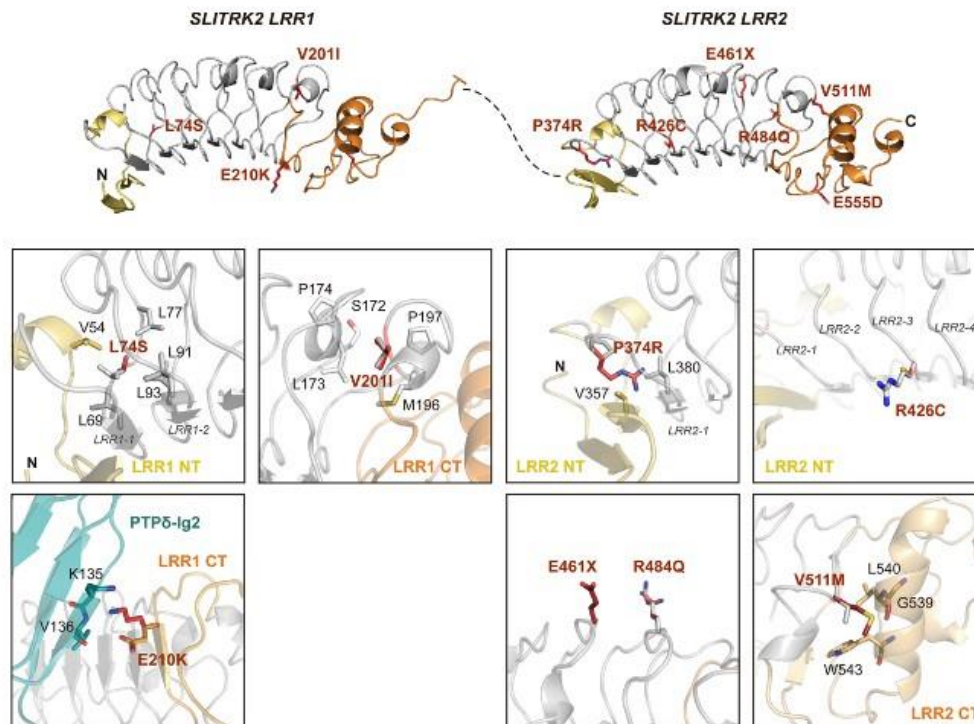
V201I variant, we included two other variants—c.967A>G p.S323N and c.1765G>A p.V589I—present in 6 and 139 males from gnomAD respectively and therefore presumed to be benign (Table S1). Notably, P374 and V511 residues in human *SLITRK2* are conserved in other human *SLITRK* proteins at equivalent positions, but none of the other residues exhibit complete sequence identity across the six *SLITRK* members (Fig. S2a). The two variants from Decipher, S91 and G15E, are located in the signal peptide of *SLITRK2*, which is expected to be cleaved after membrane targeting. Thus, for these two variants, we generated pCMV5-*SLITRK2* vector harboring its own signal peptide, whereas the other variants were generated using the pDisplay-*SLITRK2* vector (Fig. 3a). Most of these residues are quite evolutionarily conserved among various vertebrate species, implying their possible functional significance (Fig. S2b).

We then examined the expression levels and intracellular transport properties of the various *SLITRK2* variants following expression in human embryonic kidney 293 T (HEK293T) cells (Fig. 3). Immunoblot analyses of lysates of HEK293T cells transfected with vectors encoding HA-tagged full-length *SLITRK2* wild-type (WT) or the respective variants showed that total protein expression levels of all *SLITRK2* point mutants (except E461\*) were comparable to those of *SLITRK2* WT (Fig. 3b). The E461\* variant, which resulted in a premature termination codon, yielded a truncated protein of ~60 kDa in SDS-PAGE analyses (Fig. 3b). Notably, unlike *SLITRK2* WT, which is expressed as both N-glycosylated and non-glycosylated protein species, three *SLITRK2* variants (L74S, P374R, and R426C) in HEK293T cells were detected as immature, non-glycosylated forms (Fig. S3). We next examined the surface and

intracellular protein levels of *SLITRK2* WT and the indicated *SLITRK2* variants in HEK293T cells (Fig. 3c, d). In keeping with the glycosylation profiles (Fig. S3), only three *SLITRK2* variants (L74S, P374R, and R426C) displayed significantly reduced surface levels, reflecting their complete entrapment in an intracellular compartment (Fig. 3c, d), in the absence of altered cellular membrane integrity (as monitored by transferrin receptor expression; Fig. S4). The E461\* variant also exhibited reduced surface levels (Fig. 3c, d), in agreement with the absence of a transmembrane region in this variant; thus, this truncated *SLITRK2* is likely secreted (Fig. S5). Results obtained using surface biotinylation assays were consistent with these findings, showing markedly lower surface expression levels of L74S, P374R, R426C, and E461\* variants (Fig. S6). Note that, although S91 and G15E variants are located in the *SLITRK2* signal peptide, they did not affect the membrane transport of *SLITRK2* (Fig. 3c, d).

To determine whether the various *SLITRK2* variants affected the interactions with PTP $\delta$ , an extracellular ligand of *SLITRK2*, we assayed cell-surface binding of recombinant Ig-fusion proteins of PTP $\delta$  (IgC-PTP $\delta$ ) or IgC alone (negative control) with HEK293T cells expressing HA-tagged *SLITRK2* variants (Fig. S7). IgC-PTP $\delta$  proteins avidly bound to HEK293T cells expressing *SLITRK2* variants, except those expressing a subset of *SLITRK2* variants with defective surface transport properties (Fig. S7). IgC did not bind to any tested *SLITRK2* variants (Fig. S7).

To ensure that the subset of the *SLITRK2* variants with impaired surface transport properties exhibited similar trafficking defects in neurons, we transfected cultured hippocampal neurons with WT or variant *SLITRK2* and analyzed them by immunocytochemistry. Whereas most HA-tagged *SLITRK2* variants were transported into

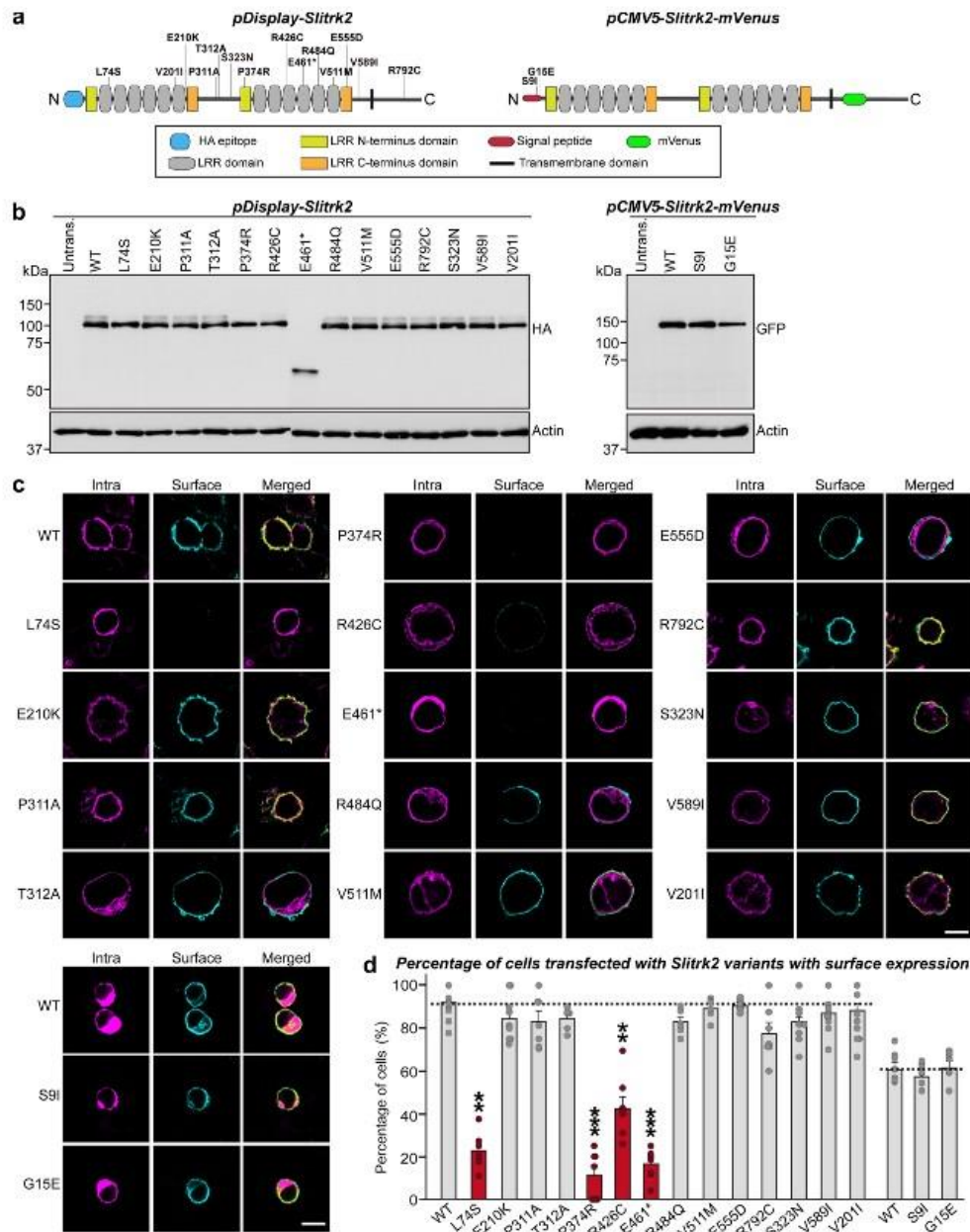


**Fig. 2 Structural model of human SLITRK2 and patient variants.** Model structure of human SLITRK2 LRR1 and LRR2 domains presented as a cartoon. Top: LRR1 domain, C33-D270 (left); LRR2 domain, P341-P579 (right). LRR N-terminus, LRR motifs, and LRR C-terminus are depicted in yellow, gray, and orange, respectively. Black dotted lines represent flexible linkers between LRR1 and LRR2 domains. Bottom: Close-up views of mutated and neighboring residues in LRR1 and LRR2 domains. Residues corresponding to patients' variants described in the current study are presented as red sticks and labeled. The structure of human PTP5 Ig2 (teal) from the crystal structure of the human SLITRK1/human PTP5 Ig1-3 (PDB ID:4RCA) complex was used as a reference for potential interactions between human SLITRK2 and PTP5 Ig2. For clarity, the human SLITRK2 model is rotated 90° around the x-axis in the close-up views of E210K, E461\*, R484Q, and V511M.

dendrites as efficiently as SLITRK2 WT (Fig. 4), three SLITRK2 variants (L74S, P374R, and R426C) were either completely retained in the cell body of transfected neurons, or displayed markedly impaired trafficking into dendrites (Fig. 4), as judged relative to immunofluorescence signals in the somatic compartment of transfected neurons. Puzzlingly, HA immunoreactive intensity in the somatic compartment was not decreased in neurons transfected with L74S, P374R, or R426C (see Fig. 4b), possibly reflecting imprecise evaluation owing to saturation of HA immunoreactivity in transfected neurons. We thus repeated the same experiments using a subset of SLITRK2 variants and quantified HA immunoreactive intensity in transfected neurons imaged using settings that prevented saturation of immunofluorescence signals (Fig. S8a, b). We found that HA immunoreactivity was significantly increased in somatic compartments of neurons expressing L74S or P374R, but not in those of neurons expressing R426C (Fig. S8a, b). Similar findings were also obtained using *Slitrk2*-deficient cultured neurons (Fig. S9). These findings are also consistent with the slight increase in R426C entrapment in an intracellular compartment (i.e., *cis*-golgi) of cultured neurons, as assessed by its colocalization with anti-GM130 (Fig. S10). No alterations in dendritic branching were observed in neurons expressing SLITRK2 WT or the indicated SLITRK2 variants (Fig. S8c). The E461\* variant also exhibited

impaired dendritic trafficking and reduced levels of HA immunofluorescence signals in the somatic compartment (Fig. 4), in agreement with the absence of a transmembrane region; in E461\*, again suggesting that the truncated SLITRK2 protein is likely secreted. These observations suggest that the surface transport deficiency observed in heterologous cells is similarly recapitulated in cultured hippocampal neurons. Taken together, these data suggest that one subset of SLITRK2 variants observed in patients with neurodevelopmental disorders causes improper biochemical processing and abnormal trafficking.

**Synaptic phenotypes of disease-associated SLITRK2 missense variants.** SLITRK2 was previously shown to trigger presynaptic differentiation when expressed in heterologous cells and cocultured in contact with axons<sup>10,12,19</sup>. To test whether the surface transport-defective SLITRK2 variants exhibit differences in presynaptic differentiation-inducing ability, we performed heterologous synapse-formation assays using HEK293T cells expressing SLITRK2 WT or the indicated SLITRK2 variants (Fig. S11). SLITRK2 WT robustly recruited clustering of both the glutamatergic presynaptic marker VGLUT1 (vesicular glutamate transporter 1) and GABAergic presynaptic marker GAD67 (glutamic acid decarboxylase 67 kDa)



around SLITRK2-expressing HEK293T cells, whereas L74S, P374R, R426C, and E461\* were functionally inactive in recruiting both VGLUT1 and GAD67, as expected given their impaired surface expression levels and PTP-binding activities (Fig. S11; see Figs. 3c, d, S6 and S7).

**Generation of *Slitrk2*-cKO mice.** To evaluate the cellular effects of SLITRK2 variants, we generated conditional *Slitrk2*-KO (*Slitrk2*-cKO) mice. Transgenic mice with a deletion of *Slitrk2* were generated by crossing homozygous *Slitrk2*-floxed mice, in which exon 2 and exon 3 are flanked by loxP sites, with a Cre

**Fig. 3 Impaired surface trafficking of a subset of SLITRK2 variants.** **a** Schematic diagrams depicting the *SLITRK2* variants analyzed in the current study. **b** Representative immunoblots from HEK293T cells transfected with the indicated WT or mutant forms of *SLITRK2*. Samples containing equal amounts of protein were resolved by SDS-PAGE and immunoblotted using anti-HA or GFP antibodies;  $\beta$ -actin was used as a loading control. Molecular mass markers are labeled in kilodaltons. The experiments were independently repeated three times. **c** Surface expression analysis of HEK293T cells expressing the indicated WT or mutant forms of *SLITRK2*. Transfected cells were immunostained with mouse anti-HA antibodies (cyan) and detected with FITC-conjugated anti-mouse secondary antibodies under non-permeabilized conditions, followed by permeabilization of cells. Cells were then stained first with rabbit anti-HA antibodies (magenta) and then with Cy3-conjugated anti-rabbit secondary antibodies. Scale bar, 10  $\mu$ m (applies to all images). **d** Quantification of the proportion of cells exhibiting surface expression. All data are shown as means  $\pm$  SEMs ('n' denotes the number of images from three independent experiments; WT, n = 10; L74S, n = 7; E210K, n = 9; P311A, n = 7; T312A, n = 6; P374R, n = 7; R426C, n = 8; E461\*, n = 8; R484Q, n = 6; V511M, n = 6; E555D, n = 6; R792C, n = 7; S323N, n = 13; V589I, n = 13; V201I, n = 16; WT, n = 7; S9I, n = 7; and G15E, n = 7; \*\*p < 0.01, \*\*\*p < 0.001; ANOVA with a non-parametric Kruskal-Wallis test). See Source data for raw data values and Supplementary Table 4 for statistical details.

recombinase driver line under control of the Nestin promoter (*Nestin-Cre*) (Fig. 5a, b). Quantitative RT-PCR analyses showed that mRNA levels of *Slitrk2*, but not those of the other *Slitrks*, were completely eliminated in *Nestin-Slitrk2* mice (Fig. 5c). *Nestin-Slitrk2* male mice weighed marginally (but significantly) less at postnatal day 30 (P30), P40, and P50 compared with control mice, a generalized metabolic phenotype of the *Nestin-Cre* driver line<sup>25</sup> (Fig. 5d), although their brain weights and size were comparable (Fig. 5e, f). Gross morphology, cortical thickness and neuron numbers, including interneuron numbers, were comparable between *Nestin-Slitrk2* mice and control littermates, as assessed by Nissl staining (Fig. 5g) and immunostaining for the neuron-specific marker NeuN (Fig. 5h, i) and GABAergic interneuron marker GAD67 (Fig. 5j, k). In addition, semi-quantitative immunoblot analyses confirmed complete elimination of detectable *Slitrk2* protein in *Nestin-Slitrk2* mice (Fig. 5l, m), as assessed with the *Slitrk2*-specific polyclonal antibody, JK177 (Fig. S12), and further showed that expression levels of various synaptic proteins in the cortex of *Nestin-Slitrk2* mice were comparable to those of littermate control mice (Fig. 5l, m). Moreover, ventricle volumes in *Nestin-Slitrk2* mice were comparable to those of control mice (Fig. S13).

***Slitrk2*-cKO induces an impairment in excitatory synaptic development that is not reversed by reintroduction of a subset of SLITRK2 patient variants.** *Slitrk2* was previously shown to specifically promote excitatory, but not inhibitory, synapse maintenance in cultured neurons and hippocampal CA1 neurons<sup>10,17,19</sup>. To address this, we first examined whether *Slitrk2* deletion induced a specific decrease in excitatory synapse numbers and synaptic transmission in cultured hippocampal neurons. Immunostaining neurons at DIV14 with antibodies to VGLUT1 and SHANK, markers for excitatory postsynaptic density, revealed a reduction in the density of excitatory synaptic puncta immunoreactive to both VGLUT1 and SHANK (Fig. S14a, b). As expected, neither the density nor size of inhibitory synaptic puncta was changed in neurons immunostained with antibodies to VGAT and GABA<sub>A</sub> $\gamma$ 2, an inhibitory postsynaptic marker (Fig. S14a, b, d).

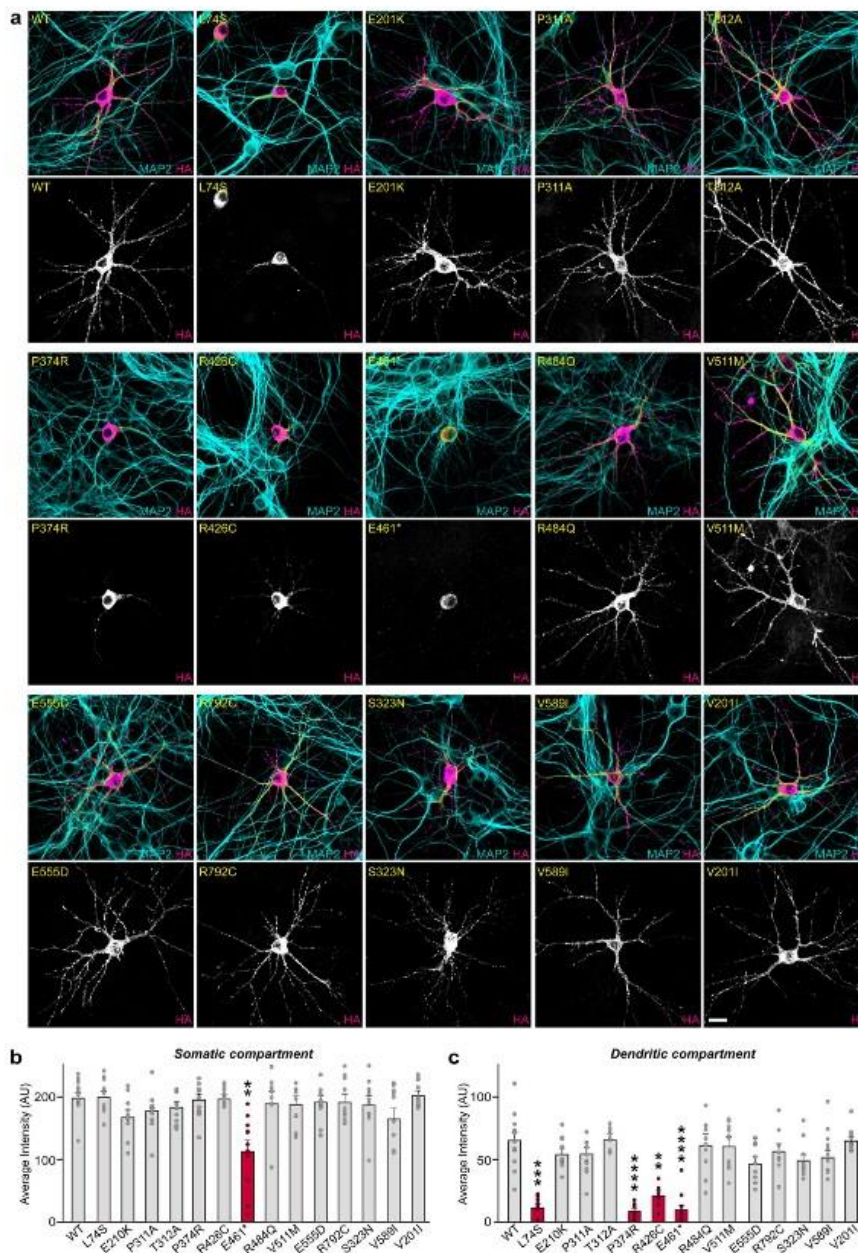
To assess whether the potential disease-causing *SLITRK2* variants affect the excitatory synapse-promoting function of *SLITRK2* WT, we infected cultured hippocampal *Slitrk2*-cKO neurons with lentiviruses expressing either inactive Cre ( $\Delta$ Cre; Control) or active Cre recombinase at DIV5 and co-transfected them with human *SLITRK2* expression variants and EGFP at DIV8–9. We subsequently immunostained these infected/co-transfected neurons with antibodies to synaptophysin, SHANK and EGFP (to visualize the transfected neurons) at DIV13–14. The density, but not the size, of excitatory synaptic puncta immunoreactive to both synaptophysin and SHANK was significantly reduced in *Slitrk2*-deficient neurons (Fig. 6a, b; Fig. S14e). Co-expression of *SLITRK2* WT completely reversed the excitatory

synaptic deficits observed in *Slitrk2*-deficient neurons (Fig. 6a, b). In contrast, co-expression of the *SLITRK2* variants, L74S, P374R, R426C, or E461\*, consistently failed to rescue the deficits in excitatory synapses (Fig. 6a, b), in agreement with the results obtained in heterologous synapse-formation assays (Fig. S11). Strikingly, expression of T312A also failed to rescue the deficits in excitatory synapses (Fig. 6a, b), despite the fact that this *SLITRK2* variant exhibited expression levels, surface transport, and pre-synaptic differentiation-inducing abilities comparable to those of *Slitrk2* WT (see Figs. 3, 4, S4, S6, and S11).

To corroborate these immunocytochemical results, we measured miniature excitatory postsynaptic currents (mEPSCs) in cultured hippocampal neurons using whole-cell electrophysiological recordings (Fig. 6c–e). We found that the frequency, but not amplitude, decay time or rise slope, of mEPSCs was markedly decreased in *Slitrk2*-deficient cultured neurons (Fig. 6c–g; see also Fig. S15). In contrast, *Slitrk2* deletion did not affect the frequency or amplitude of miniature inhibitory postsynaptic currents (mIPSCs) in cultured hippocampal neurons (Fig. S15). Co-expression of *SLITRK2* WT or the indicated *SLITRK2* variants completely reversed decrease in mEPSC frequency (Fig. 6c, e), whereas co-expression of *SLITRK2* variants defective in rescuing anatomical deficits failed to rescue mEPSC frequency (Fig. 6c, e). Consistent with results obtained in cultured hippocampal neurons, immunohistochemical analyses using *Nestin-Slitrk2* mice revealed that the density of VGLUT1 and PSD-95 puncta was significantly decreased in a subset of hippocampal CA1 layers (*stratum oriens* [SO] and *stratum radiatum* [SR]) (Figs. S16). These results collectively suggest that the synaptic phenotypes of the various *SLITRK2* variants likely reflect perturbation of diverse facets of *SLITRK2* functions.

**Reduced TrkB levels induced by *SLITRK2* WT, but not by disease-associated *SLITRK2* variants.** Previous studies have shown that *SLITRK5* acts as a co-receptor for TrkB that orchestrates brain-derived neurotrophic factor (BDNF)-TrkB-dependent trafficking and signaling<sup>13</sup>, and that an obsessive-compulsive disorder (OCD)-associated *SLITRK5* missense variant impairs *SLITRK5* binding to TrkB<sup>26</sup>. Thus, we next analyzed whether the identified *SLITRK2* variants also altered TrkB binding, assessed by coimmunoprecipitation assays in HEK293T cells. We found that disease-linked *SLITRK2* variants with impaired excitatory synapse maintenance did not affect the ability of *SLITRK2* to bind TrkB (Fig. S17a, b).

BDNF plays critical roles in neuronal survival and synaptic plasticity through activation of a full-length receptor (TrkB-FL) while truncated TrkB isoforms (TrkB-T) are presumed to competitively inhibit BDNF-mediated TrkB signaling in certain contexts<sup>27</sup>. Excitotoxic stimulation of cultured neurons with glutamate downregulated TrkB-FL, resulting in downregulation of BDNF-TrkB signaling<sup>28</sup>. In addition, calpain activation triggers the truncation of TrkB-FL, leading to downregulation of BDNF-

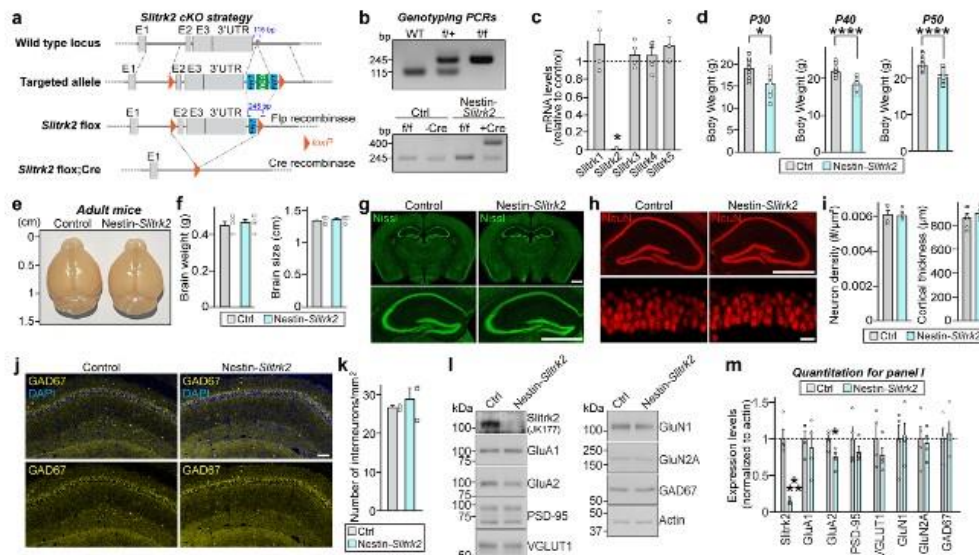


TrkB signaling<sup>29</sup>. These findings suggest that disease-linked *SLITRK2* variants could alter BDNF-TrkB signaling. To test this possibility, we infected DIV3 cortical cultured neurons with lentiviruses expressing *SLITRK2* WT or the indicated variants, and collected the infected neuronal lysates at DIV14. Immunoblot analyses revealed that overexpression of *SLITRK2* WT

significantly decreased TrkB-FL levels, but markedly increased TrkB-T levels (Fig. 7a, b), suggesting that *SLITRK2* might reduce forward trafficking of TrkB and/or trigger the truncation of TrkB-FL. Strikingly, T312A abolished the ability of *SLITRK2* WT to decrease TrkB-FL levels (Fig. 7a, b). In addition, the effects of L74S, P374R, R426C, or E461\* were similar to those of T312A



**Fig. 4 Impaired dendritic targeting of a subset of SLITRK2 variants in cultured hippocampal neurons.** **a** Representative images of hippocampal neurons transfected with the indicated WT or mutant forms of SLITRK2 at DIV10. The transfected neurons at DIV14 were double-immunostained for antibodies against the somatodendritic marker, MAP2 (cyan), and HA (magenta). Scale bar, 20  $\mu$ m (applies to all images). **b, c** Somatic (**b**) or dendritic (**c**) targeting of WT or the indicated mutant forms of SLITRK2 in hippocampal neurons was quantified by measuring average intensity of HA immunofluorescence in primary dendrites. The average intensities of SLITRK2 WT and variants in the soma region of transfected neurons were also quantified. All data are shown as means  $\pm$  SEMs ( $n$  denotes the number of neurons from three independent experiments; WT,  $n = 12$ ; L74S,  $n = 10$ ; E210K,  $n = 10$ ; P311A,  $n = 10$ ; T312A,  $n = 10$ ; P374R,  $n = 11$ ; R426C,  $n = 10$ ; E461\*,  $n = 11$ ; R484Q,  $n = 10$ ; V511M,  $n = 10$ ; E555D,  $n = 11$ ; R792C,  $n = 10$ ; S323N,  $n = 10$ ; V589I,  $n = 10$ ; V201I,  $n = 10$ ; \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ ; ANOVA with a non-parametric Kruskal-Wallis test). See Source data for raw data values and Supplementary Table 4 for statistical details.

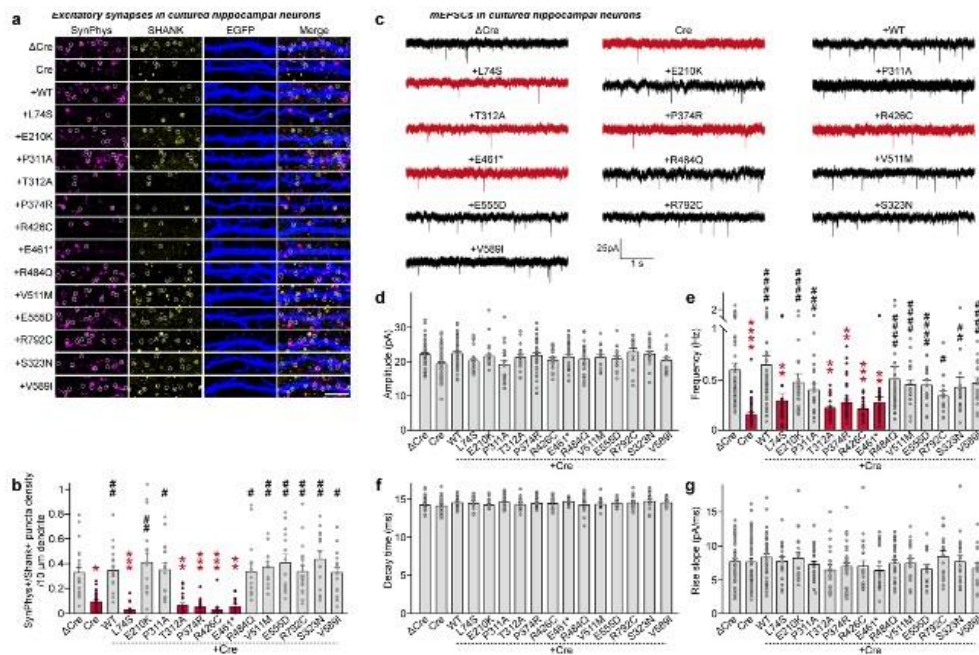


**Fig. 5 Generation of *Slitrk2*-cKO mice.** **a** Strategy used to generate *Slitrk2*-cKO mice. LoxP sites were introduced at positions flanking the neomycin gene, FLP recombinant target (FRT), and exon 2-3 (E2 and E3) of the *Slitrk2* gene. Black arrows indicate forward and reverse primers used for genotyping. Note that lacZ and neomycin cassettes are two separate markers. **b** PCR genotyping of *Slitrk2*-floxed mice (top) and Nestin-Cre;*Slitrk2*-cKO (Nestin-*Slitrk2*) mice (bottom). The band size for the *Slitrk2*-floxed allele was 245 bp. The 400-bp Cre-specific PCR product was detected only in Nestin-*Slitrk2* mice. **c** Quantitative RT-PCR analyses of *Slitrk* mRNA levels in Nestin-*Slitrk2* mice. Data are mean  $\pm$  SEMs ( $n = 4$  mice; \* $p < 0.05$ ; two-tailed Mann-Whitney  $U$  test). **d** Body weights of Nestin-*Slitrk2* and *Slitrk2*-floxed (Ctrl) mice at P30, P40, and P50. Data are means  $\pm$  SEMs ( $n$  denotes the number of mice; P30: Ctrl,  $n = 15$ ; Nestin-*Slitrk2*,  $n = 10$ ; P40: Ctrl,  $n = 20$ ; Nestin-*Slitrk2*,  $n = 14$ ; P50: Ctrl,  $n = 20$ ; Nestin-*Slitrk2*,  $n = 14$ ; \* $p < 0.05$ , \*\*\*\* $p < 0.0001$ ; two-tailed Mann-Whitney  $U$  test). **e, f** Representative photographs of whole brains (**e**) and quantification of brain weights and size (**f**) of P42 Nestin-*Slitrk2* and *Slitrk2*-floxed (Ctrl) mice. Data are means  $\pm$  SEMs ( $n$  denotes the number of mice; Ctrl and Nestin-*Slitrk2*,  $n = 5$ ; two-tailed Mann-Whitney  $U$  test). **g** Normal gross morphology of the Nestin-*Slitrk2* brain at 7 wks, as revealed by Nissl staining. Scale bar: 1 mm (applies to both top and bottom). **h** Representative images of NeuN (a neuronal marker) staining in the Nestin-*Slitrk2* brain. Normal numbers of neurons in hippocampal regions. Top, hippocampus; bottom, CA1 stratum pyramidale layer. Scale bar: 1 mm (top) and 20  $\mu$ m (bottom). **i** Summary graphs for neuron density and cortical thickness. Data are means  $\pm$  SEMs ( $n$  denotes the number of mice; neuron density: Ctrl and Nestin-*Slitrk2*,  $n = 4$ ; cortical thickness: Ctrl and Nestin-*Slitrk2*,  $n = 6$ ; two-tailed Mann-Whitney  $U$  test). **j** Representative images of GAD67 (GABAergic interneuron marker) staining in the hippocampal CA1 of Nestin-*Slitrk2* mice. Scale bar: 0.1 mm. **k** Summary data for GAD67 staining from (**j**). Data are means  $\pm$  SEMs ( $n = 3$  mice each after averaging data from 3 sections/mouse; two-tailed Mann-Whitney  $U$  test). **l, m** Representative immunoblots (**l**) of hippocampal lysates, and summary data showing synaptic protein levels (**m**) in crude synaptosomal fractions of P42 control and Nestin-*Slitrk2* brains, analyzed by semi-quantitative immunoblotting. Data are means  $\pm$  SEMs ( $n = 4$  mice/group; \* $p < 0.05$ , \*\*\* $p < 0.001$ ; two-tailed unpaired  $t$ -test). See Source data for raw data values and Supplementary Table 4 for statistical details.

(Fig. 7a, b). To corroborate these observations, we examined whether a SLITRK2 deletion affects the levels of TrkB-FL and TrkB-T in hippocampal CA1 lysates (Fig. 7c, d). We found that a deletion of *SLITRK2* induced a significant increase in TrkB-FL and decrease in TrkB-T (Fig. 7c, d), confirming the SLITRK2 gain-of-function effects. These results underscore the contribution of a normal

BDNF-TrkB activation range to the regulation of SLITRK2-mediated excitatory synaptic functions.

The accumulation of misfolded proteins is well known to trigger cellular stress responses, such as the endoplasmic reticulum (ER) stress response and unfolded protein response<sup>30–32</sup>. Thus, we hypothesized that expression of a subset of *SLITRK2* variants (i.e.,

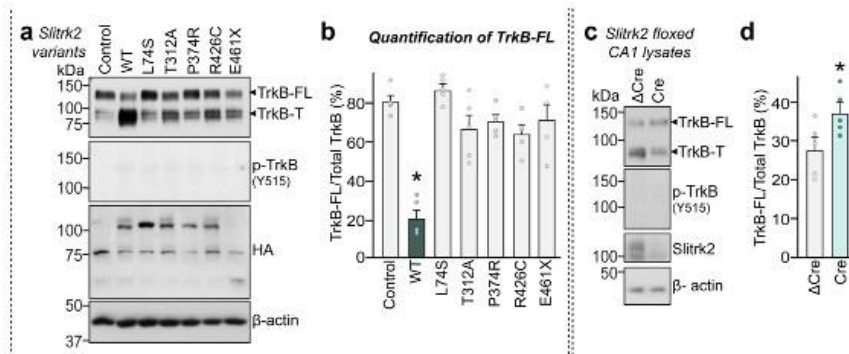


**Fig. 6** Impaired rescue of excitatory synapse development and transmission in *Slitrk2*-cKO hippocampal neurons by a subset of SLITRK2 variants. **a** Triple-immunofluorescence analysis of EGFP (blue), the presynaptic marker synaptophysin (SynPhys; magenta), and the excitatory postsynaptic marker SHANK (yellow) in mature cultured neurons (DIV14) derived from *Slitrk2*<sup>fl/fl</sup> mice infected with lentiviruses expressing  $\Delta$ Cre or Cre at DIV5. For rescue experiments, neurons infected with lentiviruses expressing Cre at DIV5 were transfected with vectors expressing WT or the indicated mutant forms of SLITRK2 together with EGFP vector at DIV8–9. Scale bar: 10  $\mu$ m (applies to all images). **b** Quantification of images in (a), showing the density of SynPhys+SHANK+ puncta. All data are shown as means  $\pm$  SEMs (*n* denotes the number of cells from five independent experiments; Control vs. experimental group:  $\Delta$ Cre, *n* = 22; Cre, *n* = 25; +WT, *n* = 18; +L74S, *n* = 12; +E210K, *n* = 17; +P311A, *n* = 16; +T312A, *n* = 18; +P374R, *n* = 16; +R426C, *n* = 13; +E461\*, *n* = 15; +R484Q, *n* = 16; +V511M, *n* = 15; +E555D, *n* = 14; +R792C, *n* = 21; +S323N, *n* = 14; +V589I, *n* = 17; \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001; Cre vs. experimental group: #*p* < 0.05, ##*p* < 0.01; ANOVA with a non-parametric Kruskal–Wallis test). **c–g** Representative mEPSC traces (**c**) and quantification of amplitudes (**d**), frequencies (**e**), decay time (**f**) and rise time (**g**) of mEPSCs recorded from hippocampal cultured neurons infected with lentiviruses expressing  $\Delta$ Cre or Cre at DIV5 and transfected with vectors expressing WT or the indicated mutant forms of SLITRK2 together with EGFP vector. All data are shown as means  $\pm$  SEMs (*n* denotes the number of cells from six independent experiments; Control vs. experimental group:  $\Delta$ Cre, *n* = 48; Cre, *n* = 46; +WT, *n* = 41; +L74S, *n* = 20; +E210K, *n* = 22; +P311A, *n* = 24; +T312A, *n* = 19; +P374R, *n* = 36; +R426C, *n* = 20; +E461\*, *n* = 22; +R484Q, *n* = 29; +V511M, *n* = 19; +E555D, *n* = 16; +R792C, *n* = 16; +S323N, *n* = 20; +V589I, *n* = 17; \**p* < 0.01, \*\**p* < 0.001, \*\*\**p* < 0.0001; experimental group; Cre vs. experimental group; #*p* < 0.05, ##*p* < 0.01, ###*p* < 0.001, ####*p* < 0.0001; ANOVA with a non-parametric Kruskal–Wallis test). See Source data for raw data values and Supplementary Table 4 for statistical details.

L74S, T312A, P374R, and R426C) exhibiting entrapment in intracellular compartments could induce an ER stress response in cultured hippocampal neurons. However, immunoblot analyses using lysates of neurons overexpressing SLITRK2 WT or the indicated SLITRK2 variants showed no difference in the levels of BiP, which are activated transcriptionally by ER stressors, between SLITRK2 WT and mutant variants (Fig. S17c, d), suggesting no effects of these variants on cellular stress responses.

***Slitrk2*-cKO mice exhibit memory deficits and abnormal gait.** We then tested whether *Slitrk2*-cKO mice exhibited the behavioral features that manifest in patients with NDD and SLITRK2 variants (see Table 1). To this end, we performed a battery of behavioral tests, including those measuring learning and memory, locomotion, anxiety, and motor coordination (Fig. 8 and Fig. S18). Ten-week-old male Nestin-*Slitrk2* mice displayed comparable exploratory behavior (assessed by open-field tests) (Fig. S18a, b),

sociability, and social novelty recognition memory (assessed by three-chamber tests) as littermate control mice, showing preference for the first novel mouse over an inanimate object (Fig. S18c, d) and preference for a novel stranger mouse over familiar mouse (Fig. S18e, f). However, Nestin-*Slitrk2* mice exhibited reduced anxiety-like behavior, as evidenced by increased time spent in open arms (with a similar number of entries into each open arm) in the elevated plus-maze test, without a change in time spent in the light compartment in the light-dark test (Fig. S18g–i). Moreover, Nestin-*Slitrk2* mice showed impaired spatial learning and memory (assessed by novel object-recognition tests), although spontaneous alternation in Y-maze tests was similar to that of littermate control mice (Fig. S18j–m). We further examined spatial reference memory, as evaluated by the Barnes maze test<sup>33</sup>. Both Control and Nestin-*Slitrk2* mice learned to locate the target hole during the course of the training period, as indicated by gradual reductions in the number of search errors, escape latency and total



**Fig. 7** Alterations of full-length TrkB protein levels in cultured neurons by a subset of SLITRK2 variants. Cultured cortical neurons were infected with lentiviruses expressing WT or the indicated mutant forms of SLITRK2. Levels of TrkB were examined by semi-quantitative immunoblotting using antibodies against TrkB or phospho-TrkB. FL full-length, T truncated. **b** Quantification of protein levels from (a). Summary data showing plots from five independent experiments. Data are mean  $\pm$  SEMs ( $^*p < 0.05$ ; ANOVA with a non-parametric Kruskal-Wallis test). **c** Hippocampal CA1 lysates from *Slitrk2*-floxed mice infected with AAVs expressing  $\Delta$ Cre or Cre were examined by semi-quantitative immunoblotting using antibodies against TrkB or phospho-TrkB. FL full-length, T truncated. **d** Quantification of protein levels from (c). Data are means  $\pm$  SEMs ( $n = 5$  mouse samples;  $^*p < 0.05$ ; two-tailed unpaired *t*-test). See Source data for raw data values and Supplementary Table 4 for statistical details.

distance traveled, indicating normal acquisition of spatial reference memory in Nestin-*Slitrk2* mice (Fig. 8a–d). We then conducted the first probe test 3 days after the last day of training and the second probe test after 17 days. In the first probe test, the number of search errors was comparable between Control and Nestin-*Slitrk2* mice (Fig. 8e). In contrast, the accuracy of spatial memory of Nestin-*Slitrk2* mice in the second probe tests was worse than that of control mice, suggesting that retention of spatial reference memory is impaired in Nestin-*Slitrk2* mice (Fig. 8f).

Because some NDD patients reported in the current study suffer from spasticity and/or dystonia and unsteady gait (Table 1), we also examined whether Nestin-*Slitrk2* mice exhibit deficits in fine motor coordination. To this end, we analyzed the juvenile (P20) and adult (P65) Nestin-*Slitrk2* mice by measuring overlap distance, stance width, step angle, and gait symmetry. At both ages, the mean stride length, stance length, and sway length of forelimbs and hindlimbs of Nestin-*Slitrk2* mice were comparable to those of control littermate mice (Fig. 8g–i). Strikingly, Nestin-*Slitrk2* mice exhibited increased front/hind paw overlap, despite normal limb strength (assessed by grip strength test) (Figs. 8g–i, S18n). Taken together, these extensive behavioral data suggest that Nestin-*Slitrk2* mice exhibit impaired learning and memory and abnormal footprint patterns, partly recapitulating the clinical phenotypes of patients with SLITRK2 variants (see Table 1).

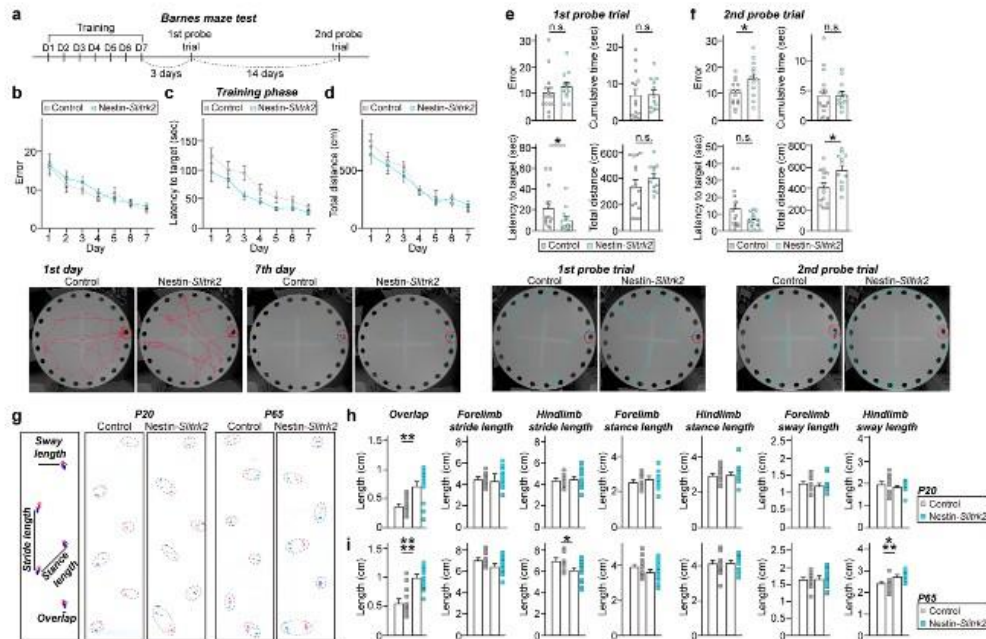
**Correlation of impaired hippocampal CA1 excitatory synapse properties of *Slitrk2*-cKO mice with deficits in spatial reference memory.** Because Nestin-*Slitrk2* mice exhibited impairment in the Barnes maze test (Fig. 8a–f), a hippocampal CA1-dependent spatial task<sup>34</sup>, we asked whether loss of Slitrk2 in the hippocampal CA1 region could recapitulate this behavioral deficit. To this end, we stereotactically injected adeno-associated viruses (AAVs) expressing Cre or  $\Delta$ Cre (Ctrl) into the CA1 of adult *Slitrk2*-floxed mice and performed semi-quantitative immunohistochemistry and electrophysiological recordings three weeks after injections (Fig. 9a, b). We found that the frequency, but not amplitude, of mEPSCs was markedly decreased in hippocampal CA1-specific *Slitrk2*-cKO mice (Fig. 9c–e). Moreover, these mice exhibited a significant reduction in the integrated intensity of VGLUT1 and PSD-95 puncta in the SO and SR layers (Fig. 9f, g), without

changes in VGAT puncta in any examined hippocampal CA1 layer (Fig. S19), in keeping with the anatomical phenotype of Nestin-*Slitrk2* mice (Fig. S16). Notably, co-expression of SLITRK2 WT, but not that of the SLITRK2 T312A variant, completely rescued the decreased integrated intensity of VGLUT1 and PSD-95 puncta in hippocampal CA1-specific *Slitrk2*-cKO mice (Fig. 9f, g). These results reinforce the phenotype of SLITRK2 T312A in hippocampal CA1 pyramidal neurons.

Finally, to determine whether CA1-specific elimination of *Slitrk2* expression leads to long-term spatial memory deficits, we performed Barnes maze tests using CA1-specific *Slitrk2*-cKO mice. Similar to the behavioral phenotype of Nestin-*Slitrk2* mice, CA1 region-specific *Slitrk2*-cKO mice also exhibited impaired retention of spatial reference memory (Fig. 9h–l). Collectively, these results support the conclusion that SLITRK2 synaptic functions in the hippocampal CA1 are essential for execution of normal spatial tasks in mice.

## Discussion

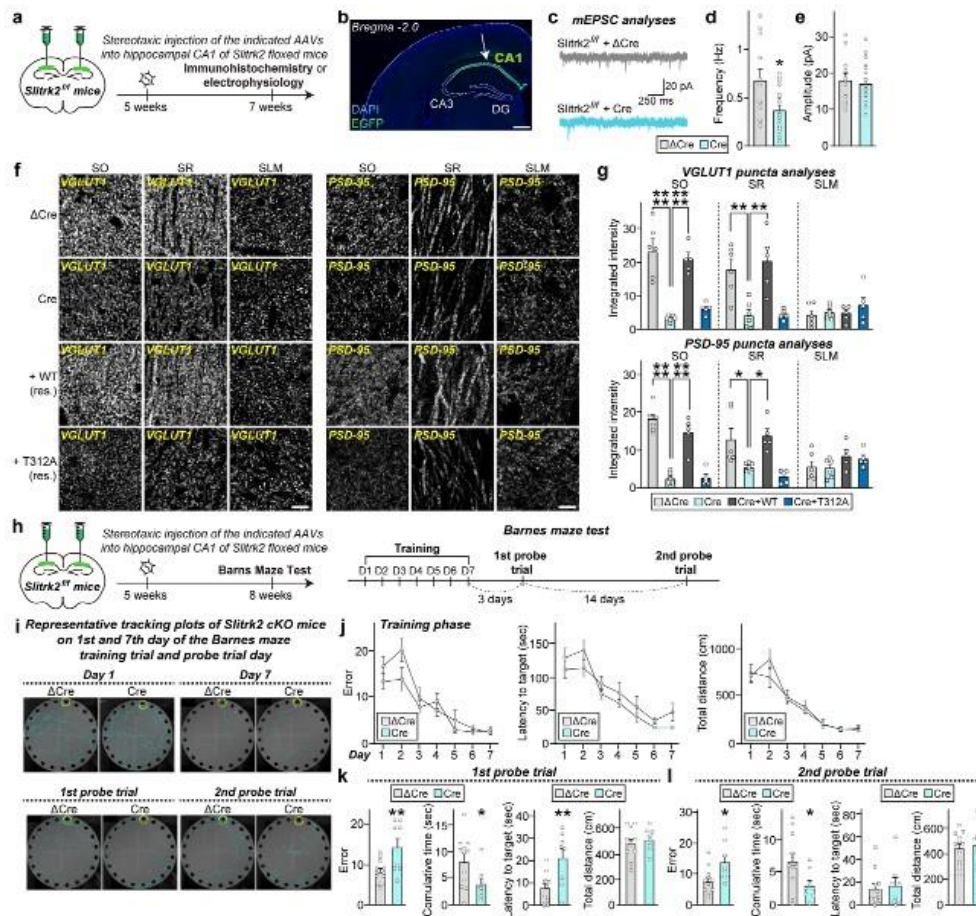
Although numerous genes encoding synaptic cell adhesion proteins (CAMs) have been identified as causative factors for a variety of neurodevelopmental and/or neuropsychiatric disorders, it had remained to be determined whether a subset of neurological symptoms of patients with these disorders are caused by alterations in synaptic adhesion signaling pathways that are not clearly established. As a member of a vertebrate-specific synaptic cell adhesion protein family, SLITRK2 specifically functions in mediating proper excitatory synapse development<sup>35</sup>. In particular, SLITRK2 is required for organizing *trans*-synaptic signaling pathways involved in excitatory presynaptic assembly and post-synaptic maintenance by binding to PTP $\sigma$  and PDZ domain-containing scaffolds<sup>10,17,36</sup>. In addition, our previous study demonstrated that a subset of disease-associated SLITRK2 missense variants induce impaired dendritic targeting of Slitrk2 in neurons, leading to dysfunctions of excitatory synapse maintenance<sup>19</sup>. The SLITRK2 V89M variant reduced excitatory synapse density, likely through its deleterious effects on surface expression and dendritic targeting<sup>19</sup>. Similarly, a series of other SLITRK variants associated with Tourette's syndrome, trichotillomania, schizophrenia, or OCD was shown to trigger loss of



**Fig. 8** *Slitrk2*-cKO mice display mild learning and memory deficits and abnormal gait. **a** Schematic depiction (top) of Barnes maze test. **b–d** Number of errors before first encountering the escape hole (**b**) escape latency (**c**) and total distance (**d**) for Control and Nestin-*Slitrk2* mice during the training session. Data are presented as means  $\pm$  SEMs (*n* denotes the number of mice; Control, *n* = 17; Nestin-*Slitrk2*, *n* = 12; two-way ANOVA with Sidak's multiple comparisons test). Representative track images (bottom) from first and last days of the training. **e, f** Number of errors before first encountering the escape hole for control and Nestin-*Slitrk2* mice during 1st (**e**) and 2nd probe (**f**) trials. Data are presented as means  $\pm$  SEMs (*n* denotes the number of mice; Control, *n* = 15; Nestin-*Slitrk2*, *n* = 12; \**p* < 0.05; two-tailed Mann-Whitney U test). Representative track images (bottom) from 1st and 2nd probe trials. **g** Representative image of the footprint patterns of juvenile (P20) and adult (P65) control and Nestin-*Slitrk2* mice. **h, i** Footprint patterns were analyzed based on overlap, stride length, stance length, and sway length. P20 (**h**) and P65 (**i**) Nestin-*Slitrk2* mice exhibited longer overlap length. Data are shown as means  $\pm$  SEMs (*n* denotes the number of mice; P20: Control, *n* = 10; Nestin-*Slitrk2*, *n* = 12; P65: Control, *n* = 14; Nestin-*Slitrk2*, *n* = 15; \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001, \*\*\*\**p* < 0.0001; two-tailed Mann-Whitney U test). See Source data for raw data values and Supplementary Table 4 for statistical details.

the respective SLITRK protein functions in cultured hippocampal neurons<sup>19,26</sup>. These results collectively suggest that pathogenic *SLITRK* genetic variants likely induce loss-of-function of corresponding SLITRKs, accounting for the pathogenesis of patients with neurological disorders. However, previous studies, which primarily employed knockdown (KD) and/or overexpression approaches, preclude firm conclusions about SLITRK2 functions. Previous *Slitrk2* loss-of-function studies performed in cultured neurons and mice produced contradictory results<sup>10,17</sup>, with *Slitrk2* KD in cultured neurons reducing excitatory synapse density, and *Slitrk2* KD in hippocampal CA1 neurons failing to affect excitatory synapse density<sup>10,17</sup>. These results were previously attributed to differences in genetic manipulations, cellular preparation, or other unknown factor(s). Extensive analyses using *Slitrk2*-cKO mouse-derived cultured neurons recapitulated our previous anatomical results using *Slitrk2* KD<sup>10</sup>, in addition to exhibiting marked reduction in excitatory synaptic transmission. Moreover, *Slitrk2* deletion in adult mouse hippocampal CA1 neurons similarly reduces the frequency of mEPSCs, in contrast to the absence of an effect of global *Slitrk2* KD on mEPSCs<sup>17</sup>. These results again underscore the importance of confirmatory analyses using a sophisticated system and approaches to elucidate functions of synaptic proteins.

In the current study, we identified *SLITRK2* variants in NDD characterized by developmental delay and speech delay, mild to severe ID, and various neurological and behavioral comorbidities. Seven of the eight individuals with potential disease-causing variants had significant anxiety, and four patients were diagnosed with autism spectrum disorder. Other behavioral diagnoses include ADHD with executive difficulties, hyperactivity, or agitation. Several individuals had unsteady gait associated with spasticity and/or dystonia and three patients had seizures, of variable type, with the severity of ID ranging from mild to severe. Such clinical heterogeneity among individuals is often reported in NDDs associated with mutations in the same gene and even precisely the same mutation<sup>37,38</sup>, suggesting that other genetic and/or environmental modifying factors might also be involved. *SLITRK2* variants have previously been reported in two females with schizophrenia<sup>18</sup>, but no patients affected by ID and/or NDD phenotypes and carrying *SLITRK2* pathogenic variants have been reported to date. Of the seven potential disease-causing variants identified in this study, five induced prominent alterations in the properties of SLITRK2. Among these, three variants (L74S, P374R, and R426C) perturbed surface transport of SLITRK2, accompanied by impaired binding to LAR-RPTPs and synaptogenic activity, similar to most of known *SLITRK1/4/5* substitutions that



**Fig. 9** Proper *Slitrk2* function in the hippocampal CA1 is required for mediating spatial reference memory in mice. **a** Experimental configuration for electrophysiology and immunohistochemistry experiments (**c–g**). **b** Representative brain sections illustrating the precise targeting of AAVs for expression of Cre recombinase in the hippocampal CA1. Scale bar: 500  $\mu$ m. **c–e** Representative traces (**c**) and summary graphs showing the frequency (**d**) and amplitude (**e**) of mEPSCs recorded from *Slitrk2*-floxed mice injected with AAVs expressing  $\Delta$ Cre or Cre. Data are shown as means  $\pm$  SEMs ( $\Delta$ Cre,  $n = 13$ ; Cre,  $n = 20$ ; \* $p < 0.05$ ; two-tailed Mann-Whitney  $U$  test). **f, g** Representative images (**f**) and summary graphs quantifying the integrated intensity (**g**) of VGLUT1 and PSD-95 puncta. Data are shown as means  $\pm$  SEMs ( $\Delta$ Cre,  $n = 6$ ; Cre,  $n = 6$ ; Cre+WT,  $n = 5$ ; and Cre+T312A,  $n = 5$  mice; \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.0001$ ; ANOVA with Tukey’s test). Scale bar, 10  $\mu$ m (applies to all images). **h** Experimental configuration for Barnes Maze tests (**i–l**). **i** Representative traces of locomotive behavior for Barnes Maze tests. Target box is indicated with a yellow circle. **j** Number of errors before first encountering the escape hole, escape latency, and total distance for *Slitrk2*-floxed mice injected with  $\Delta$ Cre (Control) or Cre during the training session. Data are presented as means  $\pm$  SEMs ( $\Delta$ Cre,  $n = 12$ ; Cre,  $n = 8$ ; two-way ANOVA with Sidak’s multiple comparisons test). **k, l** Number of errors before first encountering the escape hole for *Slitrk2*-floxed mice injected with  $\Delta$ Cre (Control) or Cre during 1st (**k**) and 2nd (**l**) probe trials. Data are presented as means  $\pm$  SEMs ( $\Delta$ Cre,  $n = 12$ ; Cre,  $n = 8$ ; \* $p < 0.05$ , \*\* $p < 0.01$ ; two-tailed unpaired  $t$ -test). See Source data for raw data values and Supplementary Table 4 for statistical details.

represent loss-of-function mutations. One variant (E461\*) in *SLITRK2* was a nonsense variant, and resulted therefore in a truncated *SLITRK2* protein. The fifth variant, T312A, induced normal surface expression, and normal extracellular ligand-binding and synaptogenic activity, albeit with markedly reduced excitatory synapse number and excitatory synaptic transmission. Remarkably, these five *SLITRK2* variants abrogated the ability of

*SLITRK2* WT to negatively regulate TrkB-FL levels in neurons, revealing the significance of interactions between *SLITRK2* and TrkB in terms of preventing *SLITRK2* dysfunction-induced pathogenic effects. Given the well-established centrality of TrkB-mediated signaling in excitatory synapse development<sup>39</sup>, it is plausible that selection of specific binding partners (i.e., LAR-RPTPs vs. TrkB) through fine regulation of *SLITRK2*-mediated

extracellular interactions constitutes an important signaling pathway for excitatory synapse organization. Our preliminary analyses suggested that *Slitrk2*, *Slitrk3*, and *Slitrk5*, but not other *Slitrks*, bind to *TrkB* (Fig. S20). Further experiments are warranted to dissect the molecular determinants underlying *Slitrk2/5-TrkB* interactions.

Our extensive behavioral analyses showed that *Slitrk2* loss-of-functions in mice recapitulated a subset of behavioral symptoms found in patients with *SLITRK2* variants. Strikingly, the current study demonstrated that selective elimination of *Slitrk2* expression in the hippocampal CA1 results in impaired retention of spatial reference memory. It is likely that specific alterations in synaptic inputs on CA1 neurons, possibly from either entorhinal cortex afferents or CA3 Schaffer-collateral afferents, are responsible for abnormalities in this behavioral task. Whether *SLITRK2* is involved in the assembly of specific hippocampal circuits remains to be determined. In addition, *Nestin-Slitrk2* mice exhibited impaired long-term memory and altered footprint patterns. However, these mice displayed normal grip strength and somewhat enhanced motor coordination and learning. In general, the cerebellum is critical for coordinating movements, and its dysfunction often leads to uncoordinated walking or gait disturbance<sup>40</sup>, a feature of patients harboring a subset of *SLITRK2* variants, even though they were not clearly ataxic and the brain MRI did not show cerebellar anomalies (except for one patient with severe cerebral and cerebellar volume loss). Although the current study primarily correlated *SLITRK2* synaptic functions in the hippocampal CA1 with spatial memory performance, future studies are warranted to identify the neuronal populations and/or neural circuits embedded in other brain areas that might underpin abnormal footprint patterns seen in *Nestin-Slitrk2* mice. Moreover, most patients harboring potential disease-causing variants in *SLITRK2* are anxious, whereas *Nestin-Slitrk2* mice appeared to exhibit reduced anxiety-like behavior in the elevated plus-maze tests. Future studies utilizing knock-in mice expressing disease-associated *SLITRK2* variants should help resolve this issue using additional behavioral paradigms that examine anxiety-like behavior. Most patients exhibit seizures and speech delay. Thus, it will be necessary to clarify whether *Nestin-Slitrk2* mice exhibit seizure-like phenotypes or communication deficits by measuring electroencephalography and measuring ultrasonic vocalization, respectively. Because our various analyses did not reveal any prominent phenotypes in the case of E210K and V511M (predicted to be deleterious by *in silico* programs), their pathogenicity remains uncertain, although it is still possible that these variants are involved in *SLITRK2*-associated pathophysiology via as-yet unidentified mechanism(s) encompassing early brain developmental processes. Given that *Slitrk2* is also expressed in specific astrocytic subtypes in adult mice<sup>41</sup>, it is reasonable to speculate that these *SLITRK2* missense variants with no clear phenotypes in current functional assays might affect specific aspects of astrocytic functions that remain to be determined.

In conclusion, we demonstrated nonsense and missense variants in *SLITRK2* in a novel form of X-linked NDD characterized by different degrees of ID associated with motor, speech, and behavioral impairments, that are caused by perturbation of *SLITRK2* function. Further studies with larger cohorts are needed to confirm the involvement of this gene in a NDD and better delineate the clinical manifestations of the latter, especially at the behavioral level.

## Methods

**Participants and genetic analysis.** Individuals were referred by clinical geneticists for genetic testing as part of routine clinical care. All patients enrolled and/or their legal representative signed informed consent for research use and authorization for

publication. All the institutions received local institutional review board (IRB) approval to use these data in research. The main IRB approval was obtained from the Ethics Committee of the Strasbourg University Hospital (CCPPRB). Individuals P1 and P2 (male siblings) were followed up in the Department of Clinical Genetics of Montpellier University Hospital, and exome sequencing was performed for P1 in the Molecular Genetics Laboratory of Strasbourg University Hospital, France. The six remaining patients (P3–P10) were recruited through an international collaborative study and data-sharing through Genematcher<sup>21</sup>. All patients were followed by a referring clinical geneticist and underwent exome sequencing in the different molecular laboratories as detailed in Supplementary information. Sanger sequencing confirmed the *SLITRK2* variants in all affected probands and was used for segregation studies. Variants have been submitted to ClinVar database. The referring physicians completed clinical forms and sent morphological pictures and brain MRI reports, when available. Clinical information is detailed in Table 1. All patients' families provided written informed consent, and all procedures performed in the studies were done in accordance with the ethical standards of the local institutional research committee and the Declaration of Helsinki. Effects of *SLITRK2* missense variants were predicted using PolyPhen-2<sup>42</sup>, SIFT<sup>43</sup>, and CADD (combined annotation-dependent depletion) scores<sup>44</sup>.

**Expression vectors.** pDisplay-*SLITRK2* missense variants (L74S, V201I, E210K, P311A, T312A, S323N, P374R, R426C, R484Q, V511M, E555D, V589I, and R792C) were generated by polymerase chain reaction (PCR)-based mutagenesis with PrimeSTAR HS DNA polymerase (Takara) using a pDisplay-*SLITRK2* wild-type (WT) construct as a template. pDisplay-*SLITRK2* E461\* was generated by PCR amplification of the indicated region of human *SLITRK2* (amino acid [aa] 22–460), followed by digestion with *Xma*I and *Sac*II and subcloning into a pDisplay vector (Invitrogen). pCMV5-*SLITRK2* missense variants (S9I and G15E) were generated by PCR-based mutagenesis with PrimeSTAR HS DNA polymerase (Takara) using a pCMV5-WT *SLITRK2*-mVenus construct as a template. pCMV6-XL4-*SLITRK2* variant-Fc constructs (L74S, P374R, and R426C) were generated by PCR-based mutagenesis with PrimeSTAR HS DNA polymerase (Takara) using the pCMV6-XL4-*SLITRK2*-Fc construct as a template. pCMV6-XL4-*SLITRK2* E461\*-Fc was generated by PCR amplification of the indicated region of human *SLITRK2* (aa 22–460) using the pCMV6-XL4-*SLITRK2*-Fc construct as a template, followed by digestion with *Nof*I and *Xba*I and subcloning into a modified pCMV6-XL4 vector in frame as a C-terminal Fc fusion protein. L-313 *SLITRK2* missense variants were constructed by amplifying the indicated full-length *SLITRK2* variants by PCR, digesting with *Nhe*I and *Bsr*GI, and subcloning the resulting products into the L-313 vector (lentiviral expression vector). pAAV-*SLITRK2* T2A-GFP variants (WT and T312A) were generated by PCR amplification of the indicated region of human *SLITRK2*, followed by digestion with *Xba*I and *Bam*HI and subcloning into the pAAV-T2A-GFP vector. pDisplay-*SLITRK3* and pDisplay-*SLITRK5* were generated by PCR amplification of mouse *Slitrk3* (aa 28–652) and human *SLITRK5* (aa 41–664), respectively, followed by digestion with *Xma*I and *Sac*II and subcloning into the pDisplay vector. pCAGG-His-HA-Nlg2 was generated by PCR amplification of rat *Nlg2* (aa 1–678) using pCMV5-Nlg2 as a template, followed by digestion with *Eco*RI and *Nof*I and subcloning into the pCAGG-His-HA vector. pCMV-SPORT6-TrkB was purchased from Korea Human Gene Bank. The following constructs were previously described: pDisplay-*SLITRK2* WT<sup>19</sup>, pCMV5-*SLITRK2*-mVenus<sup>10</sup>, pDisplay-*SLITRK1*<sup>10</sup>, pDisplay-*SLITRK4*<sup>10</sup>, pCMV-IgC PTP<sup>19</sup>, and pAAV-T2A-GFP<sup>15</sup>; FSW-ΔCre and FSW-Cre were gifts from Dr. Pascal S. Kaeser (Harvard University, Cambridge, MA, USA). pCMV6-XL4-*SLITRK2*-Fc was a gift from Dr. Davide Comolletti (Victoria University, New Zealand)<sup>46</sup>.

**Antibodies.** A polyclonal rabbit antibody against *SLITRK2* was generated using a fusion protein of glutathione S-transferase (GST) with human *SLITRK2* (aa 705–784) produced in BL21 *Escherichia coli* and purified on a glutathione-Sepharose column (GE Healthcare). Following immunization of rabbits with GST-*SLITRK2* [705–784], the *SLITRK2*-specific antibody, designated JK177 (RRID: AB\_2892626), was affinity-purified using a GST-*SLITRK2* [705–784]-immobilized Sulfolink column (Pierce). The following commercially available antibodies were used: mouse monoclonal anti-GM130 (clone 35/GM130; BD Transduction Laboratories; Cat# 610822; RRID: AB\_398141), rabbit monoclonal anti-TIR (clone EPR20584; Abcam; Cat# ab214039; RRID: AB\_2904534), guinea pig polyclonal anti-VGLUT1 (Millipore; Cat# AB5905; RRID: AB\_2301751), mouse monoclonal anti-GAD67 (clone 1G10.2; Millipore; Cat# MAB5406; RRID: AB\_2278725), mouse monoclonal anti-Synaptophysin (clone SVP-38; Sigma-Aldrich; Cat# S5768; RRID: AB\_477523), mouse monoclonal anti-PSD-95 (clone K28/43; NeuroMab; Cat# 75-028; RRID: AB\_2877189), rabbit polyclonal anti-GABA<sub>A</sub>γ2 (Synaptic Systems; Cat# 224 003; RRID: AB\_2263066), guinea pig polyclonal anti-VGAT (Synaptic Systems; Cat# 131 004; RRID: AB\_887873), mouse monoclonal anti-HA (clone 16B12; BioLegend; Cat# 901501; RRID: AB\_2565006), mouse monoclonal anti-MAP2 (clone AP-20; Sigma-Aldrich; Cat# M1406; AB\_477171), rabbit polyclonal anti-MAP2 (Abcam; Cat# ab32454; AB\_776174), rabbit monoclonal anti-TrkB (clone 80E3; Cell Signaling; Cat# 4603; RRID: AB\_2155125), rabbit monoclonal anti-phospho-TrkB (Thermo Fisher; Cat# PAS-36695; RRID: AB\_2553666), rabbit monoclonal anti-BiP (clone C50B12; Cell Signaling; Cat# 3177; RRID: AB\_2119845), goat polyclonal anti-GFP (Rockland; Cat# 600-101-215; RRID: AB\_218182), mouse monoclonal anti-GluN1 (clone 54.1; Millipore; Cat# MAB363; RRID: AB\_94946), rabbit polyclonal anti-NR2A (Millipore; Cat# 07-632; RRID:

AB\_310837), mouse monoclonal anti- $\beta$ -actin (clone C4; Santa Cruz; Cat# sc-47778; RRID: AB\_2714189), mouse monoclonal anti-NeuN (clone A60; Millipore; Cat# MAB377; RRID: AB\_2298772), HRP-goat anti-human IgG cross-adsorbed secondary antibody (Thermo Fisher; Cat# 62-8420; RRID: AB\_2533962), Cy3-AffiniPure donkey Anti-rabbit IgG antibodies (Jackson ImmunoResearch; Cat# 711-165-152; RRID: AB\_2307443), Cy3-AffiniPure donkey anti-mouse IgG antibodies (Jackson ImmunoResearch; Cat# 715-165-150; RRID: AB\_2340813), Cy3-donkey anti-human IgG antibodies (Jackson ImmunoResearch; Cat# 709-165-149; RRID: AB\_2340535), Cy3-donkey anti-guinea pig IgG antibodies (Jackson ImmunoResearch; Cat# 706-035-148; RRID: AB\_2340447), FITC-AffiniPure donkey anti-mouse IgG antibodies (Jackson ImmunoResearch; Cat# 715-095-150; RRID: AB\_2340792), FITC-AffiniPure donkey anti-goat IgG antibodies (Jackson ImmunoResearch; Cat# 705-095-147; RRID: AB\_2340401), FITC-AffiniPure Donkey anti-rabbit IgG antibodies (Jackson ImmunoResearch; Cat# 711-095-152; RRID: AB\_2315776), and goat anti-guinea pig IgG antibodies (Thermo Fisher; Cat# A-21450; RRID: AB\_1418882). The rabbit polyclonal anti-VGLUT1 (JK111; RRID: AB\_2810945) and the rabbit polyclonal anti-PSD-95 (JK016; RRID: AB\_2722693) were previously described<sup>17</sup>. Rabbit polyclonal anti-pan-Shank (1172; RRID: AB\_2810261), rabbit polyclonal anti-GluA1 (1193; RRID: AB\_2722772), and rabbit polyclonal anti-GluA2 (1195; RRID: AB\_2722773) were gifts from Dr. Eunjoon Kim (KAIST, Korea) and previously described<sup>51</sup>. See Supplementary Table 3 for detailed information of antibody dilutions.

**Cell culture.** HEK293T cells (ATCC; CRL-3216) were cultured in Dulbecco's modified Eagle's medium (DMEM; Welgeline) supplemented with 10% fetal bovine serum (FBS; Tissue Culture Biologicals) and 1% penicillin-streptomycin (Thermo Fisher) at 37 °C in a humidified 5% CO<sub>2</sub> atmosphere. All procedures were performed according to the guidelines and protocols for rodent experimentation approved by the Institutional Animal Care and Use Committee of Daegu Gyeongbuk Institute of Science and Technology (DGIST).

**Cell-surface binding assays.** Recombinant PTP $\delta$  Fc-fusion proteins were produced in HEK293T cells. HEK293T cells were transfected with a PTP $\delta$ -IgC construct or pCMV-IgC empty vector. After 72 h, the media from transfected cells were collected, and 50 mM HEPES (pH 7.4) and 0.5 mM EDTA were added. Soluble Fc-fusion proteins were purified using protein A-Sepharose beads (GE Healthcare). Pulled-down proteins were eluted with 0.1 M glycine (pH 2.2) and then neutralized with 1 M Tris-HCl (pH 8.0). HEK293T cells expressing HA-SLITRK2 and its variants were incubated with 10  $\mu$ g/ml of the indicated PTP $\delta$  Fc-fusion proteins. Images were acquired using a confocal microscope (LSM800; Zeiss).

**Biotinylation assays.** HEK293T cells were transfected with HA-tagged SLITRK2 WT or the indicated missense variants. After 48 h, cells were washed twice with ice-cold phosphate-buffered saline (PBS), incubated with 1 mg/ml SulfoNHS-LC-biotin (Pierce) in ice-cold PBS for 30 min on ice, rinsed briefly once with 0.1 M glycine in PBS, and incubated with 0.1 M glycine/PBS for 10 min at room temperature to completely quench biotin reactions. Cells were lysed with lysis buffer (1% Triton X-100, 0.1% SDS, 50 mM Tris pH 7.4, 150 mM NaCl, and protease inhibitors) and incubated for 20 min on ice. After removal of cell debris by centrifugation, 250  $\mu$ g of lysates was incubated with 30  $\mu$ l streptavidin agarose beads (Pierce) for 4 h at 4 °C. After washing beads three times with lysis buffer, surface-labeled proteins were eluted with 2X sample buffer and analyzed by immunoblotting using the indicated antibodies. Anti-transferrin receptor (TfR) and  $\beta$ -actin antibodies were used as surface and total protein loading controls, respectively.

**Staining for surface/intracellular protein levels.** HEK293T cells were transfected with expression vectors for HA-SLITRK2 WT or the indicated variants. After 48 h, cells were washed twice with PBS, fixed with 3.7% formaldehyde for 10 min at 4 °C, and blocked with 3% horse serum/0.1% bovine serum albumin (BSA; crystalline grade) in PBS for 15 min at room temperature. Surface-expressed SLITRK2 protein was then detected by staining with mouse anti-HA antibody at room temperature. After 90 min, cells were washed twice with PBS and incubated with FITC-conjugated anti-mouse antibodies for 1 h at room temperature. Cells were then permeabilized with PBS containing 0.2% Triton X-100 for 10 min at 4 °C and incubated with rabbit anti-HA antibody for 90 min at room temperature to label intracellularly expressed SLITRK2 proteins, followed by incubation with Cy3-conjugated anti-rabbit secondary antibodies.

#### Production of recombinant viruses

**Lentiviruses.** Lentiviruses were produced by transfecting HEK293T cells with three plasmids—lentivirus vectors, psPAX2, and pMD2.G—at a 2:2:1 ratio. After 72 h, lentiviruses were harvested by collecting the media from transfected HEK293T cells and centrifuging at 1000  $\times$  g to remove cellular debris, as previously described<sup>47</sup>.

**Adeno-associated viruses (AAVs).** For high-efficiency transfections, AAVs were packaged with pHelper and AAV1.0 (serotype 2/9) capsids, as previously described<sup>48</sup>. Briefly, HEK293T cells were co-transfected with pHelper and

pAAV10.2/9 vectors together with pAAV-hSyn-SLITRK2-T2A-EGFP, pAAV-hSyn-SLITRK2-T2A-EGFP (T312A), pAAV-hSyn- $\Delta$ Cre-EGFP, or pAAV-hSyn-Cre-EGFP. Cells were harvested 72 h later, lysed, mixed with 40% polyethylene glycol and 2.5 M NaCl, and centrifuged at 2000  $\times$  g for 30 min. The resulting pellets were resuspended in HEPES buffer (20 mM HEPES, 115 mM NaCl, 1.2 mM CaCl<sub>2</sub>, 1.2 mM MgCl<sub>2</sub>, 2.4 mM KH<sub>2</sub>PO<sub>4</sub>), mixed with an equal volume of chloroform, and centrifuged at 400  $\times$  g for 5 min. The supernatants were concentrated three times with a Centriprep centrifugal filter (15 ml, 4310; Millipore) at 1220  $\times$  g for 5 min each and then with an Amicon Ultra centrifugal filter (0.5 ml, 3 K MWCO; Millipore) at 16,000  $\times$  g for 10 min. The infectious titer of viruses was assessed by qRT-PCR detection of EGFP sequences with subsequent reference to a standard curve generated using the pAAV-U6-EGFP plasmid.

**Animals.** All mice were on a C57BL/6J background, produced by back-crossing with C57BL/6J wild-type (WT) mice (purchased from Jackson Research Laboratory). The mice were kept and produced in the animal facility of Daegu Gyeongbuk Institute of Science and Technology under standard, temperature-controlled laboratory settings, including a 12-h light/dark cycle (lights on at 7 am) and free access to water and food. Pregnant Sprague-Dawley rats (Daehan Biolinek) were used to prepare in vitro cultures of dissociated hippocampal neurons. All procedures were carried out in compliance with the guidelines and protocols for rodent experimentation in accordance with protocols (DGIST-IACUC-19052109-00) approved by the Institutional Animal Care and Use Committee of Daegu Gyeongbuk Institute of Science and Technology.

**Generation of *Slitrk2*-floxed mice.** The C57BL/6J mouse strain was used to produce our *Slitrk2*-cKO mice. All mice were housed in a specific pathogen-free facility.

**Targeting strategy.** The knockout strategy for conditional deletion of *Slitrk2* targeted exon 2–3 using the Cre-loxP system. Specifically, 5' loxP was inserted into intron 1, and Frt-Neo-Frt-3' loxP was inserted downstream of exon 3; mouse embryonic stem cell (mESC)/homologous recombination (HR) technology developed by Biocytogen was used to establish the *Slitrk2*-cKO mouse model.

**Donor vector construction.** The 5' and 3' homologous arms used in gene-targeting span ~3.5 kb and 3.0 kb, respectively. The targeting vector also expresses the diphtheria toxin A (DTA) gene as a negative screening tag, which can be used to deplete ESC clones with non-homologous recombination and improve ESC clone screening efficiency.

**Microinjection.** The targeting vector was linearized and electroporated into C57BL/6J mESCs. After G418 selection, a total of 1000 clones were selected in three separate rounds. Selected cell clones were verified by Southern blotting, karyotyping, and sequence validation of PCR products. Chimeric mice were generated by injecting positively selected C57BL/6J mESCs into a total of 80 (two rounds of 30 and 50) pre-implantation BALB/c mouse embryos (blastocysts).

**Southern blot analysis.** Chimeric offspring were crossed with WT C57BL/6J mice to produce the F1 generation. Germline transmission was then confirmed by Southern blotting and PCR analysis of tail DNA from F1 generation mice. The restriction enzymes EcoRI (5' Probe-A), HindIII (Neo Probe (3')), and SacI (3'Probe-A) (all from New England Biolabs) were used for Southern blot analysis of genomic DNA. After digestion, genomic DNA extracted from mouse tails was separated on a 1% agarose gels and transferred to a positively charged nylon membrane (Hybond N+; Amersham International plc). The membrane was hybridized using DIG Easy Hyb Granules (Roche Applied Science Inc.) at 42 °C overnight in a mixture containing a PCR-generated probe labeled using the PCR DIG probe synthesis kit (Roche Applied Science Inc.). Hybridization signals were detected using the DIG Luminescent Detection Kit (Roche Applied Science Inc.).

**Quantitative RT-PCR in cultured neurons.** Cultured hippocampal and cortical neurons from *Slitrk2*-floxed mice were infected with  $\Delta$ Cre or Cre lentivirus at DIV4. Infected neurons were homogenized at DIV12 in TRIzol Reagent (Invitrogen), and total RNA was extracted according to the manufacturer's protocol. cDNA was prepared from total RNA using a cDNA synthesis kit (Takara Bio). qRT-PCR was performed on a CFX96 Real-Time PCR system (Bio-Rad) with TB Green reagents (Takara Bio) using a *Slitrk2*-specific probe (forward, 5'-GCA-GAGCTTGAGTATCTCTAATT-3'; reverse, 5'-GGACCTCAGCAGGTGTGTATT-3'). A probe for *Gapdh* (forward, 5'-ACATGGTCTACATGTTCCAG-3'; reverse, 5'-TCGCTCCTGGGAAGATGGTAT-3') was used as an endogenous control.

**Structural modeling.** The domain structures of human SLITRK2 (LRR1, aa C33-D270; LRR2, aa P341-P579) were modeled with AlphaFold<sup>22,23</sup>. LRR1 and LRR2 domains of human SLITRK2 were presumed to be connected by a flexible linker based on previous negative-stain electron microscopy images of the full

ectodomain of human SLITRK1<sup>14</sup>. Protein stability of domain structures of human SLITRK2 and variants was calculated using FoldX 5 suite<sup>49</sup>.

**Nissl staining.** Mice underwent intracardiac perfusion with PBS and then with 4% paraformaldehyde. Fixed brain tissue was isolated, post-fixed for 12 h at 4 °C, and dehydrated in 30% sucrose for 6–8 d. Then, brain tissue was embedded in OCT (Optimal Cutting Temperature) compound, frozen, cryo-sectioned at 20- $\mu$ m thicknesses using a cryostat (Leica CM5120). Slices were mounted on glass slides, rinsed with PBS, and permeabilized with 0.1% Triton X-100 in PBS for 10 min. Permeabilized samples were washed twice with PBS (5 min each), then incubated for 20 min in 200  $\mu$ l of NeuroTrace 500/525 Green Fluorescent Nissl Stain (Molecular Probes), diluted 1:100 in PBS before use. Following three more washes with PBS, the specimens were dried and mounted using Vectashield Mounting Medium containing 4',6-diamidino-2-phenylindole (DAPI; Vector Laboratories). A slide scanner (Axio Scan.Z1; Carl Zeiss) was used to capture green fluorescence.

**Measurement of mouse ventricle volumes.** Acute brain slices (50  $\mu$ m thickness) were obtained from 7-wk-old mice using a vibratome. Every sixth brain section (300  $\mu$ m between sections) was collected starting from bregma +0.86 mm and continuing to the dorsal end at bregma -0.82 mm. Sections were immunostained with anti-NeuN (clone A60; Millipore), and whole-brain slice images were obtained using a slide scanner (Axio Scan.Z1; Carl Zeiss) with a  $\times 20$  objective lens. All image settings were kept constant. Lateral ventricle areas were manually sketched and calculated using MetaMorph software (Molecular Devices Corp.). Lateral ventricle volumes were calculated as the sum of sectional areas of lateral ventricles times the layer thickness.

**Heterologous synapse-formation assays.** Forty-eight hours after transfecting with the indicated expression vectors, HEK293T cells were trypsinized, seeded onto cultured hippocampal neurons at days in vitro (DIV)10, and cocultured for 48 h, as indicated. Cocultured neurons were co-immunostained with antibodies against the indicated proteins. Fluorescence images were acquired by confocal microscopy (LSM780, Carl Zeiss), and results were quantified by measuring the fluorescence intensities of synaptic marker puncta in randomly selected transfected HEK293 cells (region of interest), normalized with respect to the area of each cell. Results were quantified for both red and green channels using MetaMorph Software (Molecular Devices Corp.; RRID: SCR\_002368).

**Primary neuronal culture, transfection, immunocytochemistry, and image acquisition and analysis.** Hippocampal and cortical mouse neuron cultures were prepared from embryonic day 17 (E17) *Slitrk2*-floxed mouse embryos, as previously described<sup>50</sup>. In brief, neurons were seeded on 18-mm poly-D-lysine (0.1 mg/ml)-coated coverslips and cultured in Neurobasal media (Gibco) containing penicillin-streptomycin and 0.5 mM GlutaMax (Thermo Fisher) supplemented with 2% B-27 (Thermo Fisher) and 0.5% FBS (Hyclone). Cultured hippocampal neurons were infected with lentiviruses expressing Cre or inactive Cre ( $\Delta$ Cre) recombinase at DIV5, and immunostained at DIV12 or DIV14. For rescue experiments, hippocampal neurons were transfected at DIV8–9 with SLITRK2 WT or the indicated point variants using a CalPhos Kit (Clontech) and immunostained at DIV13–14; controls were transfected with EGFP. For immunocytochemistry, cultured rat neurons were fixed with 4% paraformaldehyde/4% sucrose in PBS for 30 min at 4 °C and permeabilized with 0.2% Triton X-100 in PBS for 30 min at 4 °C. Neurons were then blocked by incubating with 3% horse serum/0.1% BSA in PBS for 15 min at room temperature, then stained with the indicated primary and secondary antibodies for 70 min at room temperature. Z-stack images of randomly selected neurons were acquired using a confocal microscope (LSM780; Carl Zeiss) with a  $\times 63$  objective lens. Obtained Z-stack images were converted to maximal projections, and colocalized puncta densities were analyzed in a blinded manner using MetaMorph software (Molecular Devices Corp.).

**Sholl analysis.** Rat cultured hippocampal neurons were co-transfected with EGFP and SLITRK2 WT or the indicated SLITRK2 variants for 5 d at DIV9. Transfected neurons were stained with anti-EGFP and anti-HA antibodies, and costained neurons were used for quantitative analysis. Dendritic branching was analyzed by performing a Sholl analysis using the Sholl Analysis plugin installed in ImageJ software (Fiji, RRID: SCR\_002285), as previously described<sup>51</sup>. The number of branches at increasing distances from the soma was quantified.

**Stereotaxic surgery.** Mice were anesthetized with isoflurane (3–4%) and then placed in a stereotaxic device. Viral solutions (titers  $\geq 1 \times 10^{11}$  viral genomes/ml) were injected bilaterally at a rate of 0.2  $\mu$ l/min with a NanoFil syringe and a Nanoliter 2010 Injector (World Precision Instruments). The AP -2.5 mm, ML  $\pm$  1.7 mm, and DV -1.3 mm (from the dura) coordinates were used for hippocampal CA1 regions. Mice were used  $\geq 3$  weeks after stereotaxic surgeries.

**Immunohistochemistry.** Mice were transcardially perfused first with PBS for 3 min and then with 4% paraformaldehyde in PBS for 5 min. After post-fixation

overnight, mouse brains were slowly sectioned at a thickness of 40  $\mu$ m using a vibratome (VT1200S; Leica). Brain sections were permeabilized by incubation with 0.2% Triton X-100 in PBS containing 10% horse serum and 0.2% BSA for 1 h at room temperature. For immunostaining, sections were incubated overnight at 4 °C with primary anti-VGLUT1 (1:500), anti-PSD-95 (1:500), or anti-VGAT (1:500) antibodies diluted in the same blocking solution. After washing three times, sections were incubated with Cy3-conjugated secondary antibodies (Jackson ImmunoResearch) for 1 h at room temperature. Sections were then washed extensively and mounted on glass slides with Vectashield Mounting Medium (Vector Laboratories). Images were acquired by confocal microscopy (LSM700; Zeiss). Synaptic puncta were quantified using MetaMorph software (Molecular Devices), and their density and average area were measured.

**Mouse behavioral tests.** Male mice at the age of 6–10 weeks were used for all behavioral tests. Tests were conducted in a sound-proof room under dim lighting (<5 lux).

**Open-field test.** Each mouse was placed into an open-field apparatus (40  $\times$  40  $\times$  40 cm) and allowed to freely explore the field for 30 min. The total distance traveled and time spent in the center area were recorded by a top-view infrared camera and analyzed using EthoVision XT 10.5 (Noldus).

**Three-chamber test.** The testing apparatus consisted of a white acrylic box divided into three chambers (20  $\times$  40  $\times$  22 cm each) with small openings on the dividing walls. Wire cups were employed to enclose social conspecifics in the corners of both side-chambers. Mouse was placed in the central chamber for a 10-min habituation. Following habituation, an age-matched social conspecific was placed into the wire cup on the left side-chamber, and the sociability of the subjects was assessed by measuring subjects' exploration times for the enclosed conspecific and the empty cup during the second 10-min session. An exploration was counted when the subjects directed their nose into the vicinity of the wire cups. In the last 10-min session, a new social conspecific was placed into the empty wire cup on the right side-chamber, and social recognition was assessed by measuring subjects' exploration times for the familiar conspecific and the novel conspecific in a wire cup.

**Elevated plus-maze test.** The white acrylic maze with two open arms (30  $\times$  5  $\times$  0.5 cm) and two closed arms (30  $\times$  5  $\times$  30 cm) was elevated 75 cm over the floor. Mice were individually placed at the center of the elevated plus-maze and allowed to freely move for 5 min. All behaviors were recorded by a top-view infrared camera, and the time spent in each arm and the number of arm entries were analyzed using EthoVision XT 10.5 (Noldus).

**Light and dark box test.** The apparatus consists of a roofless box divided into a closed dark chamber and a brightly illuminated chamber. A small entrance allows free travel between the two chambers. The light chamber was illuminated at 350 lux. The movement of mice was recorded by a top-view infrared camera and EthoVision XT 10.5 (Noldus) was used for analysis of the time spent in each chamber and the number of transitions.

**Y-maze test.** Spatial working memory was evaluated on the basis of spontaneous alternation behavior in a Y-maze consisting of three 40-cm-long arms at a 120° angle from each other. Each mouse was placed at the center of the maze and allowed to explore freely for 8 min. An entry was counted when the center point of a mouse was within the arm. The movement of mice was recorded by a top-view infrared camera and analyzed using EthoVision XT 10.5 (Noldus).

**Novel object-recognition test.** Object recognition test was performed in the open-field box. On the first day, mice were habituated to the maze for more than 10 min. On the subsequent day, mice were allowed to explore two identical objects twice for 10 min. Exploration time for each object was measured. Twelve hours later, one of two objects was replaced with a new distinct object and mice were returned to the field and allowed to explore for 10 min to test novel-object recognition. Exploration time for each object was measured.

**Barnes maze test.** The Barnes maze test was conducted on a white circular platform, 95 cm in diameter, with 20 holes equally spaced around the perimeter. The hole above the escape box represented the target. The location of the target was consistent for a given mouse but randomized across mice. The maze was rotated daily, with the spatial location of the target unchanged with respect to the distal visual room cues, to prevent a bias based on olfactory or proximal cues within the maze. Three trials per day were conducted for 7 successive days. On day 10, the first probe trial was conducted without the escape box to confirm that the spatial task was acquired based on navigation by distal environment room cues. On day 17, the second probe trial was conducted. During acquisition and probe trials, the following parameters were measured by a top-view infrared camera and analyzed using EthoVision XT 10.5 (Noldus): velocity, error score, total distance, first visit time to escape box, and total latency.



**Grip strength test.** A grip strength meter (BIOSEB) was used to measure forelimb grip strength. A mouse was allowed to grasp the steel bar mounted on the force gauge. The gauge was reset after stabilization, and the mouse's tail was slowly pulled back by an experimenter. Maximal grip force was automatically recorded as newtons by the gauge. We performed 3 consecutive measurements per day for 3 consecutive days.

**Gait test.** The forepaws and hind paws of a mouse were coated with different colored nontoxic acrylic paints. The mice were trained to run down the runway (30-cm long and 5-cm wide) in a straight line before the test. The footprints were traced on the white paper covering a runway. The footprints were analyzed for overlap, each side of stride length, each side of stance length, and each side of sway length. All parameters were analyzed by MetaMorph software (Molecular Devices Corp.).

### Electrophysiology

**Cultured neuron electrophysiology.** Hippocampal cultured neurons obtained from *Slitrk2*-cKO mice were infected on DIV4 with lentiviruses encoding Cre-EGFP or  $\Delta$ Cre-EGFP, followed by analysis at DIV13–16. Similarly, hippocampal cultured neurons obtained from Nestin-*Slitrk2* mice were analyzed at DIV13–16. Pipettes were pulled from borosilicate glass (o.d. 1.5 mm, i.d. 0.86 mm; Sutter Instruments) using a Model P-97 pipette puller (Sutter Instruments). The resistance of patch pipettes filled with internal solution varied between 3 and 6 M $\Omega$ . The internal solution contained 145 mM CsCl, 5 mM NaCl, 10 mM HEPES, 10 mM EGTA, 0.3 mM Na-GTP and 4 mM Mg-ATP, with pH adjusted to 7.2–7.4 with CsOH and an osmolarity of 290–295 mOsmol/l. The external solution consisted of 130 mM NaCl, 4 mM KCl, 2 mM CaCl<sub>2</sub>, 1 mM MgCl<sub>2</sub>, 10 mM HEPES, and 10 mM D-glucose, with pH adjusted to 7.2–7.4 with NaOH and an osmolarity of 300–305 mOsmol/l. The whole-cell configuration was obtained at room temperature using  $\mu$ -TSC manipulators (SENSAPLEX). Electrophysiological data were acquired with a Multiclamp 700B amplifier (Axon instrument) and pCLAMP software, and were digitized using an Axon DigiData 1550B data acquisition board (Axon Instruments). mEPSCs were recorded at a holding potential of  $-70$  mV. Synaptic currents were analyzed offline using Clampfit 10.8 software (Molecular Devices). For mEPSC recordings, the external solution contained 1  $\mu$ M TTX and 50  $\mu$ M picrotoxin to block GABA<sub>A</sub> receptor and Na<sup>+</sup> currents, respectively. For mIPSC recordings, the external solution contained 1  $\mu$ M TTX, 10  $\mu$ M CNQX, and 50  $\mu$ M D-AP5 to block Na<sup>+</sup> currents, AMPA receptors, and NMDA receptors, respectively.

**Hippocampal CA1 pyramidal neuron electrophysiology.** Whole-cell voltage-clamp recordings were obtained from acute brain slices. Brain slices were transferred to a recording chamber and perfused with a bath solution of aerated (O<sub>2</sub> 95%/CO<sub>2</sub> 5% mixed gas) artificial cerebrospinal fluid (aCSF) consisting of 124 mM NaCl, 3.3 mM KCl, 1.3 mM NaH<sub>2</sub>PO<sub>4</sub>, 26 mM NaHCO<sub>3</sub>, 11 mM D-glucose, 2.5 mM CaCl<sub>2</sub>, and 1.5 mM MgCl<sub>2</sub> at 28–30 °C. For measuring postsynaptic currents, patch pipettes (open pipette resistance, 3–5 M $\Omega$ ) were filled with an internal solution consisting of 145 mM CsCl, 5 mM NaCl, 10 mM HEPES, 10 mM EGTA, 4 mM Mg-ATP, and 0.3 mM Na-GTP. Whole-cell recordings of mEPSCs were performed on CA1 pyramidal neurons, voltage clamped at  $-70$  mV, and currents were pharmacologically isolated by bath applications of 50  $\mu$ M picrotoxin (Tocris), 1  $\mu$ M tetrodotoxin (Tocris). Electrophysiological data were acquired using pCLAMP software and a Multiclamp 700B (Axon Instruments) and were digitized using an Axon DigiData 1550B data acquisition board (Axon Instruments). Data were sampled at 10 kHz and filtered at 4 kHz. Data were discarded if the series resistance was >30 M $\Omega$ , or the series resistance differed by more than 20%.

**Data analyses.** Data analysis and statistical tests were performed using GraphPad Prism8.0 software (RRID: SCR\_002798). Heterologous synapse-formation assays and surface-binding assays were quantified by randomly selecting transfected HEK293T cells as the region of interest. The fluorescence intensities of synaptic marker puncta or Fc-fusion proteins were normalized to transfected protein signal intensities using MetaMorph software (Molecular Devices Corp.). All data are expressed as means  $\pm$  standard error of the mean (SEM) unless stated otherwise, and significance is indicated with an asterisk. All experiments were performed using at least three independent mice, cultures, and/or cohorts of grouped mice, and the normality of data distributions was evaluated using the Shapiro–Wilk test. Data were compared using Student's *t*-test or one-way analysis of variance (ANOVA) using a non-parametric Kruskal–Wallis test, followed by Dunn's multiple comparison test for post hoc group comparisons, *t*-test, Mann–Whitney *U* test, or Fisher's least significance difference; 'n' values used are indicated in the figure legends, and numbers shown indicate replicates. Tests used to determine statistical significance are stated in the text and legends of figures depicting the results of the respective experiments. A *p*-value <0.05 was considered statistically significant, and individual *p*-values are indicated in the respective figure legend. All experiments were performed and analyzed in a blind manner by the experimenter.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### Data availability

Data sets presented in this study are included in full wherever possible, including the display of individual data points. All relevant data supporting the findings of this study are available from the corresponding authors upon reasonable request. Biological materials, including mutant mice and custom antibodies generated for this study, will be shared upon request within the limits of respective material transfer agreements for as long as they are available in the laboratory. Source data are provided with this paper.

### Code availability

All relevant code supporting the findings of this study is available from the corresponding authors upon request.

Received: 17 September 2021; Accepted: 22 June 2022;

Published online: 15 July 2022

### References

- Geetz, J., Shoubridge, C. & Corbett, M. The genetic landscape of intellectual disability arising from chromosome X. *Trends Genet.* **25**, 308–316 (2009).
- Tarpey, P. S. et al. A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation. *Nat. Genet.* **41**, 535–543 (2009).
- Kaufman, L., Ayub, M. & Vincent, J. B. The genetic basis of non-syndromic intellectual disability: a review. *J. Neurodev. Disord.* **2**, 182–209 (2010).
- Ropers, H. H. X-linked mental retardation: many genes for a complex disorder. *Curr. Opin. Genet. Dev.* **16**, 260–269 (2006).
- Lubs, H. A., Stevenson, R. E. & Schwartz, C. E. Fragile X and X-linked intellectual disability: four decades of discovery. *Am. J. Hum. Genet.* **90**, 579–590 (2012).
- Piton, A., Redin, C. & Mandel, J. L. XLID-causing mutations and associated genes challenged in light of data from large-scale human exome sequencing. *Am. J. Hum. Genet.* **93**, 368–383 (2013).
- Neri, G., Schwartz, C. E., Lubs, H. A. & Stevenson, R. E. X-linked intellectual disability update 2017. *Am. J. Med. Genet. A* **176**, 1375–1388 (2018).
- Aruga, J. & Mikoshiba, K. Identification and characterization of *Slitrk*, a novel neuronal transmembrane protein family controlling neurite outgrowth. *Mol. Cell Neurosci.* **24**, 117–129 (2003).
- Aruga, J., Yokota, N. & Mikoshiba, K. Human *SLITRK* family genes: genomic organization and expression profiling in normal brain and brain tumor tissue. *Gene* **315**, 87–94 (2003).
- Yim, Y. S. et al. *Slitrks* control excitatory and inhibitory synapse formation with LAR receptor protein tyrosine phosphatases. *Proc. Natl Acad. Sci. USA* **110**, 4057–4062 (2013).
- Beaubien, F., Raja, R., Kennedy, T. E., Fournier, A. E. & Cloutier, J. F. *Slitrk1* is localized to excitatory synapses and promotes their development. *Sci. Rep.* **6**, 27343 (2016).
- Takahashi, H. et al. Selective control of inhibitory synapse development by *Slitrk3*-PTPdelta trans-synaptic interaction. *Nat. Neurosci.* **15**, 389–398 (2012).
- Song, M. et al. *Slitrk5* mediates BDNF-dependent TrkB receptor trafficking and signaling. *Dev. Cell* **33**, 690–702 (2015).
- Um, J. W. et al. Structural basis for LAR-RPTP/*Slitrk* complex-mediated synaptic adhesion. *Nat. Commun.* **5**, 5423 (2014).
- Yamagata, A. et al. Structure of *Slitrk2*-PTPdelta complex reveals mechanisms for splicing-dependent trans-synaptic adhesion. *Sci. Rep.* **5**, 9686 (2015).
- Li, J. et al. A conserved tyrosine residue in *Slitrk3* carboxyl-terminus is critical for GABAergic synapse development. *Front. Mol. Neurosci.* **12**, 213 (2019).
- Han, K. A. et al. *Slitrk2* controls excitatory synapse development via PDZ-mediated protein interactions. *Sci. Rep.* **9**, 17094 (2019).
- Piton, A. et al. Systematic resequencing of X-chromosome synaptic genes in autism spectrum disorder and schizophrenia. *Mol. Psychiatry* **16**, 867–880 (2011).
- Kang, H. et al. *Slitrk* missense mutations associated with neuropsychiatric disorders distinctively impair *slitrk* trafficking and synapse formation. *Front. Mol. Neurosci.* **9**, 104 (2016).
- Leblond, C. S. et al. Meta-analysis of SHANK mutations in autism spectrum disorders: a gradient of severity in cognitive impairments. *PLoS Genet.* **10**, e1004580 (2014).

21. Sobreira, N., Schiettecatte, F., Valle, D. & Hamosh, A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Hum. Mutat.* **36**, 928–930 (2015).
22. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
23. Tunyasuvunakool, K. et al. Highly accurate protein structure prediction for the human proteome. *Nature* **596**, 590–596 (2021).
24. Marino, S. M. & Gladyshev, V. N. Cysteine function governs its conservation and degeneration and restricts its utilization on protein surfaces. *J. Mol. Biol.* **404**, 902–916 (2010).
25. Galichet, C., Lovell-Badge, R. & Rizzotti, K. Nestin-Cre mice are affected by hypopituitarism, which is not due to significant activity of the transgene in the pituitary gland. *PLoS ONE* **5**, e11443 (2010).
26. Song, M. et al. Rare synaptogenesis-impairing mutations in SLITRK5 are associated with obsessive compulsive disorder. *PLoS ONE* **12**, e0169994 (2017).
27. Li, Y. X. et al. Expression of a dominant negative TrkB receptor, T1, reveals a requirement for presynaptic signaling in BDNF-induced synaptic potentiation in cultured hippocampal neurons. *Proc. Natl Acad. Sci. USA* **95**, 10884–10889 (1998).
28. Gomes, J. R. et al. Excitotoxicity downregulates TrkB FL signaling and upregulates the neuroprotective truncated TrkB receptors in cultured hippocampal and striatal neurons. *J. Neurosci.* **32**, 4610–4622 (2012).
29. Qiu, L. L. et al. Dysregulation of BDNF/TrkB signaling mediated by NMDAR/Ca(2+)/calpain might contribute to postoperative cognitive dysfunction in aging mice. *J. Neuroinflammation* **17**, 23 (2020).
30. Harding, H. P., Calton, M., Urano, F., Novoa, I. & Ron, D. Transcriptional and translational control in the mammalian unfolded protein response. *Annu. Rev. Cell Dev. Biol.* **18**, 575–599 (2002).
31. Ron, D. Translational control in the endoplasmic reticulum stress response. *J. Clin. Invest.* **110**, 1383–1388 (2002).
32. Martinez, G., Khatiwada, S., Costa-Mattioli, M. & Hetz, C. ER proteostasis control of neuronal physiology and synaptic function. *Trends Neurosci.* **41**, 610–624 (2018).
33. Barnes, C. A. Memory deficits associated with senescence: a neurophysiological and behavioral study in the rat. *J. Comp. Physiol. Psychol.* **93**, 74–104 (1979).
34. Fox, G. B., Fan, L., LeVasseur, R. A. & Faden, A. I. Effect of traumatic brain injury on mouse spatial and nonspatial learning in the Barnes circular maze. *J. Neurotrauma* **15**, 1037–1046 (1998).
35. Um, J. W. & Ko, J. LAR-RPTPs: synaptic adhesion molecules that shape synapse development. *Trends Cell Biol.* **23**, 465–475 (2013).
36. Loomis, C. et al. Identification of MAGUK scaffold proteins as intracellular binding partners of synaptic adhesion protein Slitrk2. *Mol. Cell Neurosci.* **103**, 103465 (2020).
37. Piton, A. et al. Mutations in the calcium-related gene IL1RAPL1 are associated with autism. *Hum. Mol. Genet.* **17**, 3965–3974 (2008).
38. Baer, S. et al. Wiedemann-Steiner syndrome as a major cause of syndromic intellectual disability: A study of 33 French cases. *Clin. Genet.* **94**, 141–152 (2018).
39. Luikart, B. W. & Parada, L. F. Receptor tyrosine kinase B-mediated excitatory synaptogenesis. *Prog. Brain Res.* **157**, 15–24 (2006).
40. Ataullah, A. H. M. & Naqvi, I. A. in *StatPearls* (Treasure Island (FL), 2022).
41. Batiuk, M. Y. et al. Identification of region-specific astrocyte subtypes at single cell resolution. *Nat. Commun.* **11**, 1220 (2020).
42. Adzhubei, I. A. et al. A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249 (2010).
43. Ng, P. C. & Henikoff, S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* **31**, 3812–3814 (2003).
44. Kircher, M. et al. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **46**, 310–315 (2014).
45. Kim, S. et al. Npas4 regulates IQSEC3 expression in hippocampal somatostatin interneurons to mediate anxiety-like behavior. *Cell Rep.* **36**, 109417 (2021).
46. Ranaivosoa, F. M. et al. A proteomic screen of neuronal cell-surface molecules reveals IgLONs as structurally conserved interaction modules at the synapse. *Structure* **27**, 893–906 (2019).
47. Han, K. A. et al. LAR-RPTPs directly interact with neurexins to coordinate bidirectional assembly of molecular machineries. *J. Neurosci.* **40**, 8438–8462 (2020).
48. Kim, J. et al. LRRTM3 regulates activity-dependent synchronization of synapse properties in topographically connected hippocampal neural circuits. *Proc. Natl Acad. Sci. USA* **119**, e2110196119 (2022).
49. Schymkowitz, J. et al. The FoldX web server: an online force field. *Nucleic Acids Res.* **33**, W382–W388 (2005).
50. Um, J. W., Han, K. A., Choi, S. Y. & Ko, J. Protocol for quantitative analysis of synaptic vesicle clustering in axons of cultured neurons. *STAR Protoc.* **1**, 100095 (2020).
51. Han, K. A. et al. Neurotrophin-3 regulates synapse development by modulating TrkC-PTPsigma synaptic adhesion and intracellular signaling pathways. *J. Neurosci.* **36**, 4816–4831 (2016).

### Acknowledgements

We thank Jinha Kim (DGIST, Korea) for technical assistance, Vaidichi Jobanputra (Columbia University, USA) and Lenet Watt Skovström (Amplexa Genetics, Denmark) for genomic analysis. This study was supported by grants from the National Research Foundation of Korea (NRF) funded by the Ministry of Science and Future Planning (2019R1A2C1086048 and 2020R1A4A1019009 to J.W.U.; 2021R1A2C1091863 to J.Ko.), the DGIST R&D Program of the Ministry of Science and ICT (22-CoE-BT-01 to J.Ko. and J.W.U.), Institute for Basic Science (IBS-R030-C1 to H.M.K.), NIH (NS106298 to M.C.K.), and SFARI and JPB Foundation (to W.K.C.).

### Author contributions

Conceptualization: J.Ko, A.P., and J.W.U. Methodology: K.A.H., D.K., G.J., D.L., H.Y.K., J.Kim., and A.r.H. Formal analysis: A.P., S.B., M.C.K., J.W., W.K.C., G.V., I. C., M.P., J.G., K.H., M.W., A.B., A.F.J., A.S., S.M., A.P.M.d.B., A.V.S., M.A., J.S., S.K., B.I., B.C., M.N., C.F., J.M., E.T., D.K.G., and M.W. Writing: original draft, S.E.C., J.Ko., A.P., and J.W.U.; review and editing, S.E.C., K.A.H., H.M.K., J.Ko., A.P., and J.W.U. Supervision: J.Ko., A.P., and J.W.U. Funding acquisition: H.M.K., M.C.K., W.K.C., J.Ko., and J.W.U.

### Competing interests

The authors declare no competing interests.

### Additional information


**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-31566-z>.

**Correspondence** and requests for materials should be addressed to Jaewon Ko, Amélie Piton or Ji Won Um.

**Peer review information** *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer review reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022

<sup>1</sup>Service de Génétique Médicale, Institut de Génétique Médicale d'Alsace (IGMA), Hôpitaux Universitaires de Strasbourg, Strasbourg, France. <sup>2</sup>Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), INSERM U1258, CNRS-UMR7104, Université de Strasbourg, Illkirch-Graffenharden, France. <sup>3</sup>Laboratoire de Génétique Médicale, UMRS\_1112, Institut de Génétique Médicale d'Alsace (IGMA), Université de Strasbourg et INSERM, Strasbourg, France. <sup>4</sup>Department of Brain Sciences, Daegu Gyeongbuk Institute of Science and Technology (DGIST), Daegu 42988, Korea. <sup>5</sup>Pediatric Movement Disorders Program, Division of Pediatric Neurology, Barrow Neurological Institute, Phoenix

### III. Troisième partie : poursuite des explorations par WGS chez des patients avec DI supposée liée à l’X non-résolue par le séquençage ciblé ou l’ES.

#### 1. Cohorte des 15 patients avec DI présumée liée à l’X

##### 1. Matériel et méthodes

###### a) Financement, recrutement des patients, circuits des ADN

###### ★ Financement

En amont de ce travail, j’ai déposé un projet et obtenu le prix de l’association Xtraordinaire (<https://www.xtraordinaire.org/>) permettant de financer cette partie du projet de thèse, notamment le séquençage de génome chez les 15 patients avec DI supposée liée à l’X.

###### ★ Recrutement des patients

Pour tenter de recruter suffisamment de patients présentant une déficience intellectuelle avec présomption de transmission liée au chromosome X, j’ai d’abord étudié la cohorte de patients analysés dans notre laboratoire de diagnostic génétique en séquençage d’exome ou de panel de 456 gènes de DI (détails dans la section résultats). J’ai étudié les dossiers des patients négatifs (absence de variant de classe 3 à 5) un à un, afin de réaliser un « pré-tri » des familles chez lesquelles la probabilité transmission liée à l’X était la plus forte. Les patients étaient catégorisés de la présomption la plus forte à la moins forte d’une liaison à l’X.

J’ai par la suite contacté les cliniciens en vue de recueillir des données anamnestiques, cliniques et iconographiques. Les critères de sélection de ces patients étaient les suivants : présence dans une même famille d’au minimum 3 individus atteints de DI de sexe masculin, sur 2 générations et reliés par une femme, symptomatique ou non. Malheureusement, à l’issue de cette première étape, très peu de patients ont été sélectionnés pour les raisons suivantes :

- Famille ne remplissant pas les critères d’inclusion.
- Patient décédé ou perdu de vue par le clinicien.
- Panel ou exome en cours.
- Absence de réponse du clinicien.
- Données cliniques manquantes
- Absence de consentement signé pour participer à une étude de recherche.
- Familles complexes où ségrégaient plusieurs maladies rares avec des phénotypes mal caractérisés ou difficilement individualisables (par exemple, identification d’un variant

pathogène hétérozygote dans un gène de DI chez un individu et absence de ce dernier chez les deux frères également atteints de DI).

J'ai donc initié plusieurs appels à collaboration nationaux, ce qui a permis de recruter des patients supplémentaires. A l'issue de cette phase, **quinze familles ont été sélectionnées**. En parallèle, un **recueil des échantillons d'ADN des patients et de leurs parents** a été réalisé suivi de leur stockage au laboratoire de diagnostic génétique des HUS après obtention d'un consentement signé et adapté à la réalisation de techniques de haut débit en recherche (Figure 55).

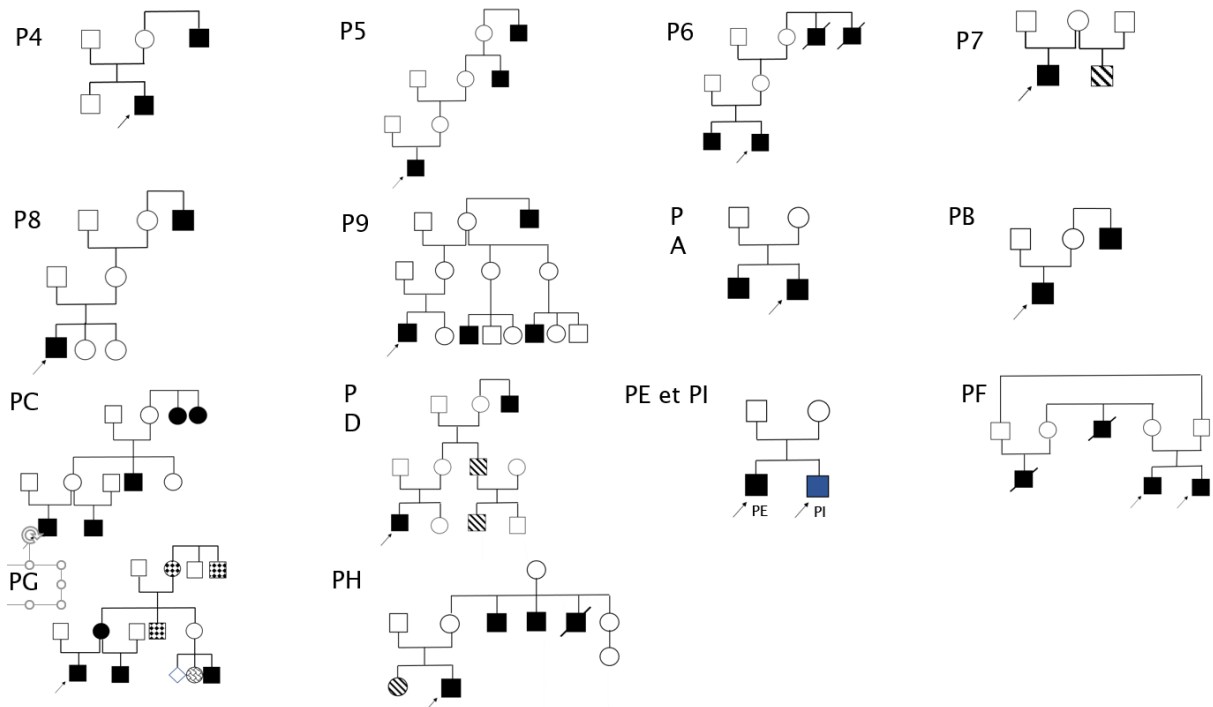


Figure 55: Familles recrutées dans le présent travail

En noir : individus présentant une DI, en hachuré ou damier : individus présentant un TND sans DI

Pour chaque patient, j'ai effectué le **recensement des différentes analyses génétiques** déjà réalisées. Pour les patients non-encore explorés par le biais de techniques de haut débit, nous avons **réalisé un séquençage d'exome ou d'un panel de 556 gènes impliqués dans la DI** au sein de notre laboratoire de diagnostic génétique. Ces analyses se sont révélées négatives chez tous.

### ★ Circuit des ADN

Tous les échantillons de sang reçus pour cette étude ont bénéficié d'une procédure d'extraction d'ADN réalisée au sein du laboratoire de diagnostic des HUS selon une méthode préalablement validée par la plateforme de séquençage du Centre National de Recherche en Génomique Humaine (CNRGH, Evry). Après extraction, des aliquotes de 3 µg d'ADN de chacun des quinze patients inclus (*soló*), marquées d'un code-barres anonyme, ont été envoyées au CNRGH par transporteur à température ambiante.

Plusieurs contrôles de qualité étaient alors effectués par le CNRGH incluant une quantification fluorimétrique de l'ADN, une mesure de la qualité à l'aide du numéro d'intégrité de l'ADN, un test d'amplification par PCR et le contrôle du sexe. Si l'ADN passait favorablement les contrôles qualité, 1,1 µg d'ADN était fragmenté et les librairies étaient préparées (sans capture) selon un protocole spécifique du CNRGH pour le WGS (kits de préparation de librairies Illumina TruSeq PCR-free). Le séquençage était réalisé en *paired-end* 2 x 150 b, *via* une plateforme Illumina NovaSeq 6000, avec une couverture de 30 X.

## *b) Analyses bioinformatiques*

### **1 - Identification des variants**

Les séquences génomiques ont été analysées sur un pipeline développé et validé par la plateforme de séquençage du CNRGH. Les données étaient produites sous forme de fichiers FASTQ générés par cette plateforme. Les séquences étaient ensuite alignées sur le génome humain de référence (GRCh38) à l'aide du logiciel BWA (H. Li et Durbin, 2009) et mises à disposition sous forme de fichiers BAM. Un réalignement des séquences autour des sites d'insertion et de délétion et le recalibrage de la qualité des bases étaient effectués à l'aide du logiciel GATK (McKenna *et al.*, 2010). Le *variant calling* des SNV et d'*indel* était effectué à l'aide du logiciel Haplotype Caller de GATK. Après cette partie du pipeline bioinformatique réalisée au CNRGH, j'ai utilisé les outils Manta (X. Chen *et al.*, 2016), Smoove (<https://github.com/brentp/smoove>) et CNVpytor (Suvakov *et al.*, 2021) pour identifier les variants de structure (SV) incluant les anomalies du nombre de copies (CNV) (Figure 56).

### **2 - Annotation et filtres utilisés pour les variants ponctuels (SNV/indel)**

Les SNV/indel identifiés et stockés sous forme de fichiers VCF, ont été annotés *via* le logiciel VaRank (Geoffroy *et al.*, 2015) (V Geoffroy, T Alouane, JB Lamouche – U1112 et IGBMC). VaRank utilise Alamut Batch (Interactive Biosoftware) pour rassembler les annotations génomiques et leurs effets de prédiction au niveau nucléotidique et protéique. VaRank permet ainsi de réunir un ensemble d'informations pour chaque variant, provenant de différents logiciels ou bases de données, telles que la présence dans les bases de données ClinVar et gnomAD, les prédictions d'un potentiel effet fonctionnel du changement d'acide aminé (SIFT, PolyPhen-2, Mutation Taster, AGVGD, Grantham), d'un effet sur l'épissage (Human Splicing Finder, MaxEntScan, NNSplice) compilées sous la forme d'un item dénommé « local splice effect », et les scores de conservation nucléotidique (Phastcons, PhyloP). VaRank effectue de manière automatisée une compilation de l'ensemble de ces informations et calcule un score de pathogénicité potentiel (score VaRank : de 0 à 110) pour chaque SNV/indel et les potentiels variants candidats sont classés en fonction de ce score. D'autres outils de prédiction d'épissage ont été utilisés pour ce

travail tels que SpliceAI (Jaganathan *et al.*, 2019) et SPiP (Splicing Prediction Pipeline) (Leman *et al.*, 2022). Concernant SpliceAI, nous avons conservé les variants dont au moins un des 4 « deviation scores (DS) » (gain ou perte de site accepteur (DS-AG, DS-AL) ou de site donneur (DS-DG, DS-DL)) était supérieur à 0,5 (>0,5). Cette valeur seuil pouvait être revue à la baisse si aucun variant candidat n'était identifié. Pour SPiP, le score de 0,452 était utilisé pour filtrer les variants (Leman *et al.*, 2022).

Les SNV/indel ont ensuite été filtrés pour :

- 1 - Exclure les variants avec fréquence > 0,5% dans gnomAD
- 2 - Exclure les variants avec nombre d'homozygotes > 5 dans gnomAD
- 3 - Exclure les variants de mauvaise qualité (non PASS)
- 4 - Exclure les variants bénins dans ClinVar (benign, likely benign)
- 5 - Conserver les variants pathogéniques dans ClinVar/HGMD (ClnSig : « pathogenic » ou « likely\_pathogenic » ou « conflicting interpretation\_of pathogenicity ») même s'ils ne répondaient pas aux critères 1-3.

### 3 - Annotation des variants structuraux (SV dont CNV)

Les variants de structure (SV) étaient annotés *via* le logiciel AnnotSV v3.1. (Geoffroy *et al.*, 2018) comme décrit plus haut (Introduction. III.3.2.c) (Figure 56).

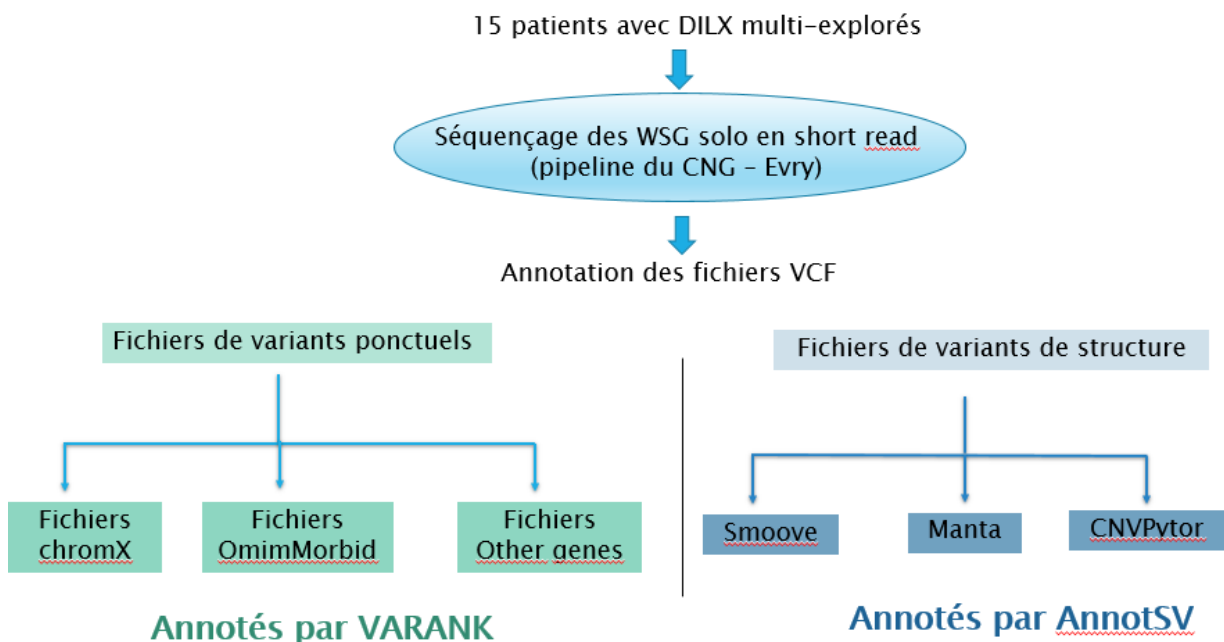


Figure 56: Annotation des variants issus du séquençage de génome

## 2. Résultats

Le travail sur les données de WGS a consisté dans un premier temps à rechercher des variants dans **les régions codantes** non identifiés par les analyses précédentes (défaut de couverture, erreur d'interprétation, gène non-encore décrit, région non incluse dans le cas de séquençage de panels de gènes) dans les gènes du chromosome X, puis dans les gènes OMIM. J'ai focalisé mon analyse sur les variants ayant un effet sur la séquence protéique et un potentiel effet sur l'épissage. La qualité des variants candidats était examinée et les *reads* étaient visualisés en utilisant le logiciel IGV (Robinson *et al.*, 2011, 2023), ce qui permettait de valider ou au contraire d'éliminer certains variants comme des faux positifs par exemple. Si un variant candidat était identifié, ce dernier était d'abord confirmé par une technique Sanger, puis les recherches ont été approfondies en réalisant une étude de ségrégation familiale, en contactant des équipes ayant travaillé sur le gène en question, parfois par le biais de GeneMatcher (Sobreira *et al.*, 2015), afin d'avancer sur son interprétation.

En l'absence de variant candidat identifié dans les régions codantes, il a fallu étudier les **régions non-codantes**: analyses des SNV dans les gènes du chromosome X, puis dans les gènes OMIM (variants introniques profonds ayant une bonne conservation nucléotidique et pouvant avoir un effet sur l'épissage ou la transcription, promoteur, éléments régulateurs), et analyse des variants de structure (déséquilibrées: CNV, équilibrées: inversions, translocations).

J'élargissais l'analyse au fur et à mesure aux gènes non encore décrits en DI mais pouvant représenter de bons candidats potentiels du fait de leur forte expression cérébrale, des interactions avec d'autres protéines codées par des gènes impliqués dans la DI, ou du rôle de la protéine concernée (ex: protéine synaptique etc.).

### *a) Analyse des variants situés dans les régions codantes*

#### **1- Application des différents filtres et stratégie d'interprétation**

Le sujet de mon travail étant la DI liée à l'X, je me suis d'abord et surtout focalisée sur le chromosome X en ne sélectionnant que les variants rares présents à l'état hémizygote chez les patients atteints.

J'ai pu observer que parmi les 15 individus séquencés, chaque patient portait en moyenne 8439 variants sur le chromosome X dont 846 variants rares. On observe donc que l'application de ces différents filtres a permis de réduire le nombre de variants d'environ 90 %. Il restait néanmoins environ 850 variants rare en moyenne par patient à analyser sur le chromosome X (Figure 57), ce qui souligne la difficulté de l'interprétation d'un génome.

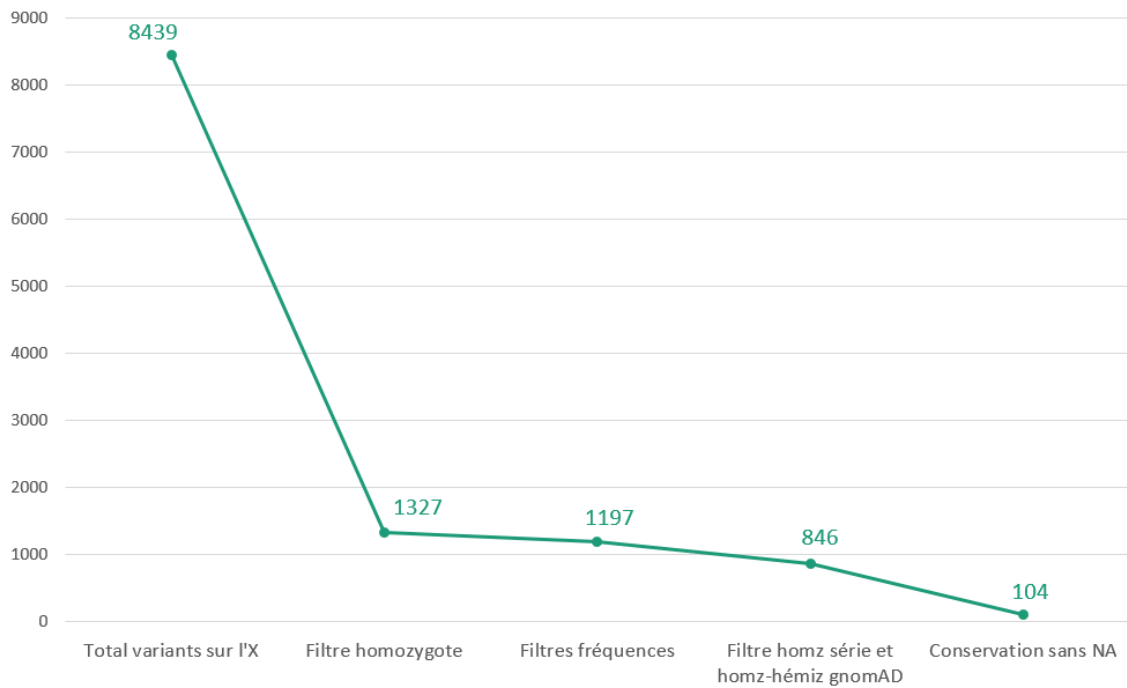


Figure 57: Nombre de variants rares par patient sur l'X après application de filtres

Dans un premier temps j'ai ciblé mon analyse sur les variants situés dans des gènes impliqués dans des pathologies rares avec TND référencées dans la base de données SysNDD (anciennement SysID) (Kochinke *et al.*, 2016) (Figure 58). Une fois cette analyse faite et si aucun variant prometteur n'était identifié, les autres gènes du X étaient analysés.

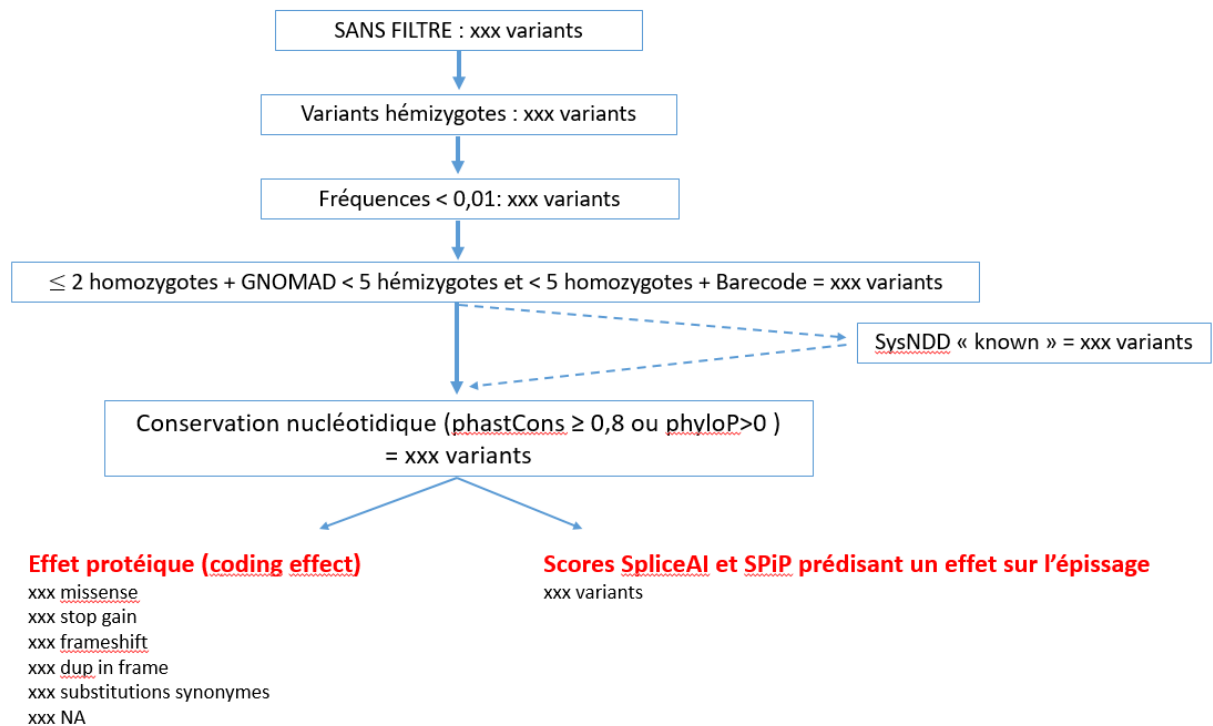


Figure 58: Stratégie d'interprétation des variants situés sur le chromosome X



J'ai ensuite étudié la distribution de ces variants selon leur localisation au sein des gènes et ai observé que seul 0,35 % des variants étaient localisés dans les régions codantes, 0,18 % et 0,42% étaient situés dans les régions 5'UTR ou 3'UTR, respectivement et 24,4 % étaient introniques. Le reste des variants (75 %) n'étaient pas annotés (NA) par VaRank, car situés dans des régions intergéniques, à proximité ou non des gènes (Figure 59).

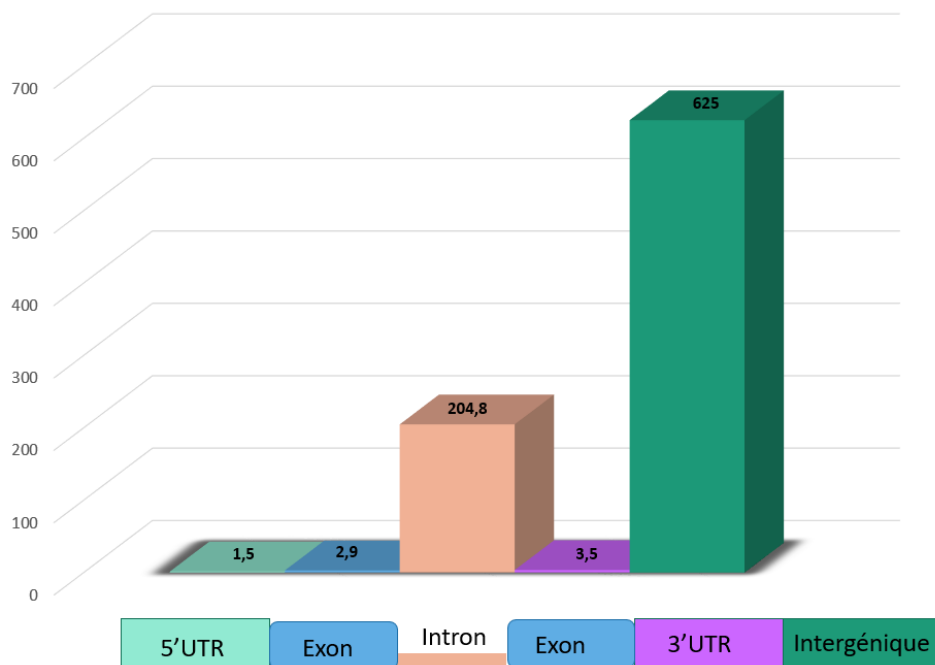


Figure 59: Distribution des variants rares de l'X selon leur localisation

La majorité des variants rares de l'X sont des substitutions (84 %), suivies par des délétions (8 %), des duplications (5 %) puis des insertions (3 %) (Figure 60).

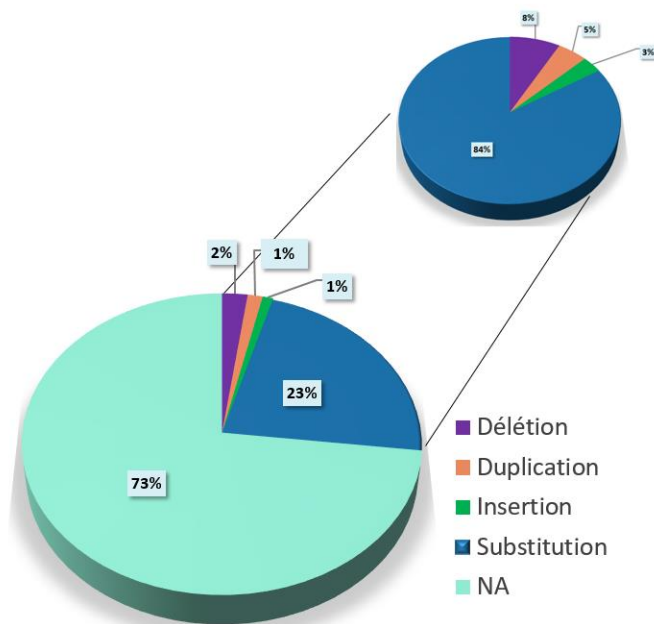


Figure 60: Répartition des variants de l'X selon leur type

Après cette étape, je me suis, dans un premier temps, intéressée aux variants situés dans les régions codantes, notamment les variants tronquants et non-synonymes, et à ceux affectant des sites canoniques d'épissage. Je n'ai pas identifié de variant stop ou *frameshift* sur l'X chez ces patients, mais il faut tenir compte du fait que ces derniers avaient été préalablement multi-investigués (panels de gènes de DI, exome). En revanche, j'ai identifié quelques variants faux-sens rares (1,13 variant faux sens en moyenne) (Figure 61).

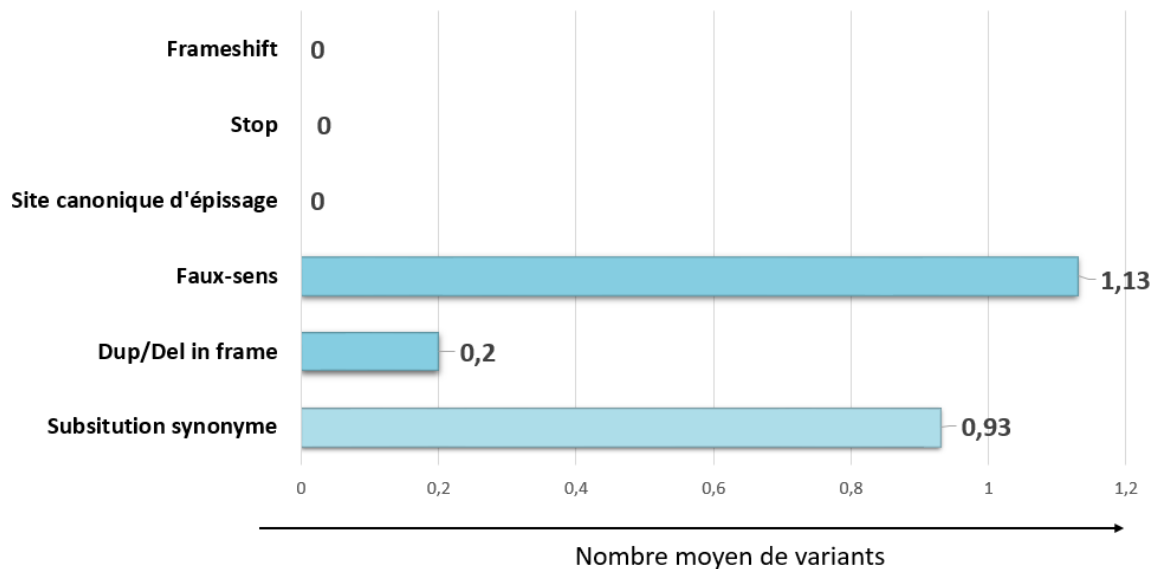


Figure 61: Représentation des variants codants selon l'effet protéique

Les variants n'affectant pas la séquence protéique étant très nombreux, j'ai utilisé des scores de conservation nucléotidique pour prioriser ceux qui pourraient être les plus prometteurs. Pour cela, j'ai utilisé en particulier les scores phyloP et phastCons, méthodes de prédiction se basant sur la conservation nucléotidique qui estiment la probabilité que chaque nucléotide appartienne à un élément conservé, sur la base de l'alignement multiple (phastCons  $\geq 0,8$ ) ou individuel (phyloP  $>0$ ) des génomes des espèces de vertébrés/mammifères (Pollard *et al.*, 2010; Siepel *et al.*, 2005). L'application combinée des deux filtres étant trop stringente et excluant un trop grand nombre de variants, dont certains étaient potentiellement intéressants, j'ai paramétré les filtres de manière à sélectionner les variants qui respectaient l'un ou l'autre de ces deux scores. On note malgré tout qu'en filtrant sur les scores de conservation (« NA » exclus) le nombre de variants passait de 846 à 104 par patient (Figure 57).

## 2- Les variants candidats identifiés

Nous n'avons pas identifié de SNV clairement pathogènes parmi les patients inclus dans l'étude. En revanche, plusieurs variants candidats ont été identifiés et seront détaillés dans ce chapitre, dont deux variants qui n'avaient pas été identifiés en exome.

### ❖ Variant candidat dans le gène *BCORL1*

Nous avons identifié chez le patient P5 (Pr Van Maldergem), un variant faux sens à l'état hémizygote (ChrX(GRCh37): g.129146999T>G, NM\_001184772.2 : c.251T>G, p.(Leu84Arg)) dans l'exon 3 du gène *BCORL1* (Xq26.1) hérité de sa mère asymptomatique et absent chez la grand-mère maternelle (ségrégation en cours chez le grand-père maternel). Le grand-oncle et l'oncle maternels de la mère du patient présentaient également une DI, avec TSA pour ce dernier, qui est décédé à l'âge de 20 ans d'une cause non-déterminée (Figure 62). Il n'a pas été possible de disposer de l'ADN de ces deux apparentés atteints.

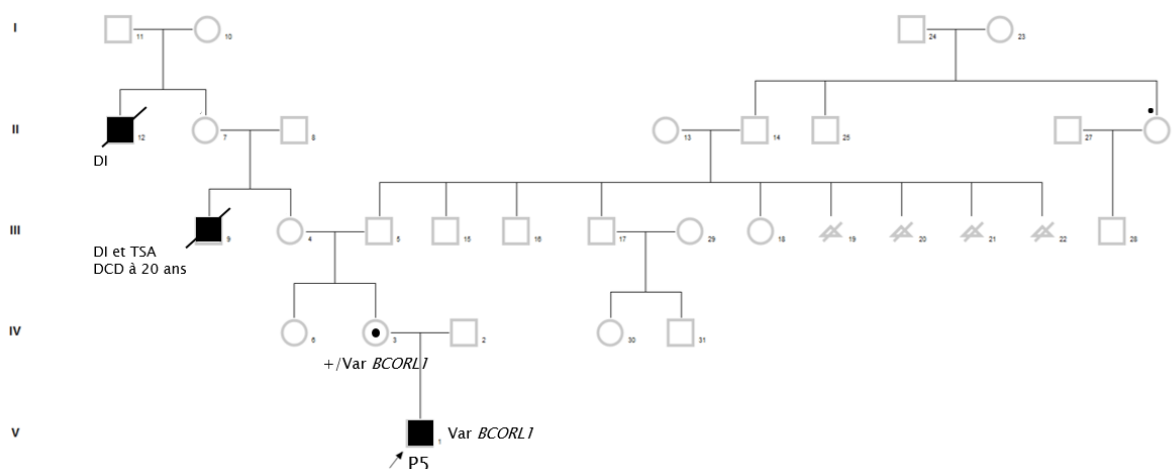


Figure 62: Arbre généalogique du patient P5.

Les symboles noirs correspondent aux patients présentant une DI. Une barre oblique signifie que l'individu est décédé.

Les signes cliniques et les examens complémentaires préalablement réalisés (négatifs) chez le patient sont résumés dans le Tableau 10.

**Tableau 10: Signes cliniques et résultats des examens complémentaires réalisés chez P5**

Signes cliniques	Examens complémentaires
PN 4280, TN 54 cm, PCN 36 cm Hypotonie précoce, marche à 4,5 ans, retard langage, Poids : 0 DS, Taille : -0,5 DS, PC : +0,5 DS Peu dysmorphique Syndrome cérébelleux (tremblement du chef, dysmétrie, ataxie cérébelleuse, strabisme convergent, voix rauque, troubles de la motricité fine)	CGH-array, X fragile, Panel DI 451 gènes Steinert, Gènes <i>ARX</i> et <i>MECP2</i> X-exome en 2017 Belgique Chromatographie des acides aminés plasmatiques, CPK IRM cérébrale (élargissement des espaces sous-arachnoïdiens)

Il s'agissait d'un variant rare, non retrouvé dans gnomAD, touchant une leucine bien conservée et qui n'avait pas été identifié lors du séquençage d'exome en 2017, vraisemblablement parce que le gène n'était pas encore identifié comme étant associé aux TND. Les outils de prédictions *in silico* étaient majoritairement en faveur d'un effet délétère :

Outils	Effet prédit
Align GVGD	65
SIFT	Deleterious (0)
PolyPhen-2	Deleterious
CADD	PHRED 23,5

Le gène *BCORL1* comporte 13 exons et code pour la protéine BCL-6 corepressor-like protein 1, constituée de 1711 acides aminés, qui est un corepresseur transcriptionnel. Elle contient un domaine CBD de liaison au CtBP (621-627), deux répétitions d'ankyrine en tandem, un signal NLS de localisation nucléaire, deux motifs LXXLL de recrutement de récepteurs nucléaires, et un domaine PUF (Figure 63). Ce gène a été impliqué à de nombreuses reprises dans le cancer (variations pathogènes somatiques) (Attia *et al.*, 2020; Coskun *et al.*, 2020) mais très peu dans les TND. En effet, trois publications rapportent au total 9 patients de sexe masculin issus de 5 familles présentant de manière variable un retard moteur, un retard de langage, une DI modérée à sévère, des troubles du comportement, un TDAH, des stéréotypies des mains, une épilepsie (4 patients sur 7), un strabisme et une dysmorphie non spécifique (Schuurs-Hoeijmakers *et al.*, 2013; Shukla *et al.*, 2019; Muthusamy *et al.*, 2021). Les mensurations étaient variables avec une tendance à l'insuffisance pondérale, le PC pouvant être normal ou inférieur à - 2 DS. Parmi les 3 patients ayant bénéficié d'une IRM cérébrale, un patient présentait une anomalie (atrophie cérébelleuse). Cinq variants pathogènes ont été rapportés et il s'agissait uniquement de faux sens (p.Asn820Ser, Val782Glu, Ser496Phe, Pro32Leu, p.Arg1265Cys). Aucun d'entre eux, tout comme le variant que nous avons identifié, ne touchaient les domaines CBD, PUF, NLS ou ANK de la protéine (Figure 63).

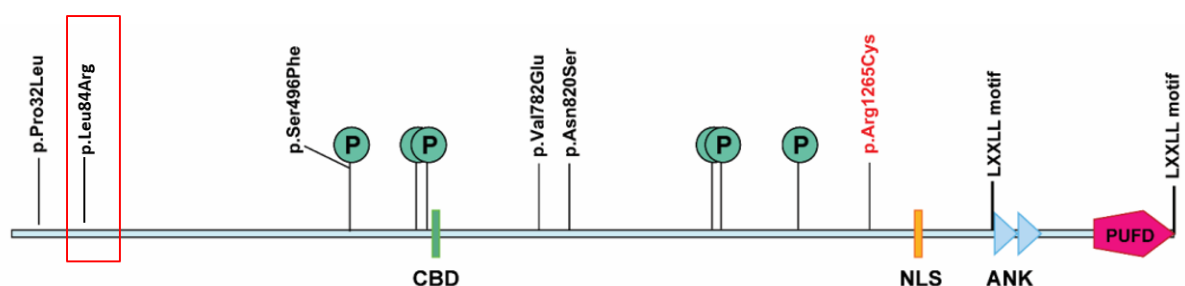


Figure 63: Variants hémizygotés potentiellement pathogènes rapportés dans *BCORL1*. Le variant p.(Leu84Arg) identifié chez le patient P5 est encadré en rouge (D'après Muthusamy *et al.*, 2021).

De manière intéressante, l'étude de l'inactivation de l'X réalisée chez la mère du patient montrait un biais important (4%-96%) et celle réalisée la grand-mère maternelle montrait une inactivation aléatoire (56%-44%).

Après un appel à collaboration *via* GeneMatcher pour tenter d'identifier d'autres patients porteurs de variants candidats dans ce gène et avancer dans son interprétation, le patient P5 a été inclus, avec l'accord de la famille, dans une cohorte qui était à ce moment-là en cours de constitution pour ce gène (Dr Noriko Miyake – Japon).

#### ❖ Variant candidat dans le gène *MED14*

Le séquençage de génome a permis d'identifier chez le patient P6 (Pr Isidor) un variant faux sens candidat dans le gène *MED14* (Xp11.4) à l'état hémizygote (ChrX(GRCh37):g.40586063G>C, NM\_004229.3:c.283C>G, p.(Leu95Val)). Ce variant était également présent chez son frère, atteint de DI et présentant un phénotype et une dysmorphie faciale très similaires, et chez leur mère asymptomatique (Figure 64). Deux oncles maternels de la mère, décédés, présentaient une DI pour laquelle le clinicien n'avait pas connaissance des détails cliniques. Nous n'avions pas d'ADN en banque pour ces derniers.

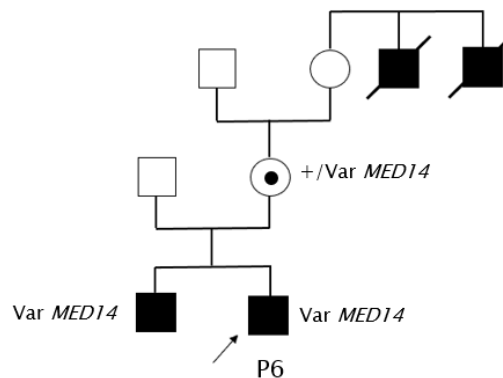


Figure 64: Arbre généalogique du patient P6.

Les signes cliniques et les examens complémentaires préalablement réalisés (négatifs) chez le patient sont résumés dans le Tableau 11.

**Tableau 11: Signes cliniques et résultats des examens complémentaires réalisés chez P6**

Signes cliniques	Examens complémentaires
Grossesse : Retard de croissance <i>in utero</i> Terme : 35 SA, PN : 1545 g, TN : 38 cm, PC : 28,4 cm. Hypospade, urétérocèle, CIV, cryptorchidie	CGH-array, X fragile, Gène <i>FGD1</i>

Retard langage, marche à 25 mois RGO sans diff alimentaire ni trouble du sommeil Poids : -1 DS, Taille : -1 DS, PC : -2,5 DS Brachydactylie, ongles dystrophiques Dysmorphie faciale (philtrum long, narines anteversées, arc de cupidon effacé)	Exome en 2013 au Japon IRM cérébrale (normale)
--	---

Il s'agissait d'un variant rare, non retrouvé dans gnomAD, touchant également une leucine très conservée. Les scores de prédiction *in silico* étaient en partie discordants, bien que plutôt en faveur d'un effet délétère :

Outils	Effet prédit
Align GVD	0
SIFT	Tolerated (0,11)
Polyphen	Deleterious
Mutation Taster	Disease causing (prob : 1)
CADD	PHRED 24,2

Le gène *MED14* n'est pas encore décrit en pathologie humaine à notre connaissance. Aucun variant pathogène et seize variants de signification incertaine sont rapportés dans ClinVar dans ce gène. Ce sont tous des variants faux-sens, qui sont situés tout le long de la protéine (Figure 65).

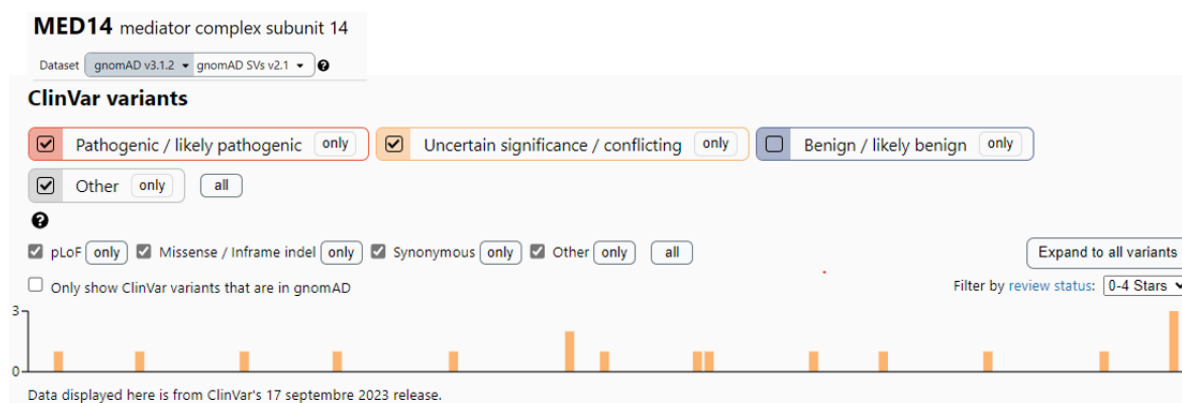


Figure 65: vue des variants du gène *MED14* répertoriés dans ClinVar (D'après www.gnomad-sg.org)

Par ailleurs, le variant du patient P6 était situé deux acides aminés plus loin qu'un variant rare faux-sens rapporté dans gnomAD une fois à l'état hétérozygote et jamais à l'état

hémizygote et ayant des scores de prédiction en faveur d'un effet délétère (ChrX(GRCh37):g.40586069G>A; c.277C>T; p.(Arg93Cys)).

Le gène *MED14* code pour une protéine qui est un composant clé du complexe Mediator, composé de très nombreuses protéines de la famille MED et agissant comme un co-activateur transcriptionnel (Figure 66). La protéine MED14, qui est nucléaire, est recrutée au niveau du promoteur *via* des interactions directes avec protéines régulatrices de la transcription et sert d'échafaudage pour l'assemblage d'un complexe de pré-initiation de la transcription avec l'ARN polymérase II et d'autres facteurs de transcription (complexe Mediator) (Cevher *et al.*, 2014). Il a été démontré que MED14 possède également un rôle clé dans la survie et le maintien des cellules souches embryonnaires pluripotentes induites, indépendamment de son rôle dans la transcription (Burrows *et al.*, 2015).

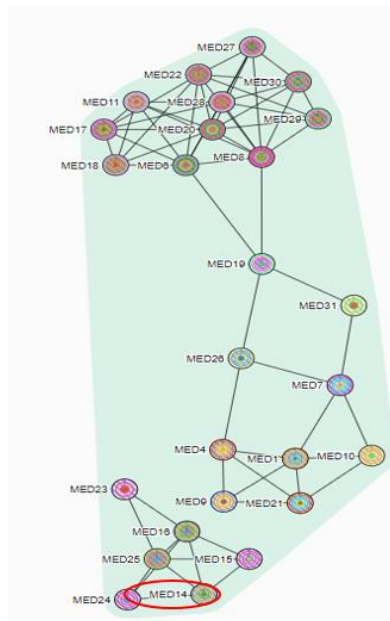


Figure 66: Complexe Mediator

Les protéines CCNC, CDK8 and CDC2L6/CDK11 n'apparaissent pas sur ce schéma. D'après Complex Portal - CPX-3227 (ebi.ac.uk).

*MED14* est un gène non encore décrit à ce jour en pathologie humaine mais dont la fonction et les interactions en font un bon gène candidat de DILX. En effet, il appartient à une grande famille comportant près d'une trentaine de gènes MED, dont plusieurs sont impliqués dans la DI comme les gènes *MED12* (Xq13.1) (syndromes de Lujan-Fryns syndrome (MIM 309520), d'Ohdo (MIM 300895) et d'Opitz-Kaveggia (MIM 305450), *MED13* (17q23.2) (Intellectual developmental disorder, autosomal dominant 61), *MED23* (6q23.2) (Mental retardation, autosomal recessive 18 (MIM 618009)), *MED25* (19q13.33) (syndrome de Basel-Vanagait-Smirin-Yosef ou eye-intellectual disability syndrome (MIM 616449), et *MED17* (11q21)

(Microcéphalie postnatale, atrophie cérébrale et épilepsie (MIM 613668)). Le gène *MED14* échapperait à l'inactivation de l'X (Yoshikawa *et al.*, 1998).

La protéine MED14 semble interagir avec plusieurs des protéines sus-citées (Figure 67). Elle est ubiquitaire et s'exprime dans de nombreux tissus, en particulier le muscle squelettique (<https://genevisible.com/tissues/HS/Gene%20Symbol/MED14>) (Yoshikawa *et al.*, 1998).

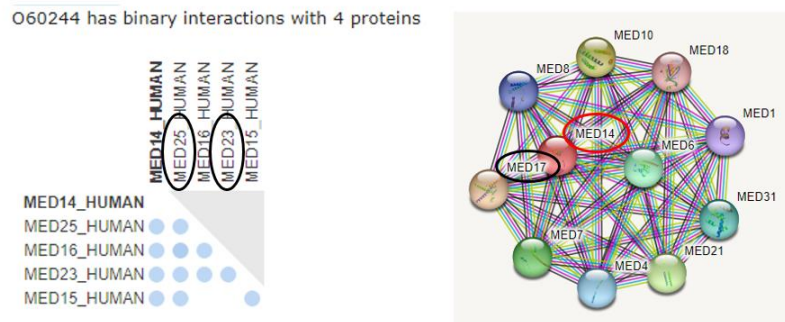


Figure 67: Interactions de la protéine MED14 avec les autres protéines de la famille MED

D'après MED14 - Mediator of RNA polymerase II transcription subunit 14 - Homo sapiens (Human) | UniProtKB | UniProt et MED14 protein (human) - STRING interaction network (string-db.org)

Suite à l'identification de ce variant, nous avons fait un appel à collaboration qui a permis d'identifier un autre patient nantais portant un variant faux-sens hémizygote dans ce gène. Une cohorte est en cours de constitution au niveau international (Dr Deshwar – Toronto), les 12 patients inclus jusqu'à présent ayant un phénotype extrêmement variable (myopathie, syndromes malformatifs au premier plan). Le patient américain porte un variant intronique hémizygote et l'étude en RNA-seq a montré l'absence de l'exon 19 dans certains transcrits aboutissant à un décalage du cadre de lecture et l'apparition d'un codon stop prématuré.

#### ❖ Variant candidat dans le gène *ARHGEF9*

Nous avons identifié chez le patient PH (Dr Demeer) un variant candidat faux-sens hémizygote dans le gène *ARHGEF9* (Xq11.1) (ChrX(GRCh37):g.62875430C>G, NM\_001353923.1:c.1283G>C, p.(Arg428Pro)), hérité de la mère et de la grand-mère maternelle. La mère a quitté le collège à l'âge de 14 ans puis a obtenu un CAP, un BEP et un baccalauréat professionnel. Elle est décrite comme ayant de discrètes particularités morphologiques (lèvre supérieure fine, sourcils horizontaux) retrouvés chez ses apparentés atteints de DI. Ce variant était également présent chez les deux oncles maternels du patient présentant une DI et absent chez la sœur du patient qui présentait une épilepsie isolée. La mère, la grand-mère maternelle et la sœur ne présentait pas de biais d'inactivation de l'X (Figure 68). Un troisième oncle maternel présentait un retard de développement global puis une DI avec épilepsie ayant débutée à l'âge de 9 mois. Il n'a jamais acquis le langage oral et est décédé au décours d'un malaise à l'âge adulte dont la cause n'est pas connue.



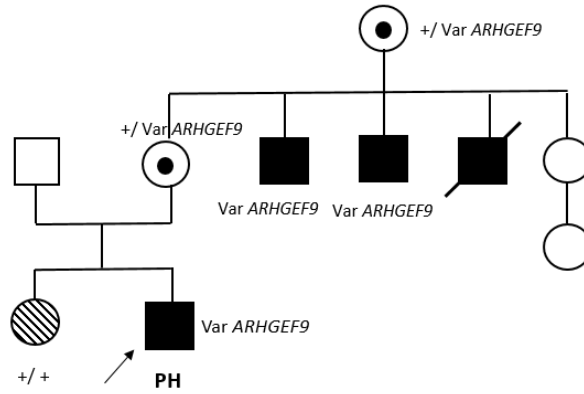


Figure 68: Arbre généalogique du patient PH.  
Les hachures correspondent à la présence d'une épilepsie.

Les signes cliniques et les examens complémentaires préalablement réalisés (négatifs) chez le patient PH sont résumés dans le Tableau 12.

**Tableau 12: Signes cliniques présentés par le patient PH et ses oncles maternels**

Signes cliniques	Ex. complémentaires
<p>Patient PH</p> <p>Grossesse sans particularité Terme : 40 SA, PN : 3720 g, TN : 50 cm, PC : 36 cm. Absence de langage, assis à 12 mois, marche à 20 mois QIT = 59 homogène Trois crises convulsives fébriles Poids : +2 DS, Taille : +2,5 DS, PC : +2 DS Agitation motrice, stéréotypies, difficultés en motricité globale Dysmorphie faciale (front haut, fentes palpébrales vers le bas et le dehors, lèvre supérieure en chapeau de gendarme) Pas d'IRM cérébrale réalisée</p>	<p>CGH-array, X fragile, Gène <i>FGD1</i> Exome en 2013 au Japon IRM cérébrale (normale)</p>
<p><u>Oncle maternel 1 :</u> Epilepsie : début à 9 mois Absence de langage, DI, Scolarisé en IME Décédé au décours d'un malaise</p> <p><u>Oncle maternel 2 :</u> Grossesse normale, poids : 4150 g, taille : 50,5 cm. Marche à 24 mois, retard de langage, DI, scolarisé en IME (écriture et lecture non-acquises), Dysmorphie : cou large, rétrognathisme, lèvre supérieure fine Poids 65,8 kg, Taille 180 cm, PC 56 cm (- 0,5 DS).</p> <p><u>Oncle maternel 3 :</u> Grossesse normale, poids : 3,5 kg, taille : 51 cm, Marche à 15 mois, retard de langage, SEGPA puis IME, CAP horticulture. A acquis la lecture et l'écriture. Dysmorphie : visage allongé, lèvre supérieure fine, enophtalmie, sourcils horizontaux, cou large, kyste branchial Poids 66 kg, Taille 182,5 cm, PC 55,5 cm (- 0,8 DS).</p>	

Ce variant avait été identifié préalablement par panel chez le patient PH en 2016, mais son interprétation restait délicate du fait de l'absence du variant chez la sœur et de l'absence de biais d'inactivation chez la mère et la grand-mère porteuses du variant. Il s'agissait d'un variant rare absent de gnomAD, touchant une Arginine relativement bien conservée et dont les scores de prédiction étaient majoritairement en faveur d'un effet pathogène, ce qui s'est confirmé plusieurs années plus tard :

Outils	Effet prédit
Align GVDG	C65
SIFT	Deleterious (0,00)
Polyphen	Deleterious
Mutation Taster	Uncertain
CADD	PHRED 29,1

Le gène *ARHGEF9* comporte 10 exons et code pour la collybistine, un facteur d'échange de nucléotides (GEF) spécifique du cerveau qui appartient à une famille de Rho-like GTPases qui passent de l'état actif lié au GTP à l'état inactif lié au GDP. La collybistine joue un rôle essentiel dans la formation de groupes de récepteurs postsynaptiques à la glycine et au GABA au niveau de certaines synapses inhibitrices et est également connue pour affecter la plasticité et la polarité du cytosquelette d'actine et pour modifier les voies de transduction du signal intracellulaire (inhibition de la voie mTORC1) (Hall et Nobes, 2000; Harvey *et al.*, 2004; Machado *et al.*, 2016).

Des variants pathogènes hémizygotés du gène *ARHGEF9*, incluant des SNV et des grandes délétions, ont été identifiés chez des patients atteints d'un TND lié à l'X (MIM 300607) caractérisé par un phénotype variable, allant de la DI modérée sans épilepsie à l'encéphalopathie épileptique, pouvant s'associer à des éléments dysmorphiques, une hyperekplexia et de l'anxiété (Alber *et al.*, 2017; Shimojima *et al.*, 2011). A ce jour, 29 variants pathogènes ont été publiés dans ce gène chez 28 garçons atteints et 14 filles atteintes. Le profil d'inactivation de l'X était biaisé chez 7 filles et aléatoire chez 6 filles (Nashabat *et al.*, 2019; Yao *et al.*, 2020; Scala *et al.*, 2021; Ghesh *et al.*, 2021).

Chez les garçons, ces variants consistaient en 6 variants perte de fonction (2 non-sens, 1 variant de *splice* et 3 grandes délétions) et 10 variants faux-sens (Tableau 13).

**Tableau 13: Récapitulatif des variants pathogènes du gène ARHGEF9 publiés à ce jour**

Garçons atteints	Filles atteintes
10 faux-sens (8 H, 2 dn) 2 non-sens (1 H, 1 dn) 1 variant de <i>splice</i> (H) 3 grandes délétions (3 dn)	2 faux-sens (2 dn) 1 non-sens (faible mosaïque chez le père) 5 grandes délétions (5 dn) 3 réarrangements chromosomiques complexes (une inversion, deux translocations X;autosome) (3 dn) 2 délétions intragéniques (2 dn) 1 variant de <i>splice</i> (dn)

H : hérité, dn : *de novo*

Ce variant est d'autant plus intéressant qu'il ségrège dans cette famille chez les oncles maternels atteints de DI. Par ailleurs, il est situé dans le domaine d'homologie à la pleckstrine (PH) au sein duquel plusieurs variants faux-sens pathogènes ont déjà été rapportés (dont trois touchent une arginine), dont le variant R357I rapporté comme altérant l'accumulation sous-membranaire de géphyrine médiée par ARHGEF9 (Yao *et al.*, 2020) et le variant Arg338Gln rapporté par Long et coll. comme affectant la fonction de ARHGEF9 et sa liaison à PI3P (Long *et al.*, 2016) (Figure 69). Dans ClinVar, 9 variants faux sens pathogènes ou probablement pathogènes sont répertoriés.

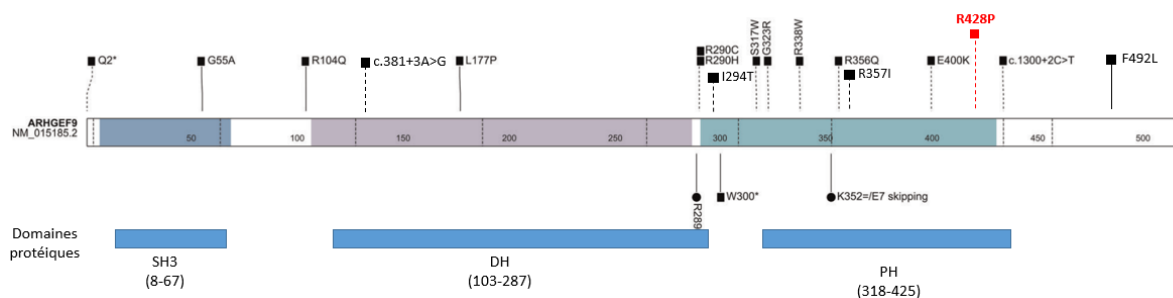


Figure 69: Schéma de la protéine ARHGEF9 et des variants pathogènes publiés.

Les ronds représentent les filles et les carrés les garçons. Les traits pointillés représentent les variants hérités et les traits pleins les variants *de novo*. Le variant identifié chez le patient PH apparaît en rouge (Adapté à partir de Gshesh 2021).

Le phénotype des patients porteurs de variant pathogènes du gène *ARHGEF9* étant très variable, bien que plusieurs arguments penchent en faveur du caractère délétère de notre variant, il n'a pas été possible de conclure de manière certaine sur sa pathogénicité. Des études fonctionnelles sont envisagées pour avancer sur son interprétation et prouver son implication dans le phénotype du patient PH et ses apparentés (J.W. Um, Corée).

Pour les autres patients, aucun variant rare non-synonyme sur l'X n'a pu être identifié. J'ai donc, chez ces derniers, élargi mon analyse aux gènes situés sur les autosomes. Plusieurs autres variants candidats ont ainsi été identifiés dans des gènes de DI au sein de la présente cohorte mais aucun n'a pu être considéré comme étant clairement pathogène : On peut citer par exemple :

- Un variant candidat faux-sens hétérozygote dans le gène *DYRK1A*, *de novo*, identifié préalablement en panel chez le patient PE. La recherche de signature mise au point pour ce gène (Courraud *et al.*, 2021) était négative et le patient ne présente pas les signes cliniques caractéristiques du syndrome DYRK1A (Dr A Piton – IGBMC).
- Un variant candidat faux-sens hétérozygote dans le gène *ZBTB18* (1q44) a été identifié et semblait prometteur, mais l'étude parentale a montré qu'il était hérité de la mère asymptomatique.

### *b) Analyse des variants pouvant avoir un effet sur l'épissage*

N'ayant identifié aucun variant non-synonyme convaincant pour la plupart des familles, j'ai entrepris d'étudier les variants pouvant avoir un effet sur l'épissage au-delà d'un effet sur les sites canoniques en utilisant les outils SpliceAI (Jaganathan *et al.*, 2019) et SPiP (Splicing Prediction Pipeline) (Leman *et al.*, 2022). J'ai ainsi pu noter que seuls deux variants étaient prédits pour avoir un effet sur l'épissage dans les gènes *COL4A5* et *ARL13A* chez deux patients (Tableau 14).

Le gène *COL4A5* est impliqué dans le syndrome d'Alport lié à l'X, pathologie associant une atteinte rénale, auditive et ophtalmologique (Knebelmann *et al.*, 1996). Le gène *ARL13A* n'est pas décrit en maladie humaine contrairement à son paralogue, le gène *ARL13B* impliqué dans le syndrome de Joubert (autosomique récessif) (Cantagrel *et al.*, 2008). La protéine Arl13a est une GTPase dont le rôle exact n'est pas connu. Chez le poisson zèbre, il a été démontré que Arl13a et Arl13b se localisent tous deux sur les microtubules des cellules ciliées et les cellules en division de l'embryon mais que les gènes ont évolué vers des rôles différents, arl13a n'ayant pas de fonction dans la rétine du poisson zèbre (Song et Perkins, 2018). La pLI étant proche de 0 pour le gène *ARL13A* et le phénotype du patient PN n'étant pas cohérent avec un syndrome d'Alport même atypique, nous n'avons pas initié d'études fonctionnelles pour ces variants (Tableau 14, Figure 70).

**Tableau 14: scores de prédictions des deux variants identifiés dans les gènes *COL4A5* et *ARL13A***

	<i>Variant COL4A5</i>	<i>Variant ARL13A</i>
Patient	PN	PB
pLI	0,99	0,00059
Localisation	Intron 5	Intron 3
Variant	ChrX(GRCh38):g.108568718A>C NM_033380.3: c.322-41A>C, p.?	ChrX(GRCh38):g.100985553G>A NM_001162491.1: c.131-114G>A, p.?
GnomAD	0,0034%	0,0046%
Local splice effect (VaRank)	Cryptic Acceptor Strongly Activated	Cryptic Acceptor Strongly Activated

SPiP		<b>0,680</b>	<b>0</b>
SPLICEAI	DS_AG :0, DS_AL :0, DS_DG :0, DS_DL :0		DS_AG :0,31, DS_AL : <b>0,55</b> , DS_DG :0, DS_DL :0
PhyloP		0,435	3,27
PhastCons		0	0,866

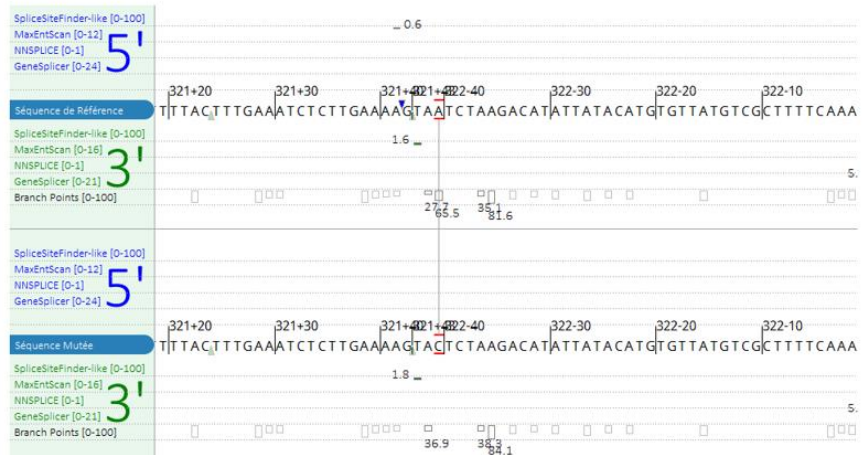
Alamut Visual Plus Browser

Gene: COL4A5 - Transcript: NM\_033380.3 - Variant: c.322-41A>C

**Splicing Predictions**

Splicing predictions at nearest natural junction:

No predicted impact.



Gene: ARL13A - Transcript: NM\_001162491.1 - Variant: c.131-114G>A

**Splicing Predictions**

Splicing predictions at nearest natural junction:

No predicted impact.

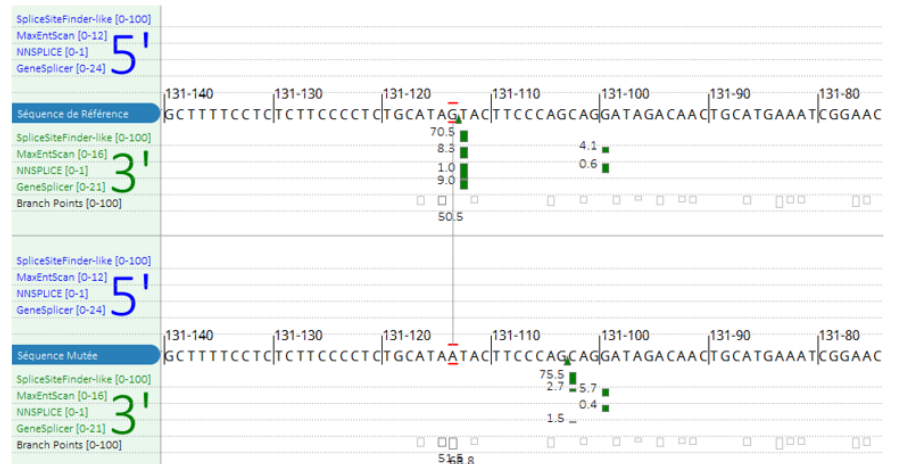


Figure 70: Prédications d'effet sur l'épissage des variants des gènes COL4A5 et ARL13A (logiciel Alamut Visual Plus)

### c) Analyse des variants structuraux

L'analyse sur le chromosome X a montré que chaque patient portait en moyenne 339 SV (incluant 152 délétions, 21 duplications, 71 insertions, 24 inversions, 71 translocations). La taille des délétions était en moyenne de 48 kb et celle des duplications de 837 kb.

J'ai pu noter qu'il y avait peu de chevauchement entre les différents programmes d'identification de CNV (Manta, Smoove et CNVpytor), puisque peu de CNV étaient identifiés de manière commune par deux outils : en moyenne 12 CNV par patient, 6,6 délétions (dont 1,4 contiennent un ou plusieurs gènes) et 5,4 duplications (dont 3,5 contiennent 1 ou plusieurs gènes), et encore moins par les trois (6 CNV sur toute la cohorte), nous conduisant à penser qu'il existe très probablement de nombreux faux-positifs.

Cette complexité d'interprétation liée au non-chevauchement des différents outils est bien illustrée par le diagramme de Venn ci-dessous concernant le patient P7 (Figure 71), où l'on voit qu'un seul CNV est identifié à la fois par les trois outils (une délétion de 11 kb ne comportant pas de gènes), et 9 CNV par deux outils (Manta et Smoove).

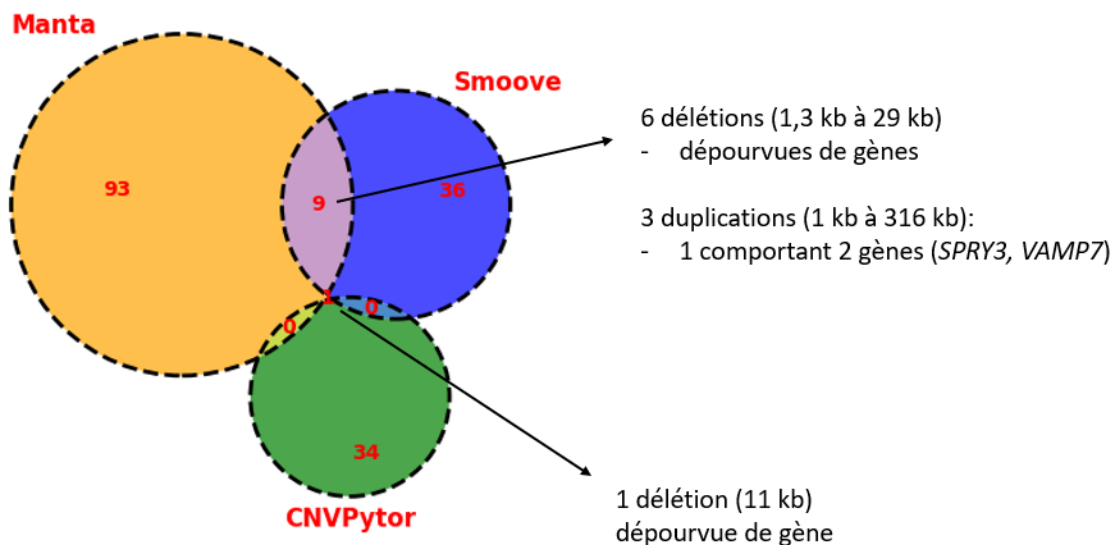


Figure 71: Diagramme de Venn illustrant le nombre de CNV identifiés par les outils Manta, Smoove et CNVpytor chez le patient P7

- ✚ Pour l'ensemble de la cohorte, seuls 6 CNV ont été identifiés par les 3 outils (5 duplications et 1 délétion, chez des patients différents), 2 CNV comportant des gènes :
  - Chez le patient PG : une duplication de la région Xq23 d'une taille de 196 290 pb contenant 3 gènes non référencés en pathologie humaine (*IL13RA2*, *LRCH2* et *SNORA35B*).

- Chez le patient PF : une duplication interstitielle en tandem de la région Xp11.22 d'une taille de 992 539 pb (X :53045501-54038043) contenant 15 gènes, dont les gènes morbides *HSD17B10*, *HUWE1*, *IQSEC2*, *KDM5C*, *PHF8*, et *SMC1A*, impliqués dans la DI (Figure 72).

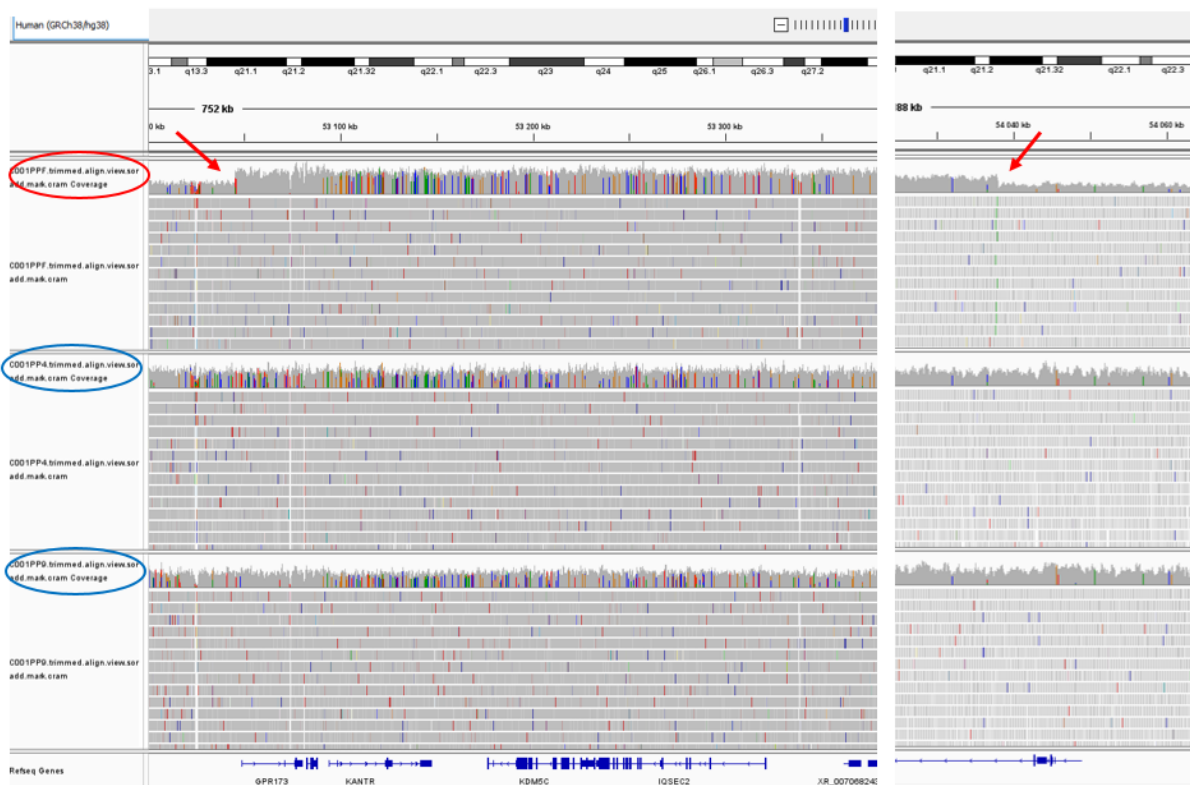


Figure 72: Visualisation sur IGV de la duplication Xp11.22 identifiée chez PF  
 Visualisation de la duplication Xp11.22 (flèches) (X:53045501-54038043) (GRCh38) détectée chez PF (ellipse rouge), absente chez les autres patients (ellipses bleues). L'alignement des *reads* est affichés par paires (Thorvaldsdottir *et al.*, 2013). Visualisation des gènes *IQSEC2*, *KDM5C* et *PHF8* inclus dans la duplication.

Le patient et son frère sont actuellement âgés de 39 ans et 35 ans, respectivement et présentent tous deux une DI sévère et des troubles du comportement qui s'aggravent progressivement avec le temps. Ils ne présentent pas d'épilepsie, tout comme leur oncle et leur cousin maternels atteints, décédés tous les deux (tableau neurodégénératif décrit chez l'oncle (Figure 73). Morphologiquement, PF et son frère présentent une dysmorphie faciale très similaire associant un hypertélorisme, une ensellure nasale haute, un synophris, un visage hypotonique avec un prognathisme et des lèvres épaisses. Ils présentent également une atrophie testiculaire, des mamelons petits et ombiliqués, une répartition tronculaire des graisses (tablier adipeux abdominal) et des hallux proéminents.

Le patient PF avait bénéficié d'un « X-exome » qui était négatif mais n'avait pas bénéficié de CGH-array.

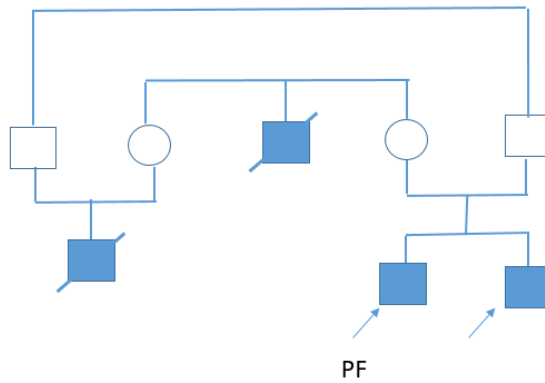


Figure 73: Arbre de la famille du patient PF

Ce gain de copie inclue la région correspondant à un CNV récurrent pathogène responsable du syndrome de duplication Xp11.22 (MIM 300705) (Figure 74) qui se caractérise par la présence d'une DI modérée à sévère, un retard de langage, une possible épilepsie, des troubles neuro-psychiatriques et des particularités morphologiques retrouvées chez PF et son frère atteint et incluant de manière variable un hypertélorisme, un synophris, une base de nez large, des lèvres épaisses avec visage hypotonique et bouche ouverte et des hallux longs (Bonnet *et al.*, 2006; Santos-Reboucas *et al.*, 2015; Grams *et al.*, 2016; Moey *et al.*, 2016; Giorda *et al.*, 2009) (Figure 75). Une atrophie testiculaire, une répartition tronculaire des graisses et des mamelons ombiliqués sont également décrits dans la littérature chez certains patient (Santos-Reboucas *et al.*, 2015; Grams *et al.*, 2016; Q. Wang *et al.*, 2020).

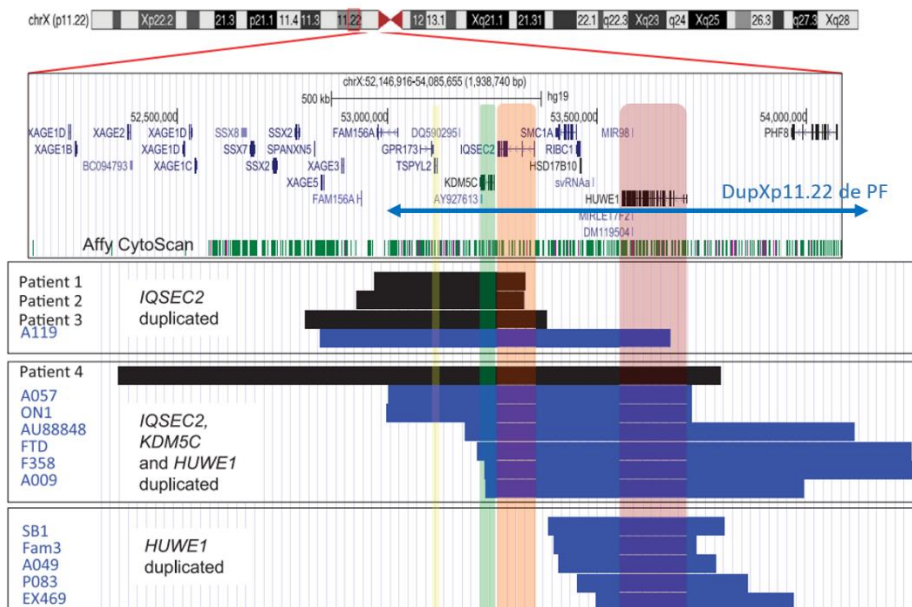


Figure 74: Duplication Xp11.22 identifiée chez PF et les autres patients de la littérature

La flèche bleue indique la duplication de PF, les barres noires et bleues correspondent aux duplications de patients rapportés dans la littérature (d'après Moey 2016).





Figure 75: Phénotype des patients porteurs d'une duplication Xp11.22  
 Les patients présentent des particularités morphologiques incluant de manière variable un synophris et des sourcils arqués, un visage hypotonique, un nez proéminent (d'après Giorda *et al.*, 2009)

Nous avons pu démontrer que cette duplication survient en tandem chez PF, comme cela est décrit chez les patients atteints de la littérature (Santos-Reboucas *et al.*, 2015) (Figure 76).

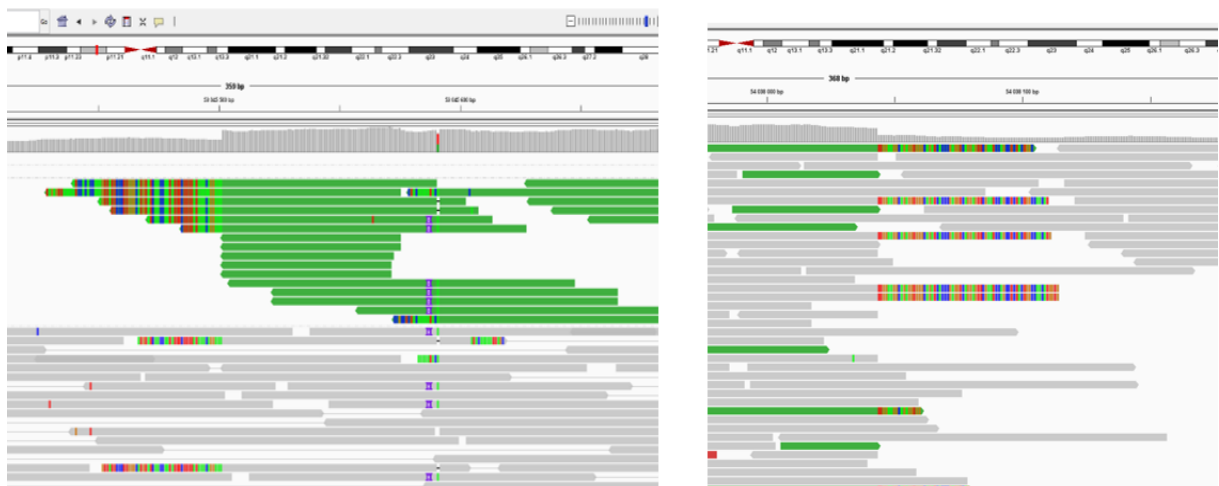


Figure 76: Mise en évidence sur IGV des paires de *reads* aux points de cassures  
 Alignement des *reads* en mode *soft clipped bases* montrant l'aspect typique (paires de *reads* pointant dans la direction opposée) pour les duplications en tandem (<https://igv.org/doc/desktop/#UserGuide>).

Nous avons également pu confirmer en qPCR la présence de cette duplication chez PF ainsi que chez son frère et son cousin germain maternel atteints (Dr Scheidecker) (Figure 73, Figure 77). Malheureusement, l'ADN de la mère et de la tante maternelle asymptomatiques, ainsi que de l'oncle paternel atteint et décédé n'étant pas disponible, nous n'avons pas pu faire cette recherche chez ces derniers.

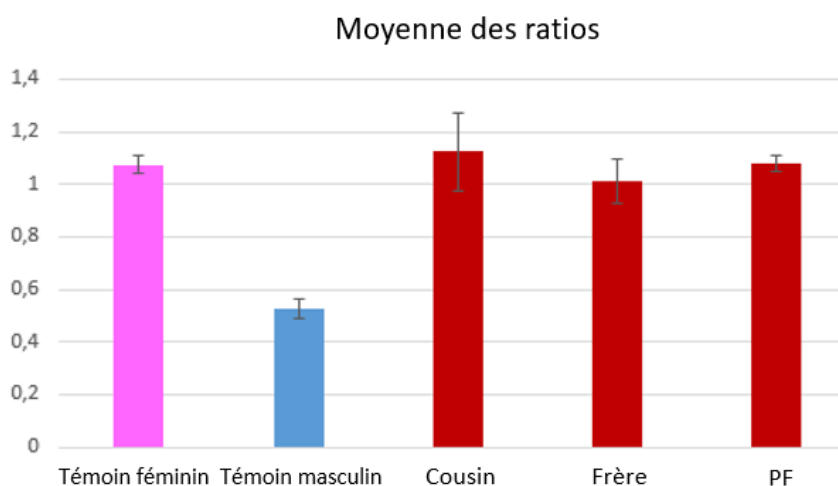


Figure 77: Ratios obtenus en PCR quantitative pour PF, son frère et son cousin maternel atteints concernant la région Xp11.22

Présence de 2 copies chez le témoin féminin et d'une copie chez le témoin masculin. On observe la présence de 2 copies chez les 3 individus de sexe masculin atteints de DI permettant de confirmer la présence de la duplication Xp11.22 chez ces derniers. Les positions des amorces utilisées sont 53045501-54038043 (GRCh38) (kit SyberGreen Mastermix Quantitect (Qiagen)). Méthode par quantification absolue. Moyenne calculée à partir de 2 ratios (gènes de référence *GAPDH* et *RPPH1*).

- ✚ **Concernant les CNV identifiés par au moins deux outils** j'ai pu recenser 12 CNV en moyenne par patient de type délétion (6,6) ou duplication (5,4), dont 1,4 délétions et 3,5 duplications qui contiennent un ou plusieurs gènes.

Parmi ces derniers, 9 délétions et 11 duplications étaient annotées par AnnotSV comme étant de classe 5 selon la classification ACMG/ClinGen.

Parmi les délétions, on notait :

- Une délétion Xq28 de 1,2 Mb (33 gènes) détectée chez PD (X 154598418-155803864). La vue d'IGV ne retrouvait pas un profil en faveur d'une délétion au niveau de la borne proximale et était hétérogène, comme chez les autres patients. La vue de la région de la borne distale semblait compatible avec une délétion mais était présente chez les autres patients (Figure 78).
- Trois grandes délétions retrouvées chez plusieurs patients, correspondant donc vraisemblablement à des faux-positifs :
  - Une délétion Xp11.23p11.22 de 4,2 Mb (100 gènes), détectée chez P5, P6, P8, et PH.
  - Une délétion Xq21.31q12 de 22,5 Mb (136 gènes), détectée chez P6 et PI.
  - Une délétion Xq21.1q13.1 de 8,7 Mb (98 gènes), détectée chez P9 et PE.

Parmi les 11 duplications on notait :

- La duplication Xp11.22 pathogène identifiée chez PF.
- Trois grandes duplications retrouvées chez plusieurs patients et présentant les mêmes bornes que les 3 délétions précédentes :

- Une duplication Xp11.23p11.22 de 4,2 Mb (100 gènes), détectée chez P6 et PH.
- Une duplication Xq21.31q12 de 22,5 Mb (136 gènes), détectée chez PF et PI.
- Une duplication Xq21.1q13.1 de 8,7 Mb (98 gènes), détectée chez P6, P9, PB, PE, PI

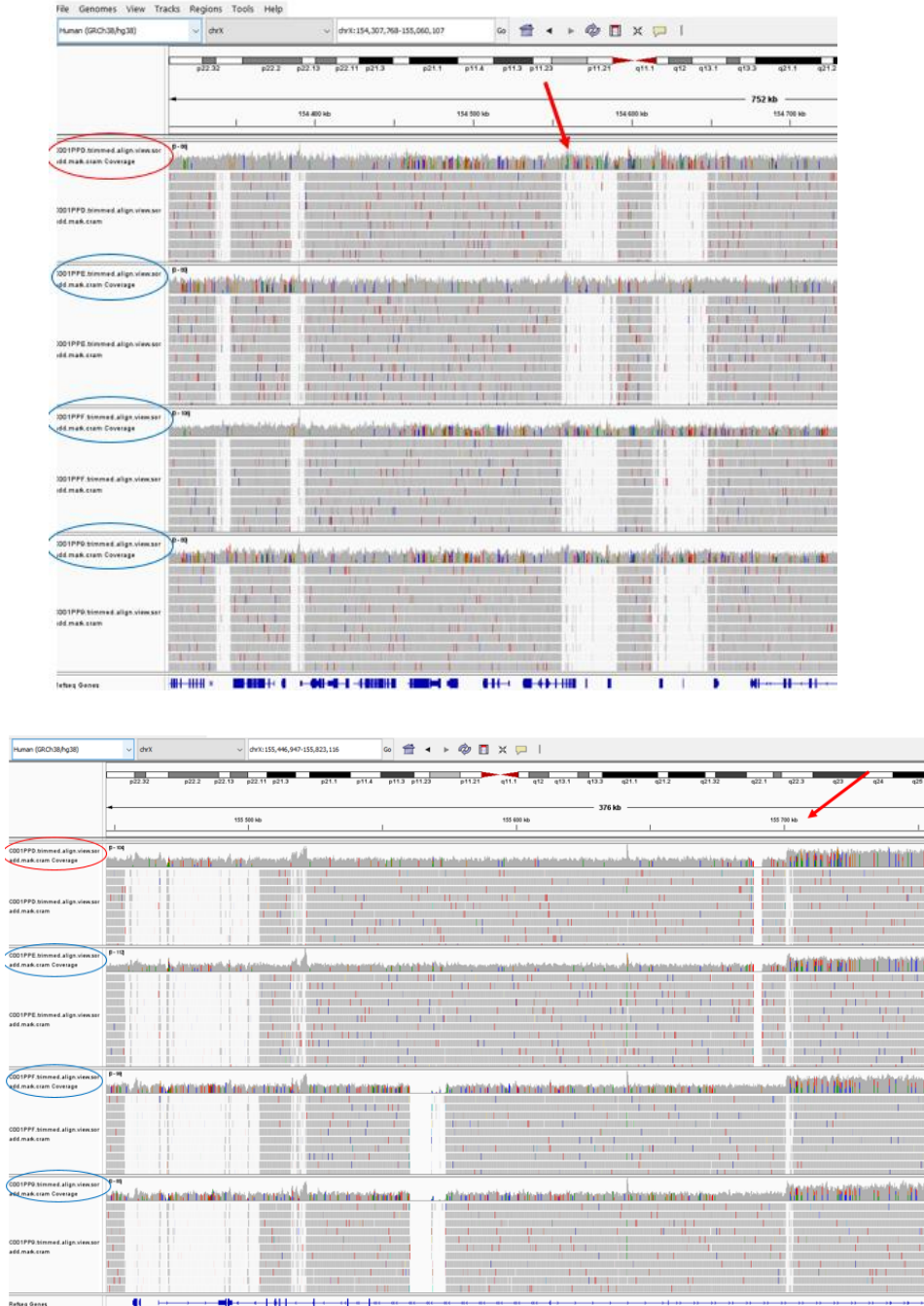


Figure 78: Vue IGV de la délétion Xq28 de 1,2 Mb détectée chez le patient PD

Visualisation des bornes proximale (haut, flèche) et distale (bas, flèche) (X : 154598418-155803864) de la délétion Xq28 détectée chez PD (ellipse rouge) et des profils d'autres patients (ellipses bleues).

Je poursuis actuellement l'analyse des CNV prédits par AnnotSV comme étant de classe 3 ACMG.

## 2. Patients suivis dans le cadre de ma pratique clinique et ayant bénéficié d'un WGS en *trio*

Dans le cadre de ma pratique clinique, j'ai été amenée à suivre plusieurs patients ayant bénéficié récemment d'un séquençage de génome entier à but diagnostique.

Parmi ces derniers, je me suis intéressée à analyser les résultats qui concernaient des variants pathogènes identifiés dans des gènes du chromosome X, chez huit patients.

De manière intéressante, sur ces 8 variants de l'X (5 garçons et 3 filles), 6 sont survenus *de novo* (chez 3 garçons et 3 filles) et 2 sont hérités de mères asymptomatiques.

### **Les variants pathogènes *de novo* sur l'X:**

En ce qui concerne les filles, il s'agit de SNV hétérozygotes pathogènes *de novo* dans les gènes *CASK* (Intellectual developmental disorder and microcephaly with pontine and cerebellar hypoplasia, OMIM 300749) (variant tronquant - *frame shift*), *MECP2* (syndrome de Rett, OMIM 312750) (variant faux-sens), et *IL1RAPL1* (Intellectual developmental disorder, X-linked 21, OMIM 300143) (variant tronquant – *frame-shift*).

En ce qui concerne les garçons, les variants pathogènes *de novo* consistent en:

- Un CNV hémizygote : délétion de la région Xp22.12p22.112 de 1657 pb emportant l'exon 3 du gène *CNKSR2* (Intellectual developmental disorder, X-linked syndromic, Hougé type, OMIM 301008), entraînant un décalage du cadre de lecture et l'apparition d'un codon STOP prématuré.
- Deux SNV hémizygotes : un variant faux-sens dans le gène *HUWE1* (Intellectual developmental disorder, X-linked syndromic, Turner type, OMIM 309590), et un variant non-sens dans *CNKSR2*.

### **Les variants pathogènes hérités d'une mère asymptomatique :**

Ces derniers ont été identifiés chez des garçons. Il s'agit d'un variant faux-sens dans le gène *BRWD3* (Intellectual developmental disorder, X-linked 93, OMIM 300659) et d'une duplication récurrente de 24 pb responsable d'une insertion de 8 alanines dans le gène *ARX* (syndrome de Partington, OMIM 309510). Il s'agit de deux patients pour lesquels il n'existe aucune histoire familiale de DI dans la branche maternelle. Cet exemple très concret illustre très bien la difficulté à reconnaître et identifier les familles avec DILX, en l'absence d'histoire familiale, cette dernière étant peu fréquente en réalité.

En ce qui concerne les variants de classe 3 sur l'X (plusieurs patients portent un second variant de classe 3 sur un gène autosomique), nous avons pu réaliser une étude de ségrégation familiale concernant un patient porteur d'un variant de classe 3 dans le gène *MECP2*, qui nous a permis de reclasser ce variant et d'exclure sa pathogénicité. D'autres études familiales de ce type seraient nécessaires pour avancer sur l'interprétation des autres variants de classe 3 mais elles sont malheureusement rarement réalisables en pratique (grands parents décédés, pas d'hommes atteints dans la branche maternelle, pas d'hommes sains disponibles pour être testés (par exemple si la mère n'a pas de frère ou n'a plus de contact avec ses apparentés pour diverses raisons).

## IV. Quatrième partie

En parallèle de cela et afin de me former à l'analyse des données de séquençage à haut débit et mieux comprendre les spécificités de l'analyse d'ES et de WGS, j'ai entrepris et/ou participé à différents travaux qui seront présentés ci-dessous:

### 1. L'interprétation de séquençage d'exome de patients atteints de DI suivis dans le service de génétique médicale des HUS.

Il s'agissait de patients ayant bénéficié d'un ES à titre diagnostique après de nombreux examens complémentaires négatifs. Parmi ces patients nous pouvons citer le cas d'une patiente atteinte d'une maladie très rare associant un retard de croissance intra-utérin (RCIU) sévère non-rattrapé (taille adulte d'1,47 m), une dysmorphie faciale incluant un microrétrognathisme sévère et une histoire obstétricale marquée par 9 épisodes de fausses-couches. L'examen fœtopathologique réalisé pour deux fœtus avait mis en évidence un RCIU sévère avec rétrognathisme, raccourcissement des os longs, fente palatine et anomalies des organes génitaux externes. L'examen placentaire révélait une insuffisance placentaire. L'ES réalisé chez la patiente a mis en évidence un variant *frameshift* hétérozygote (c.1001dup, (p.Asp334Glufs\*2)) dans le gène *ARCN1*, présent également chez les deux fœtus. Il s'agissait d'un variant non-rapporté dans gnomAD, et qui n'a pas pu être recherché chez les parents de la patiente. Le gène *ARCN1* (pLI = 1) n'avait alors été rapporté que deux fois, avec des variants tronquants (Izumi *et al.*, 2016; Verloes *et al.*, 1997). Avec l'accord de la patiente, le variant a pu être inclus dans le travail de recherche du Dr K Izumi (Ritter *et al.*, 2022) (Annexe 3).

## 2. L'interprétation d'ES de patients atteints de maladies dermatologiques très rares

Dans le cadre d'un projet API jeune chercheur (DERMASEQ – Pr LIPSKER), l'ES a permis d'identifier chez un patient une variation pathogène dans le gène *C1R* confirmant le diagnostic clinique de syndrome d'Ehlers Danlos parodontal (SEDp). Grâce à plusieurs appels à collaboration nationaux j'ai pu colliger 15 nouveaux patients atteints de SEDp, ce qui correspond à la première cohorte française pour ce syndrome rare, qui se caractérise par une parodontopathie sévère d'apparition très précoce avec perte prématurée des dents, une hyperpigmentation prétibiale et une fragilité cutanée. Ce travail a permis de réaliser une description phénotypique fine de ce syndrome avec un focus sur les anomalies vasculaires, à ce jour rarement rapportées, et de souligner l'importance d'un diagnostic précoce du SEDp pour débiter très tôt des soins dentaires appropriés et préciser le conseil génétique. Ce travail a fait l'objet d'une publication (El Chehadeh *et al.*, 2021) (Annexe 4).

## 3. L'interprétation de WGS de patients présentant l'association de mélanomes multiples et lipomatose mésosomatique

Dans le cadre du projet DERMASEQ, huit patients présentant l'association de mélanomes multiples (sporadiques) et d'une lipomatose mésosomatique autosomique dominante ont été inclus pour un séquençage de génome. Après le séquençage, j'ai focalisé mes analyses sur une extraction des données de l'exome, dans un premier temps. Ce travail préliminaire m'a permis d'apprendre le maniement des différents outils bioinformatiques et des bases de données nécessaires à l'analyse et l'interprétation du WGS. Ceci m'a également permis d'appréhender les difficultés rencontrées (très nombreux variants, nombreux logiciels de détection/annotation, mécanismes mutationnels encore mal connus, patients séquencés en *solo*, particularités liées aux gènes de prédisposition au cancer (pénétrance incomplète et expressivité variable rendant difficile la détermination du statut de « sain », surtout chez les individus jeunes). Nous n'avons pas identifié de variant pathogène ou probablement pathogène chez ces patients.

## 4. La description phénotypique de syndromes rares avec DI

Durant ces cinq années, en parallèle de ce travail d'interprétation, j'ai continué d'effectuer des travaux de recherche clinique comprenant :

- La description phénotypique d'une pathologie très rare, le syndrome de Frank-Ter Haar, dont était atteinte l'une des patientes que je suivais, et dont les variants pathogènes ont pu être identifiés au sein du laboratoire de recherche U1112. Ce travail a abouti à une publication (Durand *et al.*, 2020) (Annexe 5).
- La contribution à l'identification d'un gène de lipodystrophie partielle syndromique avec atteinte neurologique périphérique et cognitive, le gène *PLAAT3* (11q12.3-q13.1). Ce travail collaboratif a abouti à une publication en co-premier auteur (sous presse - *Nature Genetics*) (Annexe 6).

# Discussion

---

Le séquençage de génome réalisé chez 15 patients présentant une DI présumée liée à l'X et ayant bénéficié préalablement d'un bilan étiologique négatif a permis d'identifier un variant pathogène ou probablement pathogène dans des gènes de l'X chez 4 patients soit 26,6% d'entre eux. Plusieurs variants candidats sont toujours en cours d'étude.

## Les nombreux *challenges* de l'interprétation du WGS

### *Les très nombreux variants*

L'intérêt du WGS a été clairement démontré dans l'identification de variants pathogènes dans les régions codantes, incluant des SNV/indel, des SV de type CNV (délétions emportant un ou plusieurs gènes par exemple), et des remaniements chromosomiques complexes (insertion-duplication intragénique), avec un taux diagnostique de plus de 40% dans la DI (Gilissen *et al.*, 2014). Cependant, malgré ses performances diagnostiques indiscutables, plusieurs limites existent et rendent difficile l'interprétation des données issues du WGS. En effet, Gilissen *et coll.* rapportaient plus de 4 millions de SNV/indel par individu (Gilissen *et al.*, 2014). Plus concrètement, parmi 50 patients, 84 SNV (incluant les *indel*) *de novo* étaient identifiés dans les régions codantes.

J'ai été confrontée à cette difficulté de devoir trier et prioriser un nombre très important de variants par individu (8439 variants sur le chromosome X dont 846 variants rares et 339 CNV), nécessitant l'utilisation d'outils robustes permettant de filtrer et d'annoter efficacement ces derniers.

Le défi de rechercher LE variant causal parmi des millions de variants en WGS est d'autant plus vrai qu'il a été démontré que le chromosome X présente un enrichissement en gènes de pathologies du neurodéveloppement et en gènes ayant un rôle important dans le développement cérébral. En effet, en 2022, Leitão *et coll.* ont montré que les gènes du chromosome X étaient significativement plus fréquemment associés à des phénotypes neurologiques que les gènes autosomiques (77 % *versus* (vs) 55-76% ;  $p = 7 \times 10^{-3}$ ). Plus spécifiquement, ils ont observé que le chromosome X était enrichi en gènes associés à la DI (58% vs 27-45% ;  $p = 5,92 \times 10^{-11}$ ), l'épilepsie (46% vs 23-38% ;  $p = 1,12 \times 10^{-4}$ ) et les troubles du langage (32% vs 11-24% ;  $p = 5,59 \times 10^{-3}$ ) (Leitao *et al.*, 2022).



### ***Le challenge des variants non-codants***

Par ailleurs et surtout, l'identification des variants dans les régions non-codantes représente un véritable *challenge*. En effet, initialement, l'utilisation du WGS dans la DI s'était surtout concentrée sur l'identification de variants codants pour des protéines, or, compte tenu de l'importance d'un contrôle étroit de l'expression des gènes pour le développement cérébral, les variations qui affectent les régions régulatrices non-codantes du génome sont rapidement apparues comme étant susceptibles de jouer un rôle important dans les TND, faisant émerger le terme de « *missing heritability* » dans les maladies monogéniques (Sadée *et al.*, 2014; Zhang et Lupski, 2015; Wanke *et al.*, 2018).

En particulier, au sein des régions non-codantes du génome, les éléments régulateurs tels que les promoteurs, les *enhancers*, et les régions non traduites (UTR) apparaissent comme d'excellentes régions candidates pour des variations pathogènes car ils déterminent le lieu, le moment et le niveau d'expression d'un gène. En effet, les variants de régulation qui entraînent une réduction significative de l'expression des gènes peuvent mimer des variants perte de fonction. Inversement, les variants entraînant une augmentation de l'expression d'un gène, ou qui modifient la chronologie ou la localisation tissulaire de cette dernière, peuvent également être préjudiciables et responsables d'un phénotype (Abelson *et al.*, 2005; Haraksingh et Snyder, 2013; Stranger *et al.*, 2007). Ainsi, depuis quelques années, plusieurs publications ont démontré l'intérêt du WGS dans l'identification de variants non-codants, parmi lesquels on peut citer à titre d'exemple dans les TND :

- des variants situés dans les régions 3'UTR altérant un site de liaison aux micro-ARNs ou aux protéines de liaison à l'ARN (Suhl *et al.*, 2015; Devanna *et al.*, 2018; Devanna *et al.*, 2018) notamment les variants distaux ou mal définis (Wright *et al.*, 2021).

- des variants situés dans les régions 5'UTR responsables de la création d'un codon d'initiation AUG : il a été démontré que ces variants peuvent être soumis à une forte sélection négative et représentent une cause importante de maladies mendéliennes telles que la neurofibromatose de type 1 ou le syndrome de Van der Woude (de Lima *et al.*, 2009; Whiffin *et al.*, 2020). Wright *et coll.* ont étudiés 9858 sujets de l'étude DDD à la recherche de variants pathogènes *de novo* dans les régions 5'UTR de gènes impliqués dans les TND par un mécanisme d'haploinsuffisance. Ils ont identifié quatre SNV et deux CNV en amont du gène *MEF2C* chez dix patients et ont démontré que ces variants provoquaient un TND par le biais de trois mécanismes de perte de fonction (perturbant la transcription, la traduction et/ou la fonction protéique). Ces variants des régions non-codantes représentaient 23 % des variants pathogènes identifiés dans *MEF2C* dans la cohorte DDD (Wright *et al.*, 2021). Ce type de variants a également été décrit sur des gènes de l'X, Kumar *et coll.* ayant rapporté un variant

pathogène familial dans la région 5'UTR de *DLG3* responsable d'une altération du repliement de l'ARNm (Kumar *et al.*, 2016).

- des variants situés au sein d'ARNs longs non-codants reliés à un surrisque de pathologies cardio-vasculaires (Azodi *et al.*, 2020; Ishii *et al.*, 2006).

- des variants altérant la structure de TADs (Topologically associated domains) (délétion, inversion ou duplication) responsables de malformations des membres (Lupianez *et al.*, 2015).

- des variants dans les éléments cis-régulateurs (CRE-SV) et des variants perturbant la régulation transcriptionnelle et la régulation des *RNA-binding proteins* chez patients atteints de trouble du spectre de l'autisme (Brandler *et al.*, 2018; Zhou *et al.*, 2019).

Lorsqu'on sait que seulement 2% du génome humain code pour des protéines (Xue *et al.*, 2015), bien que 90% de ce dernier soient transcrits en ARN (Azodi *et al.*, 2020), on comprend aisément la difficulté d'interprétation des variants situés dans des régions non-codantes. A titre d'exemple, parmi 50 patients atteints de DI, Gilissen *et coll.* ont identifié 43 variants *de novo* dans des régions non-codantes de 528 gènes de DI (promoteur (1 variant), 5'UTR (2 variants), introns (38 variants), site non canonique d'épissage (1 variant), 3'UTR (1 variant)) (Gilissen *et al.*, 2014).

De plus, l'intégralité des éléments régulateurs et le rôle précis de chacun reste encore mal connu. La difficulté réside ainsi dans la nécessité de prouver le caractère causal de ces variants non-codants, en affirmant qu'ils présentent un impact fonctionnel sur la régulation d'un gène. Différentes approches fonctionnelles ont déjà été utilisées afin de confirmer l'activité régulatrice de régions candidates telles que les MPRA (Massively Parallel Reporter Assays), le STARR-seq (self-transcribing active regulatory region sequencing), ou l'utilisation de la technologie CRISPR-Cas9 (Rojano *et al.*, 2019; Santiago-Algarra *et al.*, 2017). Par ailleurs, plusieurs scores et outils de prédiction *in silico* de variant non-codants ont été créés et comparés (Liu *et al.*, 2017) tels que CADD (Kircher *et al.*, 2014), FATHMM-MKL (Shihab *et al.*, 2015), fitCons (Gulko *et al.*, 2015), ReMM (Smedley *et al.*, 2016), Fire (Functional Inference of Regulators of Expression) (Ioannidis *et al.*, 2017) et NCBoost (Caron *et al.*, 2019). Cependant, si ces scores fournissent des informations pour une position donnée du génome, ils ne présagent pas toujours des conséquences d'un variant situé dans cette région régulatrice sur un gène cible qui peut se situer très à distance de cette dernière. Il peut être par ailleurs difficile de relier un effet fonctionnel à un phénotype.

Ceci est bien illustré par le cas du patient suivi dans notre service chez lequel a été identifié en WGS une délétion dépourvue de gènes mais située en amont du gène *PQBP1* (Xp11.23), et qui reste pour le moment de signification incertaine (classe 3 ACMG), donc non-utilisable pour le conseil génétique, bien qu'intéressante en termes de cohérence phénotypique.

### ***Le cas particulier des expansions et des éléments mobiles***

Les répétitions en tandem (TR) représentent l'une des catégories de variations les plus abondantes dans le génome humain (plus d'un million), qui sont polymorphes par nature, et peuvent être très instables en fonction de leur longueur. Elles sont principalement localisées dans les régions non-codantes du génome, bien que des répétitions de triplets puissent également être présentes dans les parties codantes des gènes (Gymrek, 2017; Willems *et al.*, 2014). L'augmentation de la longueur des répétitions au fil des générations est un processus bien connu et responsable de diverses pathologies affectant principalement le système nerveux central. Il existe deux principaux types de TR:

- les expansions affectant les régions codantes, conduisant principalement à des tronçons anormalement longs de polyglutamines (polyQ, principalement codés par les codons CAG) (La Spada *et al.*, 1991) ou de polyalanines (polyA, codés par les codons GCN) (Koide *et al.*, 1999).
- les expansions atteignant les régions non-codantes des gènes: par exemple, des répétitions pentanucléotidiques non-codantes TTTTA et insertions de TTCA dans six gènes différents (*SAMD12*, *STARD7*, *MARCHF6*, *YEATS2*, *TNRC6A*, et *RAPGEF2*) sont responsables d'épilepsie myoclonique familiale de l'adulte (Hannan, 2018; Corbett *et al.*, 2023).

Les premières expansions identifiées chez l'homme concernaient les gènes *FMR1* (CGG) (Syndrome de l'X fragile, MIM 300624), *AR* (CAG) (maladie de Kennedy, MIM 313200), *DMPK* (CTG) (dystrophie myotonique de Steinert, MIM 160900), *HTT* (CAG) (maladie de Huntington, MIM 143100) et *FXN* (GAA) (ataxie de Friedreich, MIM 229300). Au moins 50 maladies associées à des expansions ont été décrites depuis, la moitié d'entre elles n'ayant été reconnue qu'au cours des dix dernières années, du fait de limites techniques ne permettant pas leur détection (Depienne et Mandel, 2021). Le développement récent des technologies de *long read* offre une occasion unique d'étudier systématiquement la contribution des TR et des expansions de répétitions dans l'apparition de la maladie. Jusqu'à récemment, la plupart des outils n'étaient capables que d'appeler des génotypes à des loci précis pour un motif donné. Cette situation a récemment changé grâce au développement d'outils tels que TRhist95 (Doi *et al.*, 2014) et ExpansionHunter DeNovo (Dolzhenko *et al.*, 2020) qui évaluent la présence d'expansions répétées à l'échelle du génome. Néanmoins, la détection des expansions en l'absence d'hypothèses préalables sur le motif concerné reste un défi en raison du nombre considérable de motifs possibles, de l'abondance des TR dans le génome humain et de la difficulté à distinguer clairement les expansions pathogènes des allèles polymorphiques (Depienne et Mandel, 2021).

Les éléments mobiles (EM) ou éléments transposables sont des séquences d'ADN qui peuvent être copiées ou déplacées de manière autonome dans le génome. Ils constitueraient

environ 45% à 69% de la séquence du génome humain (de Koning *et al.*, 2011), mais une minorité reste active et capable de créer de nouvelles insertions (IEM) (Mills *et al.*, 2007). On estime qu'une naissance sur 114 environ présente une IEM *de novo* (Gardner *et al.*, 2019), ces dernières représentent ainsi un élément majeur de l'évolution de la modification de l'architecture génomique (Thung *et al.*, 2014; Torene *et al.*, 2020). Les EM peuvent être classés en deux catégories en fonction de leur mode de transposition. Les rétrotransposons de classe I passent par un intermédiaire ARN par « copier-coller » et comprennent les EM de type L1 (6 kb, 500 000 copies), de type Alu (300 bp, 1 000 000 copies) et de type SVA (2 kb, 3 000 copies), tandis que les transposons de classe II passent par un intermédiaire ADN et se déplacent généralement par « couper-coller » (Thung *et al.*, 2014). La transposition des EM se produit souvent dans la lignée germinale ou au cours de l'embryogenèse précoce et peut être à l'origine de variations pathogènes en s'insérant dans des régions importantes sur le plan fonctionnel, perturbant ainsi la fonction des gènes (Cordaux et Batzer, 2009).

La première insertion d'EM (IEM) pathologique chez l'homme a été identifiée à l'état hémizygote dans le gène *FVIII* chez deux patients atteints d'hémophilie A (Kazazian *et al.*, 1988). Depuis, plus de 124 IEM pathogènes ont été découvertes, représentant 0,03% à 0,04 % des variants pathogènes identifiés dans des cohortes de près de 10000 et 39000 patients avec TND, respectivement, étudiés en ES (majorité d'insertions d'éléments Alu), permettant une augmentation du taux diagnostique d'environ 0,15% (Delvallee *et al.*, 2021; Gardner *et al.*, 2019; Sen *et al.*, 2006; Torene *et al.*, 2020). Leur structure de séquences hautement répétées les rend difficiles à détecter, ce qui a conduit à la création d'outils de détection d'EM tels que Mobster, MELT, ou SCRAMble (Soft Clipped Read Alignment Mapper) (Thung *et al.*, 2014; Gardner *et al.*, 2017; Torene *et al.*, 2020), ces derniers ayant fait leur preuve en permettant d'augmenter le nombre de diagnostic dans des cohortes de patients non-résolus ayant bénéficié d'un ES (Wijngaard *et al.*, 2023).

En ce qui concerne ce travail de thèse, la recherche d'expansions et d'EM débute seulement et constitue une des perspectives à atteindre.

### ***L'évolution rapide des outils bioinformatiques***

Il existe de multiples logiciels et outils de prédiction *in silico*, qui aident considérablement à l'interprétation de ces nombreux variants. Cependant, ces outils évoluent très régulièrement, et certains deviennent rapidement obsolètes. J'ai en effet été amenée à utiliser certains outils au début de mon travail, qui ont progressivement été remplacés par d'autres, plus performants, qu'il a fallu apprendre à manier.

Ainsi, le caractère évolutif de la génétique moléculaire, notamment l'évolution rapide des techniques de bioinformatique ainsi que l'identification régulières des nouveaux gènes de DI,

imposent aux bioinformaticiens de devoir gérer les mises à jour continues des outils, logiciels et bases de données, afin de garantir aux biologistes moléculaires les informations les plus récentes et actualisées possibles pour leur travail d'interprétation (Lelieveld *et al.*, 2015, 2016). Cela implique également de la part des laboratoires qu'ils puissent mettre en œuvre des stratégies d'automatisation pour la réanalyse des données existantes (Fung *et al.*, 2020; Liu *et al.*, 2019), ce qui, aujourd'hui, n'est faisable pas en pratique de manière systématique dans tous les laboratoires de diagnostic français, faute de moyens le plus souvent.

Or cette réanalyse a fait ses preuves en augmentant le rendement diagnostique de 31 % à 53 % dans l'étude menée par Schobers *et coll.*, où les nouveaux diagnostics ont été faits grâce à la reclassification de variants précédemment identifiés (10 %), à l'amélioration des pipelines bioinformatiques (19 %), à une meilleure couverture après le reséquençage (29%) et à de nouvelles associations gène-maladie (42 %) (Schobers *et al.*, 2022). Dans l'étude menée par de Ligt en 2012, la réévaluation moléculaire réalisée 2 ans après le premier ES chez 100 patients a également permis d'identifier ou de reclasser des variants pathogènes, notamment dans les gènes *DDX3X* et *MECP2* (de Ligt *et al.*, 2012; Gilissen *et al.*, 2014).

Ces observations soulignent la nécessité pour les biologistes interprétant ces données d'ES et WGS, de bénéficier d'outils et de bases de données de variants les plus à jour possible, et de bénéficier de moyens supplémentaires pour automatiser cette réanalyse des données. Ce d'autant que nombre de patients ne reprennent pas contact avec le généticien, après un premier résultat négatif, pour une réévaluation clinique, comme cela leur est pourtant proposé. Or, la majorité des nouveaux diagnostics concernent des patients qui n'ont pas recontacté leur généticien, ce qui prouve l'intérêt d'une réanalyse systématique (Ji *et al.*, 2021; Schobers *et al.*, 2022). Une réévaluation clinique régulière des patients sans diagnostic reste cependant fondamentale, comme cela était illustré par le cas des patients atteints de syndrome de Simpson Golabi Behmel cités précédemment.

A titre d'exemple également, le consortium Solve-RD (<https://solve-rd.eu/>), par l'utilisation combinée de la réanalyse massive de données de plus de 19 000 patients atteints de maladies rares non résolues et de nouvelles approches « omiques », a permis de résoudre 255 cas parmi 8393 patients ayant eu un ES ou WGS négatif (Zurek *et al.*, 2021). Ce même consortium a permis, après une réanalyse basée sur les annotations de ClinVar de 3576 ES (1522 cas index et 2054 apparentés), d'identifier un variant causal chez 59 individus (3,9%) (Denomme-Pichon *et al.*, 2023). Dans toutes ces situations, les auteurs soulignent l'importance de l'interaction cliniciens-biologistes et de l'expertise clinique, notamment dans des situations où la forte suspicion diagnostique du clinicien a permis d'établir un diagnostic après une première analyse négative suivie d'une réanalyse guidée par la suspicion clinique (Corominas *et al.*, 2022).

## ***La difficulté d'identification des variants de structure en WGS***

Les variants de structure (SV) sont une cause importante de MM. Ils correspondent à des réarrangements chromosomiques de plus de 50 pb, qui peuvent être bénins ou pathogènes, déséquilibrés (CNV : délétions ou duplications, IEM) ou « apparemment équilibrés » (ABCR : insertions, inversions, translocations réciproques), isolés ou combinés (chromothripsis). Ils représentent environ 17% des événements pathogènes rares à l'échelle du génome, et leurs effets délétères sont équivalents à ceux des variants codants avec perte de fonction. Approximativement 90% de ces SV correspondent à des délétions « non-codantes » (Abel *et al.*, 2020).

Un génome contient environ 2100 à 2500 SV (1000 grandes délétions, 160 CNV, 915 insertions Alu, 51 insertions de rétrotransposons de type SVA (SINE/VNTR/Alu), 4 NUMT (nuclear mitochondrial DNA segment), et 10 inversions, affectant 20 millions de bases (Auton *et al.*, 2015; Sudmant *et al.*, 2015). L'identification des SV représente donc un réel *challenge* pour les biologistes moléculaires. Il a été démontré que le WGS, notamment la stratégie paired-end utilisée ici, est une méthode efficace pour la détection des SV (Korbel *et al.*, 2007), y compris des SV complexes tels que des délétion-inversion-duplication, duplication-inversion-duplication etc. (Sanchis-Juan *et al.*, 2018). Le WGS a également montré sa supériorité pour caractériser les points de cassures des ABCR, ces derniers pouvant avoir comme conséquence un effet de position ou une interruption de gènes, de TADs ou d'EM, responsables d'un phénotype pathologique dans environ 45% des cas (Chen *et al.*, 2010; Lupianez *et al.*, 2015; Redin *et al.*, 2017; Schluth-Bolard *et al.*, 2013).

Cependant, les difficultés d'identification des SV ont déjà été soulevées par plusieurs auteurs, notamment du fait du nombre important de faux-positifs, et de nombreux outils de détection de SV ont été créés, certains étant complémentaires entre eux et chacun ayant ses avantages et inconvénients (Mahmoud *et al.*, 2019). Par exemple, DELLY (Rausch *et al.*, 2012) intègre l'analyse des « lectures fractionnées » (*split-reads*) dans sa recherche de distances et d'orientations anormales entre les paires de lectures. Bien que cela augmente la précision du point de cassure et permette la détection de petites délétions (20 pb), les événements plus importants restent difficiles à distinguer des artefacts. Pour y remédier, certaines méthodes ont intégré des informations complémentaires notamment la profondeur de lecture. Par exemple, LUMPY (Layer *et al.*, 2014) effectue une analyse conjointe de la profondeur de lecture, de la discordance des lectures pairées et des lectures fractionnées, ce dont tient compte également Manta (Chen 2010). GRIDSS (Cameron *et al.*, 2017), quant à lui, ne retient que les lectures qui fournissent des preuves de l'existence de SV et les assemble ensuite *via* un graphe

positionnel. Ainsi, GRIDSS et LUMPY offrirait un taux de précision de 90% environ contre 70-85 % pour DELLY ou Manta (Mahmoud *et al.*, 2019; Rausch *et al.*, 2012).

J'ai moi-même utilisé plusieurs logiciels de détection de variants structuraux (Smooove, Manta et CNVpytor) pour tenter de minimiser les risques de faux positifs ou faux-négatifs. J'ai néanmoins été mise en difficulté dans cette recherche, d'une part du fait du nombre important de CNV détectés (339 CNV sur l'X par patient), et d'autre part parce que bon nombre d'entre eux n'étaient pas retrouvés après visualisation des BAMs sur IGV (et correspondaient donc vraisemblablement à des faux-positifs) ou n'étaient pas détectés par les trois logiciels. L'utilisation de ces outils a cependant permis l'identification en WGS d'un CNV pathogène chez un patient, une duplication hémizygote de la région Xp11.22 de près d'1 Mb correspondant à une duplication récurrente pathogène rare mais bien décrite. Il a également été possible de démontrer qu'il s'agissait d'une duplication en tandem, correspondant à ce qui est décrit dans la littérature (Santos-Reboucas *et al.*, 2015), ce que la CGH-array n'aurait pas permis de déterminer.

Bien que des progrès restent probablement encore à faire sur l'amélioration de ces outils de détection de SV, l'utilisation d'autres techniques de biologie moléculaire en plein essor, telles que la cartographie optique, représente une perspective prometteuse (Lam *et al.*, 2012; Mak *et al.*, 2016). En effet, outre le fait que cette technique permette d'identifier des aneuploïdies et des remaniements chromosomiques complexes tels que le chromothripsis (Neveling *et al.*, 2021), des chromosomes en anneau ou des isochromosomes (Mantere *et al.*, 2021), elle permet de détecter de manière très précise différents types de SV, incluant des variants déséquilibrés (CNV) et des variants équilibrés (inversions (de Bruijn *et al.*, 2023), translocations équilibrées pouvant mener à de potentiels gènes de fusion dans les hémopathies malignes (Smith *et al.*, 2022). Ainsi, de nombreux auteurs s'accordent à dire que cette technique s'apprête à supplanter prochainement les techniques de cytogénétiques standard, aussi bien en oncologie, pour détecter des anomalies somatiques sur tissu tumoral (Coccaro *et al.*, 2023), qu'au niveau constitutionnel chez des patients atteints de maladies rares avec anomalies du développement et/ou TND (Iqbal *et al.*, 2023; Mantere *et al.*, 2021).

## **Les défis posés par la DILX**

### ***La difficulté d'identifier des familles véritablement concernées par une DILX***

Ce travail m'a permis de me rendre compte qu'il était bien plus difficile qu'attendu d'identifier des familles réellement concernées par une DILX. En effet, lors du recrutement des patients, nous avons surtout identifié des familles avec deux frères atteints, ce qui pouvait correspondre à une hérédité liée à l'X mais également autosomique récessive. Par ailleurs, parmi les variants pathogènes identifiés en WGS dans le cadre du diagnostic, 6 sur 8 étaient de

*novo* et 2 étaient hérités de mères asymptomatiques, sans histoire familiale. Il n'était donc pas possible de « reconnaître » ou supposer une transmission liée à l'X sur la seule histoire familiale.

De la même façon, Gilissen *et coll.*, qui ont étudié 50 patients avec DI ayant bénéficié d'une CGH-array et d'un ES négatifs, ont identifié un variant pathogène chez 21 patients (42%), incluant 1 variant dans un gène de l'X chez 4 patients, *de novo* chez tous. D'une manière plus générale, on voit que 60% des variants pathogènes identifiés dans leur cohorte étaient *de novo* contre 2% de variants hérités, ce qui montre la faible représentation des familles avec variants hérités, qui plus est sur l'X, et dont l'arbre généalogique est clairement évocateur d'une histoire familiale (Gilissen *et al.*, 2014).

### ***Le phénotype des filles encore mal compris***

Alors que 10 à 12% des cas de DI chez les hommes ont été attribués à une DILX (Ropers et Hamel, 2005), une proportion croissante de DI chez les femmes est maintenant expliquée par des variants *de novo* dans des gènes situés sur le chromosome X, tels que *DDX3X* (Snijders Blok *et al.*, 2015), *ARHGEF9* (Ghesh *et al.*, 2021) ou encore *STAG2* (Provenzano *et al.*, 2022). Ainsi, la DILX représente également un défi par le caractère incomplet des connaissances et de la compréhension du lien entre certains variants dans des gènes de l'X et le phénotype des filles atteintes. De nombreux exemples existent pour illustrer la difficulté d'interprétation de ces variants tels que le MECP2DS, qui a été rapporté, bien que très rarement, chez des filles atteintes de DI (légère à sévère, et avec des duplications héritées comme *de novo*) et chez lesquelles il n'a pas été mis en évidence de corrélation entre l'XCI dans le sang, la taille de la duplication et la sévérité du phénotype. Ces filles (duplication interstitielle) présentent cependant un phénotype différent des garçons atteints, avec une DI le plus souvent isolée et non syndromique (El Chehadeh *et al.*, 2017; Peters *et al.*, 2019; Ta *et al.*, 2022). Ces observations, bien que rares, rendent délicat le conseil génétique des couples à risque, notamment en cas de grossesses de fœtus féminin et mènent régulièrement à des discussions multidisciplinaires sur la question du DPN et du DPI, pour lesquels il n'existe pas à ce jour en France de consensus.

De la même manière, la pathologie liée au gène *ARHGEF9* était initialement considérée comme étant récessive liée à l'X (Aarabi *et al.*, 2019; Wang *et al.*, 2018), jusqu'à l'identification en 2020 de SNV pathogènes chez des filles, dont des variants perte de fonction, incitant à reconsidérer encore une fois cette classification (Ghesh *et al.*, 2021; Scala *et al.*, 2021). Un autre exemple déroutant est celui du gène *MED12* (Xq13.1) dont les variants faux-sens sont associés à plusieurs syndromes de DILX dont le syndrome d'Opitz-Kaveggia, de Lujan-Fryns, d'Ohdo lié à l'X mais également à de la DI non syndromique (Graham et Schwartz, 2013; Prontera *et al.*, 2016). Or, en 2021, Li *et coll.* ont démontré que des variants perte de fonction de *MED12*



étaient responsables d'un syndrome très rare n'atteignant que les filles, le syndrome de Hardikar, qui associe des anomalies congénitales multiples mais sans DI (D. Li *et al.*, 2021), élargissant considérablement les connaissances et le spectre phénotypique lié à ce gène.

Enfin, la duplication récurrente Xp11.22 identifiée chez le patient PF a également été rapportée chez plusieurs filles atteintes de DI (Edens *et al.*, 2011; Evers *et al.*, 2015; Giorda *et al.*, 2009; Grams *et al.*, 2016), avec, le plus souvent mais pas toujours, une XCI biaisée avec inactivation préférentielle de l'X normal. Certains auteurs attribuent la variabilité phénotypique chez ces filles à la variabilité de l'XCI au niveau cérébral et au fait que plusieurs gènes inclus dans cette duplication échappent à l'XCI tels que *KDM5C*, *IQSEC2* and *SMC1A* (Moey *et al.*, 2016; Santos-Reboucas *et al.*, 2015).

Ainsi, la compréhension de la DILX est rendue d'autant plus complexe qu'il a été démontré que plusieurs autres gènes de l'X comme *DDX3X* (Snijders Blok *et al.*, 2015), *KDM6A* (Bogershausen *et al.*, 2016), ou encore *STAG2* (Garieri *et al.*, 2018; Provenzano *et al.*, 2022), échappent au phénomène d'inactivation de l'X (Brand *et al.*, 2021; Tukiainen *et al.*, 2017).

### ***L'XCI qui ne nous aide pas toujours***

Alors qu'initialement l'existence d'un phénotype chez les filles porteuses d'un variant pathogène de l'X était reliée à l'existence d'un biais complet d'inactivation de l'X, on sait maintenant qu'il ne s'agit pas d'une science exacte et que ce paramètre n'aide pas toujours à conclure sur la pathogénicité de certains variants (Ziats *et al.*, 2020). Ceci a été démontré pour la duplication du gène *MECP2* et la duplication Xp11.22 (El Chehadeh *et al.*, 2017; Evers *et al.*, 2015) mais également pour des gènes de l'X comme le gène *ARHGEF9*, pour lequel des variants pathogènes hétérozygotes ont été rapportés chez 14 filles atteintes de DI, la moitié ayant une XCI biaisé et l'autre moitié une XCI aléatoire (Ghesh *et al.*, 2021; Scala *et al.*, 2021). Les auteurs concluaient que ces résultats discordants ne remettaient pour autant pas en question la pathogénicité des variants identifiés, puisque l'analyse réalisée sur le sang peut ne pas refléter le profil d'inactivation de l'X dans les neurones post-synaptiques inhibiteurs (Scala *et al.*, 2021).

### ***La difficulté d'interprétation des variants faux-sens sur l'X en l'absence de ségrégation étendue***

Identifier des nouveaux gènes de DI sur l'X est donc un *challenge*, du fait de la difficulté pour reconnaître des familles évocatrices, et pour interpréter le phénotype, notamment chez des filles symptomatiques porteuses d'un variant de l'X hérité d'une mère asymptomatique. L'interprétation des variants liés au chromosome X, notamment les faux-sens, dans les cas sporadiques et les petites familles, reste difficile en l'absence d'une ségrégation étendue dans la famille, laquelle est rarement réalisable. Plusieurs auteurs s'accordent sur le fait que l'identification de nouvelles associations gène-maladie sur le chromosome X nécessite des

études de familles comptant plusieurs hommes atteints (Hu *et al.*, 2016; Tarpey *et al.*, 2009) et plusieurs sujets présentant des phénotypes similaires (Snijders Blok *et al.*, 2015). En 2012, de Ligt et collaborateurs publiaient dans le journal *New England Journal of Medicine*, les résultats du séquençage d'ES en *trio* de 100 patients avec DI avec un taux diagnostique de 16 %, incluant 5 patients avec un variant possiblement pathogène sur l'X (2 *de novo* et 3 hérités). En 2014, Gilissen, rapportait que deux des variants hérités, situés sur les gènes *PDHA1* et *ARHGEF9* (situés sur l'X), n'ont finalement pas été considérés comme pathogènes, suite aux résultats de l'étude de ségrégation familiale (de Ligt *et al.*, 2012; Gilissen *et al.*, 2014).

Dans ma pratique clinique, je suis fréquemment confrontée à la difficulté d'interprétation de variants de classe 3 dans un gène de l'X chez un patient pour lequel l'impossibilité de réaliser une étude de ségrégation familiale est un obstacle majeur car ne nous permettant pas de statuer sur la pathogénicité du variant en question, ni de proposer un conseil génétique fiable. En pratique, ceci signifie pour les couples une impossibilité d'avoir recours à un DPN ou DPI pour de futures grossesses, avec le retentissement psychologique négatif qu'on imagine aisément. Nous sommes confrontés à la même problématique lorsqu'aucune cohorte de réplication n'est publiée au moment de l'identification d'un variant candidat chez l'enfant atteint d'un couple ayant un projet de grossesse à court terme (voire une grossesse déjà débutée), dont la temporalité n'est pas toujours compatible avec les avancées des connaissances scientifiques. Un autre écueil est celui des biais d'interprétation des variants pour les gènes de l'X liés au fait que les pipelines bioinformatiques peuvent être mis en difficulté pour traiter la dichotomie filles-garçons (Leitao *et al.*, 2022).

## **Les limites communes aux techniques de NGS**

Ainsi, dans certains cas, malgré l'utilisation d'outils de prédiction *in silico*, l'organisation de discussions clinico-biologiques au niveau international avec des équipes expertes du gène et une étude de ségrégation familiale correctement menée, le niveau de preuve reste insuffisant pour déterminer si le variant identifié est responsable de la pathologie du patient. Cette situation est fréquente pour les variants faux-sens identifiés dans des gènes de l'X, notamment lorsqu'ils sont hérités de plusieurs femmes asymptomatiques dans une famille et qu'aucun homme sain n'est disponible pour être testé. Cette limite, liée à la non-exhaustivité de nos connaissances et de notre compréhension des mécanismes mutationnels et des fonctions de notre « ome » (génomique, transcriptome, protéome etc.), reste de mise malgré les progrès apportés par le WGS et continue à rendre nos discussions clinico-moléculaires toujours plus passionnantes et animées.

Une des façons de palier à cette limite est d'avoir recours à des tests fonctionnels, pour mieux comprendre l'impact du variant sur la transcription (étude sur ARNm) (Courraud *et al.*, 2021) ou sur l'expression ou la fonction de la protéine (études protéiques) (Mattioli *et al.*,

2019), et donc son implication dans le phénotype du patient. Karam et *coll.* ont ainsi pu élégamment démontrer, à partir de fibroblastes de patients, que des variants pathogènes du gène *BBS5* avaient un impact à la fois sur la formation du cil primaire (diminution du nombre et augmentation de la longueur des cils) et sur sa fonction (diminution de l'activation de la voie Hedgehog) (Karam *et al.*, 2023). Concernant ce travail de thèse, nous avons eu recours à des études fonctionnelles poussées dans le but de prouver l'implication du gène *SLITRK2* dans le phénotype d'une série de patients avec TND, et des études fonctionnelles sont également envisagées chez les patients P6 et PH porteurs de variants candidats hémizygotés dans les gènes *MED14* et *ARHGEF9*, respectivement.

# Conclusions

---

Ce travail, qui est décomposé en plusieurs parties, permet de confirmer l'intérêt des techniques de séquençage à très haut débit telles que l'ES et le WGS, dans la compréhension des bases moléculaires de la DILX, que ce soit *via* l'identification de nouveaux gènes candidats de TND et DI (*SLITRK2*, *MED14*, *PLAAT3*) ou *via* l'identification de variants candidats dans des gènes en cours de description par une autre équipe de recherche (*BCORL1*, *ARCN1*).

Toutefois, ce travail souligne également les limites et les difficultés que représente l'interprétation des données de WGS, en particulier de par le très grand nombre de variants, notamment introniques, identifiés par patient et le *challenge* que représente l'identification des variants non-codants. Ceci est particulièrement vrai en ce qui concerne la DILX pour laquelle l'interprétation des nombreux variants faux-sens est délicate en l'absence de ségrégation familiale étendue et aboutit fréquemment à une absence de conclusion possible, avec un variant qui restera en « classe 3 » ACMG tant qu'une étude fonctionnelle ne sera pas disponible pour aider à statuer sur sa pathogénicité. Par ailleurs, de nombreux phénomènes restent encore mal compris dans la DILX tels que la présence d'un phénotype chez les filles dans des pathologies habituellement clairement « récessives » liées à l'X et pour lesquelles l'étude de l'inactivation de l'X n'apporte le plus souvent pas d'aide à l'interprétation. Ce travail nous a également montré à quel point il peut être difficile, voire impossible, de suspecter une DILX en se fiant uniquement à la distribution de la DI dans la famille. En effet, que ce soit dans la cohorte des 15 patients ayant bénéficié d'un WGS, ou chez les patients du service de génétique médicale ayant bénéficié d'un ES ou d'un WGS, lorsqu'un variant pathogène dans un gène de l'X était identifié, les situations où l'arbre généalogique suggérait une liaison à l'X étaient exceptionnelles, les variants étant majoritairement survenus *de novo* chez le cas index ou chez sa mère, ou hérité d'une mère sans histoire familiale.

L'ensemble de ces observations montre que des progrès restent probablement à faire sur la poursuite du développement de nouveaux outils d'interprétation des données de WGS et de l'amélioration des prédictions, dans le but de réduire le plus finement possible et avec la meilleure sensibilité possible le nombre de variants réellement candidats sur lesquels nous devrions nous pencher avec une attention particulière. A l'avenir, on pourrait espérer un accès concomitant à des études complémentaires en RNAseq ou protéomiques pour affiner en temps réel l'interprétation des variants complexes.



# Références bibliographiques

---

- Aarabi, M., Kessler, E., Madan-Khetarpal, S., Surti, U., Bellissimo, D., Rajkovic, A. et Yatsenko, S. A. (2019). Autism spectrum disorder in females with ARHGEF9 alterations and a random pattern of X chromosome inactivation. *European Journal of Medical Genetics*, 62(4), 239-242. <https://doi.org/10.1016/j.ejmg.2018.07.021>
- Abel, H. J., Larson, D. E., Regier, A. A., Chiang, C., Das, I., Kanchi, K. L., Layer, R. M., Neale, B. M., Salerno, W. J., Reeves, C., Buyske, S., Matise, T. C., Muzny, D. M., Zody, M. C., Lander, E. S., Dutcher, S. K., Stitzel, N. O. et Hall, I. M. (2020). Mapping and characterization of structural variation in 17,795 human genomes. *Nature*, 583(7814), 83-89. <https://doi.org/10.1038/s41586-020-2371-0>
- Abelson, J. F., Kwan, K. Y., O'Roak, B. J., Baek, D. Y., Stillman, A. A., Morgan, T. M., Mathews, C. A., Pauls, D. L., Rasin, M.-R., Gunel, M., Davis, N. R., Ercan-Sencicek, A. G., Guez, D. H., Spertus, J. A., Leckman, J. F., Dure, L. S. 4th, Kurlan, R., Singer, H. S., Gilbert, D. L., ... State, M. W. (2005). Sequence variants in SLITRK1 are associated with Tourette's syndrome. *Science (New York, N.Y.)*, 310(5746), 317-320. <https://doi.org/10.1126/science.1116502>
- Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Kondrashov, A. S. et Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature Methods*, 7(4), 248-249. <https://doi.org/10.1038/nmeth0410-248>
- Alber, M., Kalscheuer, V. M., Marco, E., Sherr, E., Lesca, G., Till, M., Gradek, G., Wiesener, A., Korenke, C., Mercier, S., Becker, F., Yamamoto, T., Scherer, S. W., Marshall, C. R., Walker, S., Dutta, U. R., Dalal, A. B., Suckow, V., Jamali, P., ... Minassian, B. A. (2017). ARHGEF9 disease: Phenotype clarification and genotype-phenotype correlation. *Neurology Genetics*, 3(3), e148. <https://doi.org/10.1212/NXG.000000000000148>
- Amberger, J. S., Bocchini, C. A., Schiettecatte, F., Scott, A. F. et Hamosh, A. (2015). OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders. *Nucleic Acids Research*, 43(Database issue), D789-798. <https://doi.org/10.1093/nar/gku1205>
- Amendola, L. M., Jarvik, G. P., Leo, M. C., McLaughlin, H. M., Akkari, Y., Amaral, M. D., Berg, J. S., Biswas, S., Bowling, K. M., Conlin, L. K., Cooper, G. M., Dorschner, M. O., Dulik, M. C., Ghazani, A. A., Ghosh, R., Green, R. C., Hart, R., Horton, C., Johnston, J. J., ... Rehm, H. L. (2016). Performance of ACMG-AMP Variant-Interpretation Guidelines among Nine Laboratories in the Clinical Sequencing Exploratory Research Consortium. *American Journal of Human Genetics*, 98(6), 1067-1076. <https://doi.org/10.1016/j.ajhg.2016.03.024>
- Amos-Landgraf, J. M., Cottle, A., Plenge, R. M., Friez, M., Schwartz, C. E., Longshore, J. et Willard, H. F. (2006). X chromosome-inactivation patterns of 1,005 phenotypically unaffected females. *American Journal of Human Genetics*, 79(3), 493-499. <https://doi.org/10.1086/507565>
- Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T., Ntini, E., Arner, E., Valen, E., Li, K., Schwarzfischer, L., Glatz, D., Raithel, J., Lilje, B., Rapin, N., ... Sandelin, A. (2014). An atlas of active enhancers across human cell types and tissues. *Nature*, 507(7493), 455-461. <https://doi.org/10.1038/nature12787>
- Aral, B., de Saint Basile, G., Al-Garawi, S., Kamoun, P. et Ceballos-Picot, I. (1996). Novel nonsense mutation in the hypoxanthine guanine phosphoribosyltransferase gene and nonrandom X-inactivation causing Lesch-Nyhan syndrome in a female patient. *Human Mutation*, 7(1), 52-58. [https://doi.org/10.1002/\(SICI\)1098-1004\(1996\)7:1<52::AID-HUMU7>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1098-1004(1996)7:1<52::AID-HUMU7>3.0.CO;2-R)
- Ars, E., Serra, E., Garcia, J., Kruyer, H., Gaona, A., Lazaro, C. et Estivill, X. (2000). Mutations affecting mRNA splicing are the most common molecular defects in patients with neurofibromatosis type 1. *Human Molecular Genetics*, 9(2), 237-247. <https://doi.org/10.1093/hmg/9.2.237>
- Aruga, J. et Mikoshiba, K. (2003). Identification and characterization of Slitrk, a novel neuronal transmembrane protein family controlling neurite outgrowth. *Molecular and Cellular Neurosciences*, 24(1), 117-129. [https://doi.org/10.1016/s1044-7431\(03\)00129-5](https://doi.org/10.1016/s1044-7431(03)00129-5)

- Attia, D., Lurie, A., Zhai, Q. et Smalridge, R. (2020). Case of aggressive metastatic follicular variant papillary thyroid carcinoma with BRAF K601E and BCORL1 mutations. *BMJ Case Reports*, 13(6), e234208. <https://doi.org/10.1136/bcr-2019-234208>
- Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., Korbel, J. O., Marchini, J. L., McCarthy, S., McVean, G. A. et Abecasis, G. R. (2015). A global reference for human genetic variation. *Nature*, 526(7571), 68-74. <https://doi.org/10.1038/nature15393>
- Azodi, M., Kamps, R., Heymans, S. et Robinson, E. L. (2020). The Missing « Inc » between Genetics and Cardiac Disease. *Non-Coding RNA*, 6(1), E3. <https://doi.org/10.3390/ncrna6010003>
- Balaton, B. P., Dixon-McDougall, T., Peeters, S. B. et Brown, C. J. (2018). The eXceptional nature of the X chromosome. *Human Molecular Genetics*, 27(R2), R242-R249. <https://doi.org/10.1093/hmg/ddy148>
- Bamshad, M. J., Nickerson, D. A. et Chong, J. X. (2019). Mendelian Gene Discovery: Fast and Furious with No End in Sight. *American Journal of Human Genetics*, 105(3), 448-455. <https://doi.org/10.1016/j.ajhg.2019.07.011>
- Baralle, D. et Baralle, M. (2005). Splicing in action: assessing disease causing sequence changes. *Journal of Medical Genetics*, 42(10), 737-748. <https://doi.org/10.1136/jmg.2004.029538>
- BARR, M. L. et BERTRAM, E. G. (1949). A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. *Nature*, 163(4148), 676. <https://doi.org/10.1038/163676a0>
- BARR, M. L., BERTRAM, L. F. et LINDSAY, H. A. (1950). The morphology of the nerve cell nucleus, according to sex. *The Anatomical Record*, 107(3), 283-297. <https://doi.org/10.1002/ar.1091070307>
- Battaglia, A. et Carey, J. C. (2021). The delineation of the Wolf-Hirschhorn syndrome over six decades: Illustration of the ongoing advances in phenotype analysis and cytogenomic technology. *American Journal of Medical Genetics. Part A*, 185(9), 2748-2755. <https://doi.org/10.1002/ajmg.a.62341>
- Belkadi, A., Bolze, A., Itan, Y., Cobat, A., Vincent, Q. B., Antipenko, A., Shang, L., Boisson, B., Casanova, J.-L. et Abel, L. (2015). Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. *Proceedings of the National Academy of Sciences of the United States of America*, 112(17), 5473-5478. <https://doi.org/10.1073/pnas.1418631112>
- Ben Khelifa, H., Soyah, N., Ben-Abdallah-Bouhjar, I., Gritly, R., Sanlaville, D., Elghezal, H., Saad, A. et Mougou-Zerelli, S. (2013). Xp22.3 interstitial deletion: a recognizable chromosomal abnormality encompassing VCX3A and STS genes in a patient with X-linked ichthyosis and mental retardation. *Gene*, 527(2), 578-583. <https://doi.org/10.1016/j.gene.2013.06.018>
- Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G., Hall, K. P., Evers, D. J., Barnes, C. L., Bignell, H. R., Boutell, J. M., Bryant, J., Carter, R. J., Keira Cheetham, R., Cox, A. J., Ellis, D. J., Flatbush, M. R., Gormley, N. A., Humphray, S. J., ... Smith, A. J. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, 456(7218), 53-59. <https://doi.org/10.1038/nature07517>
- Berletch, J. B., Yang, F. et Disteche, C. M. (2010). Escape from X inactivation in mice and humans. *Genome Biology*, 11(6), 213. <https://doi.org/10.1186/gb-2010-11-6-213>
- Bernstein, B. E., Stamatoyannopoulos, J. A., Costello, J. F., Ren, B., Milosavljevic, A., Meissner, A., Kellis, M., Marra, M. A., Beaudet, A. L., Ecker, J. R., Farnham, P. J., Hirst, M., Lander, E. S., Mikkelsen, T. S. et Thomson, J. A. (2010). The NIH Roadmap Epigenomics Mapping Consortium. *Nature biotechnology*, 28(10), 1045-1048. <https://doi.org/10.1038/nbt1010-1045>
- Biederer, T., Kaeser, P. S. et Blanpied, T. A. (2017). Transcellular Nanoalignment of Synaptic Function. *Neuron*, 96(3), 680-696. <https://doi.org/10.1016/j.neuron.2017.10.006>
- Bienvenu, T., des Portes, V., Saint Martin, A., McDonell, N., Billuart, P., Carrie, A., Vinet, M. C., Couvert, P., Toniolo, D., Ropers, H. H., Moraine, C., van Bokhoven, H., Fryns, J. P., Kahn, A., Beldjord, C. et Chelly, J. (1998). Non-specific X-linked semidominant mental retardation by mutations in a Rab GDP-dissociation inhibitor. *Human Molecular Genetics*, 7(8), 1311-1315. <https://doi.org/10.1093/hmg/7.8.1311>
- Bijlsma, E. K., Collins, A., Papa, F. T., Tejada, M. I., Wheeler, P., Peeters, E. A. J., Gijsbers, A. C. J., van de Kamp, J. M., Kriek, M., Losekoot, M., Broekma, A. J., Crolla, J. A., Pollazzon, M., Mucciolo, M., Katzaki, E., Disciglio, V., Ferreri, M. I., Marozza, A., Mencarelli, M. A., ... Ruivenkamp, C. A. L. (2012). Xq28 duplications including MECP2 in five females: Expanding the phenotype to severe mental retardation. *European Journal of Medical Genetics*, 55(6-7), 404-413. <https://doi.org/10.1016/j.ejmg.2012.02.009>
- Billuart, P., Bienvenu, T., Ronce, N., des Portes, V., Vinet, M. C., Zemni, R., Roest Crollius, H., Carrie, A., Fauchereau, F., Cherry, M., Briault, S., Hamel, B., Fryns, J. P., Beldjord, C., Kahn, A., Moraine, C. et Chelly, J. (1998).

- Oligophrenin-1 encodes a rhoGAP protein involved in X-linked mental retardation. *Nature*, 392(6679), 923-926. <https://doi.org/10.1038/31940>
- Binquet, C., Lejeune, C., Faivre, L., Bouctot, M., Asensio, M.-L., Simon, A., Deleuze, J.-F., Boland, A., Guillemin, F., Seror, V., Delmas, C., Esp rou, H., Duffourd, Y., Lyonnet, S., Odent, S., Heron, D., Sanlaville, D., Frebourg, T., Gerard, B. et Dollfus, H. (2022). Genome Sequencing for Genetics Diagnosis of Patients With Intellectual Disability: The DEFIDIAG Study. *Frontiers in Genetics*, 12, 766964. <https://doi.org/10.3389/fgene.2021.766964>
- Bogershausen, N., Gatinois, V., Riehmer, V., Kayserili, H., Becker, J., Thoenes, M., Simsek-Kiper, P. O., Barat-Houari, M., Elcioglu, N. H., Wieczorek, D., Tinschert, S., Sarrabay, G., Strom, T. M., Fabre, A., Baynam, G., Sanchez, E., Nurnberg, G., Altunoglu, U., Capri, Y., ... Wollnik, B. (2016). Mutation Update for Kabuki Syndrome Genes KMT2D and KDM6A and Further Delineation of X-Linked Kabuki Syndrome Subtype 2. *Human Mutation*, 37(9), 847-864. <https://doi.org/10.1002/humu.23026>
- Bonneau, D., Toutain, A., Laquerriere, A., Marret, S., Saugier-veber, P., Barthez, M.-A., Radi, S., Biran-Mucignat, V., Rodriguez, D. et Gelot, A. (2002). X-linked lissencephaly with absent corpus callosum and ambiguous genitalia (XLAG): clinical, magnetic resonance imaging, and neuropathological findings. *Annals of Neurology*, 51(3), 340-349. <https://doi.org/10.1002/ana.10119>
- Bonnet, C., Gregoire, M. J., Brochet, K., Raffo, E., Leheup, B. et Jonveaux, P. (2006). Pure de-novo 5 Mb duplication at Xp11.22-p11.23 in a male: phenotypic and molecular characterization. *Journal of Human Genetics*, 51(9), 815. <https://doi.org/10.1007/s10038-006-0023-3>
- Bonnet, C., Krieger, S., Vezain, M., Rousselin, A., Tournier, I., Martins, A., Berthet, P., Chevrier, A., Dugast, C., Layet, V., Rossi, A., Lidereau, R., Frebourg, T., Hardouin, A. et Tosi, M. (2008). Screening BRCA1 and BRCA2 unclassified variants for splicing mutations using reverse transcription PCR on patient RNA and an ex vivo assay based on a splicing reporter minigene. *Journal of Medical Genetics*, 45(7), 438-446. <https://doi.org/10.1136/jmg.2007.056895>
- Brand, B. A., Blesson, A. E. et Smith-Hicks, C. L. (2021). The Impact of X-Chromosome Inactivation on Phenotypic Expression of X-Linked Neurodevelopmental Disorders. *Brain Sciences*, 11(7), 904. <https://doi.org/10.3390/brainsci11070904>
- Brandler, W. M., Antaki, D., Gujral, M., Kleiber, M. L., Whitney, J., Maile, M. S., Hong, O., Chapman, T. R., Tan, S., Tandon, P., Pang, T., Tang, S. C., Vaux, K. K., Yang, Y., Harrington, E., Juul, S., Turner, D. J., Thiruvahindrapuram, B., Kaur, G., ... Sebat, J. (2018). Paternally inherited cis-regulatory structural variants are associated with autism. *Science (New York, N.Y.)*, 360(6386), 327-331. <https://doi.org/10.1126/science.aan2261>
- Brody, E. et Abelson, J. (1985). The « spliceosome »: yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction. *Science (New York, N.Y.)*, 228(4702), 963-967. <https://doi.org/10.1126/science.3890181>
- Brunner, A. L., Johnson, D. S., Kim, S. W., Valouev, A., Reddy, T. E., Neff, N. F., Anton, E., Medina, C., Nguyen, L., Chiao, E., Oyolu, C. B., Schroth, G. P., Absher, D. M., Baker, J. C. et Myers, R. M. (2009). Distinct DNA methylation patterns characterize differentiated human embryonic stem cells and developing human fetal liver. *Genome Research*, 19(6), 1044-1056. <https://doi.org/10.1101/gr.088773.108>
- Buntinx, W., Cans, C., Colleaux, L., Courbois, Y., Debban , M., Desportes, V., Detraux, J.-J., Facon, B., Haelewyck, M.-C., Heron, D., Petitpierre, G. et Plaisance, E. (s. d.). D ficiences intellectuelles.
- Burrows, J. T. A., Pearson, B. J. et Scott, I. C. (2015). An In Vivo Requirement for the Mediator Subunit Med14 in the Maintenance of Stem Cell Populations. *Stem Cell Reports*, 4(4), 670-684. <https://doi.org/10.1016/j.stemcr.2015.02.006>
- Caceres, J. F. et Kornblihtt, A. R. (2002). Alternative splicing: multiple control mechanisms and involvement in human disease. *Trends in Genetics : TIG*, 18(4), 186-193. [https://doi.org/10.1016/s0168-9525\(01\)02626-9](https://doi.org/10.1016/s0168-9525(01)02626-9)
- Cameron, D. L., Schroder, J., Penington, J. S., Do, H., Molania, R., Dobrovic, A., Speed, T. P. et Papenfuss, A. T. (2017). GRIDSS: sensitive and specific genomic rearrangement detection using positional de Bruijn graph assembly. *Genome Research*, 27(12), 2050-2060. <https://doi.org/10.1101/gr.222109.117>
- Cantagrel, V., Silhavy, J. L., Bielas, S. L., Swistun, D., Marsh, S. E., Bertrand, J. Y., Audollent, S., Attie-Bitach, T., Holden, K. R., Dobyys, W. B., Traver, D., Al-Gazali, L., Ali, B. R., Lindner, T. H., Caspary, T., Otto, E. A., Hildebrandt, F., Glass, I. A., Logan, C. V., ... Gleeson, J. G. (2008). Mutations in the cilia gene ARL13B lead to the classical form of Joubert syndrome. *American Journal of Human Genetics*, 83(2), 170-179. <https://doi.org/10.1016/j.ajhg.2008.06.023>



- Carmignac, V., Nambot, S., Lehalle, D., Callier, P., Moortgat, S., Benoit, V., Ghoumid, J., Delobel, B., Smol, T., Thuillier, C., Zordan, C., Naudion, S., Bienvenu, T., Touraine, R., Ramond, F., Zweier, C., Reis, A., Kraus, C., Nizon, M., ... Thauvin-Robinet, C. (2020). Further delineation of the female phenotype with KDM5C disease causing variants: 19 new individuals and review of the literature. *Clinical Genetics*, *98*(1), 43-55. <https://doi.org/10.1111/cge.13755>
- Caron, B., Luo, Y. et Rausell, A. (2019). NCBoost classifies pathogenic non-coding variants in Mendelian diseases through supervised learning on purifying selection signals in humans. *Genome Biology*, *20*(1), 32. <https://doi.org/10.1186/s13059-019-1634-2>
- Carrel, L., Cottle, A. A., Goglin, K. C. et Willard, H. F. (1999). A first-generation X-inactivation profile of the human X chromosome. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(25), 14440-14444. <https://doi.org/10.1073/pnas.96.25.14440>
- Cereda, A. et Carey, J. C. (2012). The trisomy 18 syndrome. *Orphanet Journal of Rare Diseases*, *7*, 81. <https://doi.org/10.1186/1750-1172-7-81>
- Cevher, M. A., Shi, Y., Li, D., Chait, B. T., Malik, S. et Roeder, R. G. (2014). Reconstitution of active human core Mediator complex reveals a critical role of the MED14 subunit. *Nature Structural & Molecular Biology*, *21*(12), 1028-1034. <https://doi.org/10.1038/nsmb.2914>
- Chang, H., Zhang, X., Xu, K., Li, N., Xie, Y., Yan, W. et Li, Y. (2023). Phenotype-Based Genetic Analysis Reveals Missing Heritability of KIF11-Related Retinopathy: Clinical and Genetic Findings. *Genes*, *14*(1), 212. <https://doi.org/10.3390/genes14010212>
- Chen, W., Ullmann, R., Langnick, C., Menzel, C., Wotschovsky, Z., Hu, H., Doring, A., Hu, Y., Kang, H., Tzschach, A., Hoeltzenbein, M., Neitzel, H., Markus, S., Wiedersberg, E., Kistner, G., van Ravenswaaij-Arts, C. M. A., Kleefstra, T., Kalscheuer, V. M. et Ropers, H.-H. (2010). Breakpoint analysis of balanced chromosome rearrangements by next-generation paired-end sequencing. *European Journal of Human Genetics : EJHG*, *18*(5), 539-543. <https://doi.org/10.1038/ejhg.2009.211>
- Chen, X., Schulz-Trieglaff, O., Shaw, R., Barnes, B., Schlesinger, F., Kallberg, M., Cox, A. J., Kruglyak, S. et Saunders, C. T. (2016). Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics (Oxford, England)*, *32*(8), 1220-1222. <https://doi.org/10.1093/bioinformatics/btv710>
- Chérot, E., Keren, B., Dubourg, C., Carré, W., Fradin, M., Lavillaureix, A., Afenjar, A., Burglen, L., Whalen, S., Charles, P., Marey, I., Heide, S., Jacqueline, A., Heron, D., Doummar, D., Rodriguez, D., Billette de Villemeur, T., Moutard, M.-L., Guët, A., ... Mignot, C. (2018). Using medical exome sequencing to identify the causes of neurodevelopmental disorders: Experience of 2 clinical units and 216 patients. *Clinical Genetics*, *93*(3), 567-576. <https://doi.org/10.1111/cge.13102>
- Cheung, R., Insigne, K. D., Yao, D., Burghard, C. P., Wang, J., Hsiao, Y.-H. E., Jones, E. M., Goodman, D. B., Xiao, X. et Kosuri, S. (2019). A Multiplexed Assay for Exon Recognition Reveals that an Unappreciated Fraction of Rare Genetic Variants Cause Large-Effect Splicing Disruptions. *Molecular Cell*, *73*(1), 183-194.e8. <https://doi.org/10.1016/j.molcel.2018.10.037>
- Chiurazzi, P. et Pirozzi, F. (2016). Advances in understanding - genetic basis of intellectual disability. *F1000Research*, *5*, F1000 Faculty Rev-599. <https://doi.org/10.12688/f1000research.7134.1>
- Chong, J. X., Buckingham, K. J., Jhangiani, S. N., Boehm, C., Sobreira, N., Smith, J. D., Harrell, T. M., McMillin, M. J., Wiszniewski, W., Gambin, T., Coban Akdemir, Z. H., Doheny, K., Scott, A. F., Avramopoulos, D., Chakravarti, A., Hoover-Fong, J., Mathews, D., Witmer, P. D., Ling, H., ... Bamshad, M. J. (2015). The Genetic Basis of Mendelian Phenotypes: Discoveries, Challenges, and Opportunities. *American Journal of Human Genetics*, *97*(2), 199-215. <https://doi.org/10.1016/j.ajhg.2015.06.009>
- Chung, C. C. Y., Leung, G. K. C., Mak, C. C. Y., Fung, J. L. F., Lee, M., Pei, S. L. C., Yu, M. H. C., Hui, V. C. C., Chan, J. C. K., Chau, J. F. T., Chan, M. C. Y., Tsang, M. H. Y., Wong, W. H. S., Tung, J. Y. L., Lun, K. S., Ng, Y. K., Fung, C. W., Wong, M. S. C., Wong, R. M. S., ... Chung, B. H. Y. (2020). Rapid whole-exome sequencing facilitates precision medicine in paediatric rare disease patients and reduces healthcare costs. *The Lancet Regional Health. Western Pacific*, *1*, 100001. <https://doi.org/10.1016/j.lanwpc.2020.100001>
- Church, G. M. (2006). Genomes for all. *Scientific American*, *294*(1), 46-54. <https://doi.org/10.1038/scientificamerican0106-46>
- Coccaro, N., Anelli, L., Zagaria, A., Tarantini, F., Cumbo, C., Tota, G., Minervini, C. F., Minervini, A., Conserva, M. R., Redavid, I., Parciante, E., Macchia, M. G., Specchia, G., Musto, P. et Albano, F. (2023). Feasibility of Optical Genome Mapping in Cytogenetic Diagnostics of Hematological Neoplasms: A New Way to Look at DNA. *Diagnostics (Basel, Switzerland)*, *13*(11), 1841. <https://doi.org/10.3390/diagnostics13111841>

- Cock, P. J. A., Fields, C. J., Goto, N., Heuer, M. L. et Rice, P. M. (2010). The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Research*, 38(6), 1767-1771. <https://doi.org/10.1093/nar/gkp1137>
- Coe, B. P., Girirajan, S. et Eichler, E. E. (2012). The genetic variability and commonality of neurodevelopmental disease. *American Journal of Medical Genetics. Part C, Seminars in Medical Genetics*, 160C(2), 118-129. <https://doi.org/10.1002/ajmg.c.31327>
- Corbett, M. A., Depienne, C., Veneziano, L., Klein, K. M., Brancati, F., Guerrini, R., Zara, F., Tsuji, S. et Gecz, J. (2023). Genetics of familial adult myoclonus epilepsy: From linkage studies to noncoding repeat expansions. *Epilepsia*, 64 Suppl 1, S14-S21. <https://doi.org/10.1111/epi.17610>
- Corcos, L. et Solier, S. (2005). [Alternative mRNA splicing, pathology and molecular therapeutics]. *Medecine sciences : M/S*, 21(3), 253-260. <https://doi.org/10.1051/medsci/2005213253>
- Cordaux, R. et Batzer, M. A. (2009). The impact of retrotransposons on human genome evolution. *Nature Reviews. Genetics*, 10(10), 691-703. <https://doi.org/10.1038/nrg2640>
- Corominas, J., Smeekens, S. P., Nelen, M. R., Yntema, H. G., Kamsteeg, E.-J., Pfundt, R. et Gilissen, C. (2022). Clinical exome sequencing-Mistakes and caveats. *Human Mutation*, 43(8), 1041-1055. <https://doi.org/10.1002/humu.24360>
- Coskun, F., Percin-Pacal, F., Emrence, Z., Ar, M., Abaci, N., Unuvar, A., Eskazan, A., Elverdi, T., Salihoglu, A., Seflekci, Y., Karakas, Z., Soysal, T. et Sirma-Ekmekci, S. (2020). Age-Related Co-Expression of BCOR and BCORL1 mRNA in Acute Myeloid Leukemia. *Clinical Laboratory*, 66(08/2020). <https://doi.org/10.7754/Clin.Lab.2020.191119>
- Courraud, Jeremie, Chater-Diehl, E., Durand, B., Vincent, M., Del Mar Muniz Moreno, M., Boujelbene, I., Drouot, N., Genschik, L., Schaefer, E., Nizon, M., Gerard, B., Abramowicz, M., Cogne, B., Bronicki, L., Burglen, L., Barth, M., Charles, P., Colin, E., Coubes, C., ... Piton, A. (2021). Integrative approach to interpret DYRK1A variants, leading to a frequent neurodevelopmental disorder. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, 23(11), 2150-2159. <https://doi.org/10.1038/s41436-021-01263-1>
- Courraud, Jérémie, Engel, C., Quartier, A., Drouot, N., Houessou, U., Plassard, D., Sorlin, A., Brischoux-Boucher, E., Maldergem, L. V., Gouy, E., Rossi, M., Edery, P., Putoux, A., Gilbert-Dussardier, B., Kalscheuer, V., Mandel, J.-L. et Piton, A. (2022, 30 mai). Molecular consequences of PQBP1 deficiency, involved in the X-linked Renpenning syndrome. *bioRxiv*. <https://doi.org/10.1101/2022.05.29.493091>
- Coursimault, J., Cassinari, K., Lecoquierre, F., Quenez, O., Coutant, S., Derambure, C., Vezain, M., Drouot, N., Vera, G., Schaefer, E., Philippe, A., Doray, B., Lambert, L., Ghoumid, J., Smol, T., Rama, M., Legendre, M., Lacombe, D., Fergelot, P., ... Nicolas, G. (2022). Deep intronic NIPBL de novo mutations and differential diagnoses revealed by whole genome and RNA sequencing in Cornelia de Lange syndrome patients. *Human Mutation*, 43(12), 1882-1897. <https://doi.org/10.1002/humu.24438>
- Crawford, D. C., Acuna, J. M. et Sherman, S. L. (2001). FMR1 and the fragile X syndrome: human genome epidemiology review. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, 3(5), 359-371. <https://doi.org/10.1097/00125817-200109000-00006>
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G. et Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics (Oxford, England)*, 27(15), 2156-2158. <https://doi.org/10.1093/bioinformatics/btr330>
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M. et Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, 10(2), giab008. <https://doi.org/10.1093/gigascience/giab008>
- David, M., Dieterich, K., Billette de Villemeur, A., Jouk, P.-S., Counillon, J., Larroque, B., Bloch, J. et Cans, C. (2014). Prevalence and characteristics of children with mild intellectual disability in a French county. *Journal of Intellectual Disability Research: JIDR*, 58(7), 591-602. <https://doi.org/10.1111/jir.12057>
- de Bruijn, S. E., Rodenburg, K., Corominas, J., Ben-Yosef, T., Reurink, J., Kremer, H., Whelan, L., Plomp, A. S., Berger, W., Farrar, G. J., Ferenc Kovacs, A., Fajardy, I., Hitti-Malin, R. J., Weisschuh, N., Weener, M. E., Sharon, D., Pennings, R. J. E., Haer-Wigman, L., Hoyng, C. B., ... Roosing, S. (2023). Optical genome mapping and revisiting short-read genome sequencing data reveal previously overlooked structural variants disrupting retinal disease-associated genes. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, 25(3), 100345. <https://doi.org/10.1016/j.gim.2022.11.013>
- de Koning, A. P. J., Gu, W., Castoe, T. A., Batzer, M. A. et Pollock, D. D. (2011). Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genetics*, 7(12), e1002384. <https://doi.org/10.1371/journal.pgen.1002384>

- de Ligt, J., Willemsen, M. H., van Bon, B. W. M., Kleefstra, T., Yntema, H. G., Kroes, T., Vulto-van Silfhout, A. T., Koolen, D. A., de Vries, P., Gilissen, C., del Rosario, M., Hoischen, A., Scheffer, H., de Vries, B. B. A., Brunner, H. G., Veltman, J. A. et Vissers, L. E. L. M. (2012). Diagnostic exome sequencing in persons with severe intellectual disability. *The New England Journal of Medicine*, 367(20), 1921-1929. <https://doi.org/10.1056/NEJMoa1206524>
- de Lima, R. L. L. F., Hoper, S. A., Ghassibe, M., Cooper, M. E., Rorick, N. K., Kondo, S., Katz, L., Marazita, M. L., Compton, J., Bale, S., Hehr, U., Dixon, M. J., Daack-Hirsch, S., Boute, O., Bayet, B., Revencu, N., Verellen-Dumoulin, C., Vikkula, M., Richieri-Costa, A., ... Schutte, B. C. (2009). Prevalence and nonrandom distribution of exonic mutations in interferon regulatory factor 6 in 307 families with Van der Woude syndrome and 37 families with popliteal pterygium syndrome. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, 11(4), 241-247. <https://doi.org/10.1097/GIM.0b013e318197a49a>
- Delvallee, C. et Dollfus, H. (2023). Retinal Degeneration Animal Models in Bardet-Biedl Syndrome and Related Ciliopathies. *Cold Spring Harbor Perspectives in Medicine*, 13(1), a041303. <https://doi.org/10.1101/cshperspect.a041303>
- Delvallee, C., Nicaise, S., Antin, M., Leuvre, A.-S., Nourisson, E., Leitch, C. C., Kellaris, G., Stoetzel, C., Geoffroy, V., Scheidecker, S., Keren, B., Depienne, C., Klar, J., Dahl, N., Deleuze, J.-F., Genin, E., Redon, R., Demurger, F., Devriendt, K., ... Muller, J. (2021). A BBS1 SVA F retrotransposon insertion is a frequent cause of Bardet-Biedl syndrome. *Clinical Genetics*, 99(2), 318-324. <https://doi.org/10.1111/cge.13878>
- Denomme-Pichon, A.-S., Matalonga, L., de Boer, E., Jackson, A., Benetti, E., Banka, S., Bruel, A.-L., Ciolfi, A., Clayton-Smith, J., Dallapiccola, B., Duffourd, Y., Ellwanger, K., Fallerini, C., Gilissen, C., Graessner, H., Haack, T. B., Havlovicova, M., Hoischen, A., Jean-Marcais, N., ... Faivre, L. (2023). A Solve-RD ClinVar-based reanalysis of 1522 index cases from ERN-ITHACA reveals common pitfalls and misinterpretations in exome sequencing. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, 25(4), 100018. <https://doi.org/10.1016/j.gim.2023.100018>
- Depienne, C. et LeGuern, E. (2012). PCDH19-related infantile epileptic encephalopathy: an unusual X-linked inheritance disorder. *Human Mutation*, 33(4), 627-634. <https://doi.org/10.1002/humu.22029>
- Depienne, C. et Mandel, J.-L. (2021). 30 years of repeat expansion disorders: What have we learned and what are the remaining challenges? *American Journal of Human Genetics*, 108(5), 764-785. <https://doi.org/10.1016/j.ajhg.2021.03.011>
- DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis, A. A., del Angel, G., Rivas, M. A., Hanna, M., McKenna, A., Fennell, T. J., Kernysky, A. M., Sivachenko, A. Y., Cibulskis, K., Gabriel, S. B., Altshuler, D. et Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, 43(5), 491-498. <https://doi.org/10.1038/ng.806>
- Des Portes, V. et Héron, D. (2020). Troubles du développement intellectuel. *Contraste*, 51(1), 91-117. <https://doi.org/10.3917/cont.051.0091>
- Desmet, F.-O., Hamroun, D., Lalande, M., Collod-Beroud, G., Claustres, M. et Beroud, C. (2009). Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Research*, 37(9), e67. <https://doi.org/10.1093/nar/gkp215>
- de Souza, N. (2012). The ENCODE project. *Nature Methods*, 9(11), 1046-1046. <https://doi.org/10.1038/nmeth.2238>
- des Portes, V. (2020). Chapter 9 - Intellectual disability. Dans A. Gallagher, C. Bulteau, D. Cohen et J. L. Michaud (dir.), *Handbook of Clinical Neurology* (vol. 174, p. 113-126). Elsevier. <https://doi.org/10.1016/B978-0-444-64148-9.00009-0>
- Devanna, P., Chen, X. S., Ho, J., Gajewski, D., Smith, S. D., Gialluisi, A., Francks, C., Fisher, S. E., Newbury, D. F. et Vernes, S. C. (2018). Next-gen sequencing identifies non-coding variation disrupting miRNA-binding sites in neurological disorders. *Molecular Psychiatry*, 23(5), 1375-1384. <https://doi.org/10.1038/mp.2017.30>
- Devanna, Paolo, van de Vorst, M., Pfundt, R., Gilissen, C. et Vernes, S. C. (2018). Genome-wide investigation of an ID cohort reveals de novo 3'UTR variants affecting gene expression. *Human Genetics*, 137(9), 717-721. <https://doi.org/10.1007/s00439-018-1925-9>
- Di Giacomo, D., Gaildrat, P., Abuli, A., Abdat, J., Frebourg, T., Tosi, M. et Martins, A. (2013). Functional analysis of a large set of BRCA2 exon 7 variants highlights the predictive value of hexamer scores in detecting alterations of exonic splicing regulatory elements. *Human Mutation*, 34(11), 1547-1557. <https://doi.org/10.1002/humu.22428>

- Di Vona, C., Bezdan, D., Islam, A. B. M. M. K., Salichs, E., Lopez-Bigas, N., Ossowski, S. et de la Luna, S. (2015). Chromatin-wide profiling of DYRK1A reveals a role as a gene-specific RNA polymerase II CTD kinase. *Molecular Cell*, 57(3), 506-520. <https://doi.org/10.1016/j.molcel.2014.12.026>
- di Blasio, L., Puliafito, A., Gagliardi, P. A., Comunanza, V., Somale, D., Chiaverina, G., Bussolino, F. et Primo, L. (2018). PI3K/mTOR inhibition promotes the regression of experimental vascular malformations driven by PIK3CA-activating mutations. *Cell Death & Disease*, 9(2), 45. <https://doi.org/10.1038/s41419-017-0064-x>
- Dobyns, W. B., Filauro, A., Tomson, B. N., Chan, A. S., Ho, A. W., Ting, N. T., Oosterwijk, J. C. et Ober, C. (2004). Inheritance of most X-linked traits is not dominant or recessive, just X-linked. *American Journal of Medical Genetics. Part A*, 129A(2), 136-143. <https://doi.org/10.1002/ajmg.a.30123>
- Doi, K., Monjo, T., Hoang, P. H., Yoshimura, J., Yurino, H., Mitsui, J., Ishiura, H., Takahashi, Y., Ichikawa, Y., Goto, J., Tsuji, S. et Morishita, S. (2014). Rapid detection of expanded short tandem repeats in personal genomics using hybrid sequencing. *Bioinformatics (Oxford, England)*, 30(6), 815-822. <https://doi.org/10.1093/bioinformatics/btt647>
- Dolzhenko, E., Bennett, M. F., Richmond, P. A., Trost, B., Chen, S., van Vugt, J. J. F. A., Nguyen, C., Narzisi, G., Gainullin, V. G., Gross, A. M., Lajoie, B. R., Taft, R. J., Wasserman, W. W., Scherer, S. W., Veldink, J. H., Bentley, D. R., Yuen, R. K. C., Bahlo, M. et Eberle, M. A. (2020). ExpansionHunter Denovo: a computational method for locating known and novel repeat expansions in short-read sequencing data. *Genome Biology*, 21(1), 102. <https://doi.org/10.1186/s13059-020-02017-z>
- Dujardin, G., Daguene, E., Bernard, D. G., Flodrops, M., Durand, S., Chauveau, A., El Khoury, F., Le Jossic-Corcus, C. et Corcos, L. (2016). [Pre-mRNA splicing: when the spliceosome loses ground]. *Medecine sciences : M/S*, 32(12), 1103-1110. <https://doi.org/10.1051/medsci/20163212014>
- Dulac, C. (2010). Brain function and chromatin plasticity. *Nature*, 465(7299), 728-735. <https://doi.org/10.1038/nature09231>
- Dulbecco, R. (1986). A turning point in cancer research: sequencing the human genome. *Science (New York, N.Y.)*, 231(4742), 1055-1056. <https://doi.org/10.1126/science.3945817>
- Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F., Epstein, C. B., Frietze, S., Harrow, J., Kaul, R., Khatun, J., Lajoie, B. R., Landt, S. G., Lee, B.-K., Pauli, F., Rosenbloom, K. R., Sabo, P., Safi, A., Sanyal, A., ... HudsonAlpha Institute, C., UC Irvine, Stanford group (data production and analysis). (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, 489(7414), 57-74. <https://doi.org/10.1038/nature11247>
- Durand, B., Stoetzel, C., Schaefer, E., Calmels, N., Scheidecker, S., Kempf, N., De Melo, C., Guilbert, A.-S., Timbolschi, D., Donato, L., Astruc, D., Sauer, A., Antal, M. C., Dollfus, H. et El Chehadeh, S. (2020). A severe case of Frank-ter Haar syndrome and literature review: Further delineation of the phenotypical spectrum. *European Journal of Medical Genetics*, 63(4), 103857. <https://doi.org/10.1016/j.ejmg.2020.103857>
- Dutta, U. R., Rao, S. N., Pidugu, V. K., V S, V., Bhattacharjee, A., Bhowmik, A. D., Ramaswamy, S. K., Singh, K. G. et Dalal, A. (2019). Breakpoint mapping of a novel de novo translocation t(X;20)(q11.1;p13) by positional cloning and long read sequencing. *Genomics*, 111(5), 1108-1114. <https://doi.org/10.1016/j.ygeno.2018.07.005>
- Edens, A. C., Lyons, M. J., Duron, R. M., Dupont, B. R. et Holden, K. R. (2011). Autism in two females with duplications involving Xp11.22-p11.23. *Developmental Medicine and Child Neurology*, 53(5), 463-466. <https://doi.org/10.1111/j.1469-8749.2010.03909.x>
- El Chehadeh, S., Touraine, R., Prieur, F., Reardon, W., Bienvenu, T., Chantot-Bastaraud, S., Doco-Fenzy, M., Landais, E., Philippe, C., Marle, N., Callier, P., Mosca-Boidron, A.-L., Mugneret, F., Le Meur, N., Goldenberg, A., Guerrot, A.-M., Chambon, P., Satre, V., Coutton, C., ... Faivre, L. (2017). Xq28 duplication including MECP2 in six unreported affected females: what can we learn for diagnosis and genetic counselling?: Xq28 duplication including MECP2 in six unreported affected females. *Clinical Genetics*, 91(4), 576-588. <https://doi.org/10.1111/cge.12898>
- El Chehadeh, Salima, Faivre, L., Mosca-Boidron, A.-L., Malan, V., Amiel, J., Nizon, M., Touraine, R., Prieur, F., Pasquier, L., Callier, P., Lefebvre, M., Marle, N., Dubourg, C., Julia, S., Sarret, C., Francannet, C., Laffargue, F., Boespflug-Tanguy, O., David, A., ... Guibaud, L. (2016). Large national series of patients with Xq28 duplication involving MECP2 : Delineation of brain MRI abnormalities in 30 affected patients. *American Journal of Medical Genetics Part A*, 170(1), 116-129. <https://doi.org/10.1002/ajmg.a.37384>
- El Chehadeh, Salima, Han, K. A., Kim, D., Jang, G., Bakhtiari, S., Lim, D., Kim, H. Y., Kim, J., Kim, H., Wynn, J., Chung, W. K., Vitiello, G., Cutcutache, I., Page, M., Gecz, J., Harper, K., Han, A.-R., Kim, H. M., Wessels, M., ... Um, J. W. (2022). SLITRK2 variants associated with neurodevelopmental disorders impair excitatory synaptic

- function and cognition in mice. *Nature Communications*, 13(1), 4112. <https://doi.org/10.1038/s41467-022-31566-z>
- El Chehadeh, Salima, Legrand, A., Stoetzel, C., Geoffroy, V., Billon, C., Adham, S., Jeunemaitre, X., Jaussaud, R., Muller, J., Schaefer, E., Benistan, K., Gaertner, S., Bloch-Zupan, A., Courval, A., Maniere, M.-C., Petit, C., Bursztejn, A.-C., Bal, L., Reyre, A., ... Lipsker, D. (2021). Periodontal (formerly type VIII) Ehlers-Danlos syndrome: Description of 13 novel cases and expansion of the clinical phenotype. *Clinical Genetics*, 100(2), 206-212. <https://doi.org/10.1111/cge.13972>
- El Chehadeh-Djebbar, S., Faivre, L., Moncla, A., Aral, B., Missirian, C., Popovici, C., Rump, P., Van Essen, A., Frances, A.-M., Gigot, N., Cusin, V., Masurel-Paulet, A., Gueneau, L., Payet, M., Ragon, C., Marle, N., Mosca-Boidron, A.-L., Huet, F., Balikova, I., ... Callier, P. (2011). The power of high-resolution non-targeted array-CGH in identifying intragenic rearrangements responsible for Cohen syndrome. *Journal of Medical Genetics*, 48(11), e1. <https://doi.org/10.1136/jmg.2011.088948>
- Enomoto, Y., Yokoi, T., Tsurusaki, Y., Murakami, H., Tominaga, M., Minatogawa, M., Abe-Hatano, C., Kuroda, Y., Ohashi, I., Ida, K., Shiiya, S., Kumaki, T., Naruto, T., Mitsui, J., Harada, N., Kido, Y. et Kurosawa, K. (2022). Divergent variant patterns among 19 patients with Rubinstein-Taybi syndrome uncovered by comprehensive genetic analysis including whole genome sequencing. *Clinical Genetics*, 101(3), 335-345. <https://doi.org/10.1111/cge.14103>
- Evers, C., Mitter, D., Strobl-Wildemann, G., Haug, U., Hackmann, K., Maas, B., Janssen, J. W. G., Jauch, A., Hinderhofer, K. et Moog, U. (2015). Duplication Xp11.22-p14 in females: does X-inactivation help in assessing their significance? *American Journal of Medical Genetics. Part A*, 167A(3), 553-562. <https://doi.org/10.1002/ajmg.a.36897>
- Field, M., Tarpey, P., Boyle, J., Edkins, S., Goodship, J., Luo, Y., Moon, J., Teague, J., Stratton, M. R., Futreal, P. A., Wooster, R., Raymond, F. L. et Turner, G. (2006). Mutations in the RSK2(RPS6KA3) gene cause Coffin-Lowry syndrome and nonsyndromic X-linked mental retardation. *Clinical Genetics*, 70(6), 509-515. <https://doi.org/10.1111/j.1399-0004.2006.00723.x>
- Field, Michael, Tarpey, P. S., Smith, R., Edkins, S., O'Meara, S., Stevens, C., Tofts, C., Teague, J., Butler, A., Dicks, E., Barthorpe, S., Buck, G., Cole, J., Gray, K., Halliday, K., Hills, K., Jenkinson, A., Jones, D., Menzies, A., ... Raymond, F. L. (2007). Mutations in the BRWD3 gene cause X-linked mental retardation associated with macrocephaly. *American Journal of Human Genetics*, 81(2), 367-374. <https://doi.org/10.1086/520677>
- Fieremans, N., Bauters, M., Belet, S., Verbeeck, J., Jansen, A. C., Seneca, S., Roelens, F., De Baere, E., Marynen, P. et Froyen, G. (2014). De novo MECP2 duplications in two females with intellectual disability and unfavorable complete skewed X-inactivation. *Human Genetics*, 133(11), 1359-1367. <https://doi.org/10.1007/s00439-014-1469-6>
- Finishing the euchromatic sequence of the human genome* | *Nature*. (s. d.). <https://www.nature.com/articles/nature03001>
- Fishilevich, S., Nudel, R., Rappaport, N., Hadar, R., Plaschkes, I., Iny Stein, T., Rosen, N., Kohn, A., Twik, M., Safran, M., Lancet, D. et Cohen, D. (2017). GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. *Database: The Journal of Biological Databases and Curation*, 2017. <https://doi.org/10.1093/database/bax028>
- Forbes, T. A., Howden, S. E., Lawlor, K., Phipson, B., Maksimovic, J., Hale, L., Wilson, S., Quinlan, C., Ho, G., Holman, K., Bennetts, B., Crawford, J., Trnka, P., Oshlack, A., Patel, C., Mallett, A., Simons, C. et Little, M. H. (2018). Patient-iPSC-Derived Kidney Organoids Show Functional Validation of a Ciliopathic Renal Phenotype and Reveal Underlying Pathogenetic Mechanisms. *American Journal of Human Genetics*, 102(5), 816-831. <https://doi.org/10.1016/j.ajhg.2018.03.014>
- Frankish, A., Diekhans, M., Ferreira, A.-M., Johnson, R., Jungreis, I., Loveland, J., Mudge, J. M., Sisu, C., Wright, J., Armstrong, J., Barnes, I., Berry, A., Bignell, A., Carbonell Sala, S., Chrast, J., Cunningham, F., Di Domenico, T., Donaldson, S., Fiddes, I. T., ... Flicek, P. (2019). GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Research*, 47(D1), D766-D773. <https://doi.org/10.1093/nar/gky955>
- Friendewey, D. et Keller, W. (1985). Stepwise assembly of a pre-mRNA splicing complex requires U-snRNPs and specific intron sequences. *Cell*, 42(1), 355-367. [https://doi.org/10.1016/s0092-8674\(85\)80131-8](https://doi.org/10.1016/s0092-8674(85)80131-8)
- Frezal, J. (1998). Genatlas database, genes and development defects. *Comptes Rendus de l'Academie Des Sciences. Serie III, Sciences de La Vie*, 321(10), 805-817. [https://doi.org/10.1016/s0764-4469\(99\)80021-3](https://doi.org/10.1016/s0764-4469(99)80021-3)
- Fu, Y. H., Kuhl, D. P., Pizzuti, A., Pieretti, M., Sutcliffe, J. S., Richards, S., Verkerk, A. J., Holden, J. J., Fenwick, R. G. J. et Warren, S. T. (1991). Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell*, 67(6), 1047-1058. [https://doi.org/10.1016/0092-8674\(91\)90283-5](https://doi.org/10.1016/0092-8674(91)90283-5)

- Fung, J. L. F., Yu, M. H. C., Huang, S., Chung, C. C. Y., Chan, M. C. Y., Pajusalu, S., Mak, C. C. Y., Hui, V. C. C., Tsang, M. H. Y., Yeung, K. S., Lek, M. et Chung, B. H. Y. (2020). A three-year follow-up study evaluating clinical utility of exome sequencing and diagnostic potential of reanalysis. *NPJ Genomic Medicine*, 5(1), 37. <https://doi.org/10.1038/s41525-020-00144-x>
- Galupa, R. et Heard, E. (2018). X-Chromosome Inactivation: A Crossroads Between Chromosome Architecture and Gene Regulation. *Annual Review of Genetics*, 52, 535-566. <https://doi.org/10.1146/annurev-genet-120116-024611>
- Garcia, C. C., Blair, H. J., Seager, M., Coulthard, A., Tennant, S., Buddles, M., Curtis, A. et Goodship, J. A. (2004). Identification of a mutation in synapsin I, a synaptic vesicle protein, in a family with epilepsy. *Journal of Medical Genetics*, 41(3), 183-186. <https://doi.org/10.1136/jmg.2003.013680>
- Gardner, E. J., Prigmore, E., Gallone, G., Danecek, P., Samocha, K. E., Handsaker, J., Gerety, S. S., Ironfield, H., Short, P. J., Sifrim, A., Singh, T., Chandler, K. E., Clement, E., Lachlan, K. L., Prescott, K., Rosser, E., FitzPatrick, D. R., Firth, H. V. et Hurles, M. E. (2019). Contribution of retrotransposition to developmental disorders. *Nature Communications*, 10(1), 4630. <https://doi.org/10.1038/s41467-019-12520-y>
- Garieri, M., Stamoulis, G., Blanc, X., Falconnet, E., Ribaux, P., Borel, C., Santoni, F. et Antonarakis, S. E. (2018). Extensive cellular heterogeneity of X inactivation revealed by single-cell allele-specific expression in human fibroblasts. *Proceedings of the National Academy of Sciences of the United States of America*, 115(51), 13015-13020. <https://doi.org/10.1073/pnas.1806811115>
- Garret, P., Chevarin, M., Vitobello, A., Verdez, S., Fournier, C., Verloes, A., Tisserant, E., Vabres, P., Prevel, O., Philippe, C., Denommé-Pichon, A.-S., Bruel, A.-L., Mau-Them, F. T., Safradou, H., Boughalem, A., Costa, J.-M., Trost, D., Thauvin-Robinet, C., Faivre, L. et Duffourd, Y. (2023). A second look at exome sequencing data: detecting mobile elements insertion in a rare disease cohort. *European Journal of Human Genetics: EJHG*, 31(7), 761-768. <https://doi.org/10.1038/s41431-022-01250-3>
- Gelinas, R., El Khoury, N., Chaix, M.-A., Beauchamp, C., Alikashani, A., Ethier, N., Boucher, G., Villeneuve, L., Robb, L., Latour, F., Mondesert, B., Rivard, L., Goyette, P., Talajic, M., Fiset, C. et Rioux, J. D. (2017). Characterization of a Human Induced Pluripotent Stem Cell-Derived Cardiomyocyte Model for the Study of Variant Pathogenicity: Validation of a KCNJ2 Mutation. *Circulation. Cardiovascular Genetics*, 10(5), e001755. <https://doi.org/10.1161/CIRCGENETICS.117.001755>
- Genome sequence of the nematode *C. elegans*: a platform for investigating biology. (1998). *Science (New York, N.Y.)*, 282(5396), 2012-2018. <https://doi.org/10.1126/science.282.5396.2012>
- Geoffroy, V., Herenger, Y., Kress, A., Stoetzel, C., Piton, A., Dollfus, H. et Muller, J. (2018). AnnotSV: an integrated tool for structural variations annotation. *Bioinformatics (Oxford, England)*, 34(20), 3572-3574. <https://doi.org/10.1093/bioinformatics/bty304>
- Geoffroy, V., Pizot, C., Redin, C., Piton, A., Vasli, N., Stoetzel, C., Blavier, A., Laporte, J. et Muller, J. (2015). VaRank: a simple and powerful tool for ranking genetic variants. *PeerJ*, 3, e796. <https://doi.org/10.7717/peerj.796>
- Geoffroy, V., Stoetzel, C., Scheidecker, S., Schaefer, E., Perrault, I., Bar, S., Kroll, A., Delbarre, M., Antin, M., Leuvrey, A.-S., Henry, C., Blanche, H., Decker, E., Kloth, K., Klaus, G., Mache, C., Martin-Coignard, D., McGinn, S., Boland, A., ... Muller, J. (2018). Whole-genome sequencing in patients with ciliopathies uncovers a novel recurrent tandem duplication in IFT140. *Human Mutation*, 39(7), 983-992. <https://doi.org/10.1002/humu.23539>
- Ghesh, L., Besnard, T., Nizon, M., Trochu, E., Landeau-Trottier, G., Breheret, F., Thauvin-Robinet, C., Bruel, A., Kuentz, P., Coubes, C., Cuisset, L., Mignot, C., Keren, B., Bézieau, S. et Cogné, B. (2021). Loss-of-function variants in *ARHGEF9* are associated with an X-linked intellectual disability dominant disorder. *Human Mutation*, 42(5), 498-505. <https://doi.org/10.1002/humu.24188>
- Giannandrea, M., Bianchi, V., Mignogna, M. L., Sirri, A., Carrabino, S., D'Elia, E., Vecellio, M., Russo, S., Cogliati, F., Larizza, L., Ropers, H.-H., Tzschach, A., Kalscheuer, V., Oehl-Jaschkowitz, B., Skinner, C., Schwartz, C. E., Gecz, J., Van Esch, H., Raynaud, M., ... D'Adamo, P. (2010). Mutations in the small GTPase gene *RAB39B* are responsible for X-linked mental retardation associated with autism, epilepsy, and macrocephaly. *American Journal of Human Genetics*, 86(2), 185-195. <https://doi.org/10.1016/j.ajhg.2010.01.011>
- Gibbons, R. J., Picketts, D. J., Villard, L. et Higgs, D. R. (1995). Mutations in a putative global transcriptional regulator cause X-linked mental retardation with alpha-thalassemia (ATR-X syndrome). *Cell*, 80(6), 837-845. [https://doi.org/10.1016/0092-8674\(95\)90287-2](https://doi.org/10.1016/0092-8674(95)90287-2)
- Gilissen, C., Hehir-Kwa, J. Y., Thung, D. T., van de Vorst, M., van Bon, B. W. M., Willemsen, M. H., Kwint, M., Janssen, I. M., Hoischen, A., Schenck, A., Leach, R., Klein, R., Tearle, R., Bo, T., Pfundt, R., Yntema, H. G., de Vries, B. B. A., Kleefstra, T., Brunner, H. G., ... Veltman, J. A. (2014). Genome sequencing identifies major causes of severe intellectual disability. *Nature*, 511(7509), 344-347. <https://doi.org/10.1038/nature13394>

- Giorda, R., Bonaglia, M. C., Beri, S., Fichera, M., Novara, F., Magini, P., Urquhart, J., Sharkey, F. H., Zucca, C., Grasso, R., Marelli, S., Castiglia, L., Di Benedetto, D., Musumeci, S. A., Vitello, G. A., Failla, P., Reitano, S., Avola, E., Bisulli, F., ... Zuffardi, O. (2009). Complex segmental duplications mediate a recurrent dup(X)(p11.22-p11.23) associated with mental retardation, speech delay, and EEG anomalies in males and females. *American Journal of Human Genetics*, *85*(3), 394-400. <https://doi.org/10.1016/j.ajhg.2009.08.001>
- Gleeson, J. G., Allen, K. M., Fox, J. W., Lamperti, E. D., Berkovic, S., Scheffer, I., Cooper, E. C., Dobyns, W. B., Minnerath, S. R., Ross, M. E. et Walsh, C. A. (1998). Doublecortin, a brain-specific gene mutated in human X-linked lissencephaly and double cortex syndrome, encodes a putative signaling protein. *Cell*, *92*(1), 63-72. [https://doi.org/10.1016/s0092-8674\(00\)80899-5](https://doi.org/10.1016/s0092-8674(00)80899-5)
- Graham, J. M. J. et Schwartz, C. E. (2013). MED12 related disorders. *American Journal of Medical Genetics. Part A*, *161A*(11), 2734-2740. <https://doi.org/10.1002/ajmg.a.36183>
- Grams, S. E., Argiropoulos, B., Lines, M., Chakraborty, P., McGowan-Jordan, J., Geraghty, M. T., Tsang, M., Eswara, M., Tezcan, K., Adams, K. L., Linck, L., Himes, P., Kostiner, D., Zand, D. J., Stalker, H., Driscoll, D. J., Huang, T., Rosenfeld, J. A., Li, X. et Chen, E. (2016). Genotype-phenotype characterization in 13 individuals with chromosome Xp11.22 duplications. *American Journal of Medical Genetics. Part A*, *170A*(4), 967-977. <https://doi.org/10.1002/ajmg.a.37519>
- Grantham, R. (1974). Amino acid difference formula to help explain protein evolution. *Science (New York, N.Y.)*, *185*(4154), 862-864. <https://doi.org/10.1126/science.185.4154.862>
- Grasshoff, U., Bonin, M., Goehring, I., Ekici, A., Dufke, A., Cremer, K., Wagner, N., Rossier, E., Jauch, A., Walter, M., Bauer, C., Bauer, P., Horber, K., Beck-Woedl, S. et Wiczorek, D. (2011). De novo MECP2 duplication in two females with random X-inactivation and moderate mental retardation. *European Journal of Human Genetics : EJHG*, *19*(5), 507-512. <https://doi.org/10.1038/ejhg.2010.226>
- Guénet, J. L. (2005). The mouse genome. *Genome Research*, *15*(12), 1729-1740. <https://doi.org/10.1101/gr.3728305>
- Gulko, B., Hubisz, M. J., Gronau, I. et Siepel, A. (2015). Probabilities of Fitness Consequences for Point Mutations Across the Human Genome. *Nature genetics*, *47*(3), 276-283. <https://doi.org/10.1038/ng.3196>
- Guo, Y., Long, J., He, J., Li, C.-I., Cai, Q., Shu, X.-O., Zheng, W. et Li, C. (2012). Exome sequencing generates high quality data in non-target regions. *BMC Genomics*, *13*, 194. <https://doi.org/10.1186/1471-2164-13-194>
- Gymrek, M. (2017). A genomic view of short tandem repeats. *Current Opinion in Genetics & Development*, *44*, 9-16. <https://doi.org/10.1016/j.gde.2017.01.012>
- HALDANE, J. (1938). A Clinical and Genetic Study of 1280 Cases of Mental Defect, *141*, 575.
- Hall, A. et Nobes, C. D. (2000). Rho GTPases: molecular switches that control the organization and dynamics of the actin cytoskeleton. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *355*(1399), 965-970. <https://doi.org/10.1098/rstb.2000.0632>
- Han, K. A., Kim, J., Kim, H., Kim, D., Lim, D., Ko, J. et Um, J. W. (2019). Slitrk2 controls excitatory synapse development via PDZ-mediated protein interactions. *Scientific Reports*, *9*(1), 17094. <https://doi.org/10.1038/s41598-019-53519-1>
- Hannan, A. J. (2018). Tandem repeats mediating genetic plasticity in health and disease. *Nature Reviews. Genetics*, *19*(5), 286-298. <https://doi.org/10.1038/nrg.2017.115>
- Haraksingh, R. R. et Snyder, M. P. (2013). Impacts of variation in the human genome on gene regulation. *Journal of Molecular Biology*, *425*(21), 3970-3977. <https://doi.org/10.1016/j.jmb.2013.07.015>
- Harris, J. C. (2010). Advances in understanding behavioral phenotypes in neurogenetic syndromes. *American Journal of Medical Genetics. Part C, Seminars in Medical Genetics*, *154C*(4), 389-399. <https://doi.org/10.1002/ajmg.c.30276>
- Harvey, K., Duguid, I. C., Alldred, M. J., Beatty, S. E., Ward, H., Keep, N. H., Lingenfelter, S. E., Pearce, B. R., Lundgren, J., Owen, M. J., Smart, T. G., Lüscher, B., Rees, M. I. et Harvey, R. J. (2004). The GDP-GTP Exchange Factor Collybistin: An Essential Determinant of Neuronal Gephyrin Clustering. *The Journal of Neuroscience*, *24*(25), 5816-5826. <https://doi.org/10.1523/JNEUROSCI.1184-04.2004>
- Hofer, M. et Lutolf, M. P. (2021). Engineering organoids. *Nature Reviews. Materials*, *6*(5), 402-420. <https://doi.org/10.1038/s41578-021-00279-y>
- Houdayer, C., Caux-Moncoutier, V., Krieger, S., Barrois, M., Bonnet, F., Bourdon, V., Bronner, M., Buisson, M., Coulet, F., Gaildrat, P., Lefol, C., Leone, M., Mazoyer, S., Muller, D., Remenieras, A., Revillion, F., Rouleau, E., Sokolowska, J., Vert, J.-P., ... Stoppa-Lyonnet, D. (2012). Guidelines for splicing analysis in molecular diagnosis derived from a set of 327 combined in silico/in vitro studies on BRCA1 and BRCA2 variants. *Human Mutation*, *33*(8), 1228-1238. <https://doi.org/10.1002/humu.22101>

- Hsi-Yang Fritz, M., Leinonen, R., Cochrane, G. et Birney, E. (2011). Efficient storage of high throughput DNA sequencing data using reference-based compression. *Genome Research*, 21(5), 734-740. <https://doi.org/10.1101/gr.114819.110>
- Hu, H., Haas, S. A., Chelly, J., Van Esch, H., Raynaud, M., de Brouwer, A. P. M., Weinert, S., Froyen, G., Frints, S. G. M., Laumonnier, F., Zemojtel, T., Love, M. I., Richard, H., Emde, A.-K., Bienek, M., Jensen, C., Hambrock, M., Fischer, U., Langnick, C., ... Kalscheuer, V. M. (2016). X-exome sequencing of 405 unresolved families identifies seven novel intellectual disability genes. *Molecular Psychiatry*, 21(1), 133-148. <https://doi.org/10.1038/mp.2014.193>
- Huang, J., Zhu, T., Qu, Y. et Mu, D. (2016). Prenatal, Perinatal and Neonatal Risk Factors for Intellectual Disability: A Systemic Review and Meta-Analysis. *PLoS ONE*, 11(4), e0153655. <https://doi.org/10.1371/journal.pone.0153655>
- Ibarluzea, N., Hoz, A. B. de la, Villate, O., Llano, I., Ocio, I., Marti, I., Guitart, M., Gabau, E., Andrade, F., Gener, B. et Tejada, M.-I. (2020). Targeted Next-Generation Sequencing in Patients with Suggestive X-Linked Intellectual Disability. *Genes*, 11(1), E51. <https://doi.org/10.3390/genes11010051>
- Initial sequencing and comparative analysis of the mouse genome | Nature*. (s. d.). <https://www.nature.com/articles/nature01262>
- Ioannidis, N. M., Davis, J. R., DeGorter, M. K., Larson, N. B., McDonnell, S. K., French, A. J., Battle, A. J., Hastie, T. J., Thibodeau, S. N., Montgomery, S. B., Bustamante, C. D., Sieh, W. et Whittemore, A. S. (2017). FIRE: functional inference of genetic variants that regulate gene expression. *Bioinformatics (Oxford, England)*, 33(24), 3895-3901. <https://doi.org/10.1093/bioinformatics/btx534>
- Iqbal, M. A., Broeckel, U., Levy, B., Skinner, S., Sahajpal, N. S., Rodriguez, V., Stence, A., Awayda, K., Scharer, G., Skinner, C., Stevenson, R., Bossler, A., Nagy, P. L. et Kolhe, R. (2023). Multisite Assessment of Optical Genome Mapping for Analysis of Structural Variants in Constitutional Postnatal Cases. *The Journal of Molecular Diagnostics : JMD*, 25(3), 175-188. <https://doi.org/10.1016/j.jmoldx.2022.12.005>
- Ishii, N., Ozaki, K., Sato, H., Mizuno, H., Saito, S., Takahashi, A., Miyamoto, Y., Ikegawa, S., Kamatani, N., Hori, M., Saito, S., Nakamura, Y. et Tanaka, T. (2006). Identification of a novel non-coding RNA, MIAT, that confers risk of myocardial infarction. *Journal of Human Genetics*, 51(12), 1087-1099. <https://doi.org/10.1007/s10038-006-0070-9>
- Izumi, K., Brett, M., Nishi, E., Drunat, S., Tan, E.-S., Fujiki, K., Lebon, S., Cham, B., Masuda, K., Arakawa, M., Jacquinet, A., Yamazumi, Y., Chen, S.-T., Verloes, A., Okada, Y., Katou, Y., Nakamura, T., Akiyama, T., Gressens, P., ... Shirahige, K. (2016). ARCN1 Mutations Cause a Recognizable Craniofacial Syndrome Due to COPI- Mediated Transport Defects. *American Journal of Human Genetics*, 99(2), 451-459. <https://doi.org/10.1016/j.ajhg.2016.06.011>
- Jaganathan, K., Kyriazopoulou Panagiotopoulou, S., McRae, J. F., Darbandi, S. F., Knowles, D., Li, Y. I., Kosmicki, J. A., Arbelaez, J., Cui, W., Schwartz, G. B., Chow, E. D., Kanterakis, E., Gao, H., Kia, A., Batzoglu, S., Sanders, S. J. et Farh, K. K.-H. (2019). Predicting Splicing from Primary Sequence with Deep Learning. *Cell*, 176(3), 535-548.e24. <https://doi.org/10.1016/j.cell.2018.12.015>
- Jegathisawaran, J., Tsiplova, K., Hayeems, R. et Ungar, W. J. (2020). Determining accurate costs for genomic sequencing technologies-a necessary prerequisite. *Journal of Community Genetics*, 11(2), 235-238. <https://doi.org/10.1007/s12687-019-00442-7>
- Jensen, L. R., Amende, M., Gurok, U., Moser, B., Gimmel, V., Tzschach, A., Janecke, A. R., Tariverdian, G., Chelly, J., Fryns, J.-P., Van Esch, H., Kleefstra, T., Hamel, B., Moraine, C., Gecz, J., Turner, G., Reinhardt, R., Kalscheuer, V. M., Ropers, H.-H. et Lenzner, S. (2005). Mutations in the JARID1C gene, which is involved in transcriptional regulation and chromatin remodeling, cause X-linked mental retardation. *American Journal of Human Genetics*, 76(2), 227-236. <https://doi.org/10.1086/427563>
- Jensen, M. et Girirajan, S. (2017). Mapping a shared genetic basis for neurodevelopmental disorders. *Genome Medicine*, 9(1), 109. <https://doi.org/10.1186/s13073-017-0503-4>
- Ji, J., Leung, M. L., Baker, S., Deignan, J. L. et Santani, A. (2021). Clinical Exome Reanalysis: Current Practice and Beyond. *Molecular Diagnosis & Therapy*, 25(5), 529-536. <https://doi.org/10.1007/s40291-021-00541-7>
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. et Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (New York, N. Y.)*, 337(6096), 816-821. <https://doi.org/10.1126/science.1225829>
- Jinnah, H. A. (1993). HPRT1 Disorders. Dans M. P. Adam, J. Feldman, G. M. Mirzaa, R. A. Pagon, S. E. Wallace, L. J. Bean, K. W. Gripp et A. Amemiya (dir.), *GeneReviews*®. University of Washington, Seattle. <http://www.ncbi.nlm.nih.gov/books/NBK1149/>



- Johnson, D. S., Mortazavi, A., Myers, R. M. et Wold, B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. *Science (New York, N.Y.)*, 316(5830), 1497-1502. <https://doi.org/10.1126/science.1141319>
- Jolly, L. A., Parnell, E., Gardner, A. E., Corbett, M. A., Pérez-Jurado, L. A., Shaw, M., Lesca, G., Keegan, C., Schneider, M. C., Griffin, E., Maier, F., Kiss, C., Guerin, A., Crosby, K., Rosenbaum, K., Tanpaiboon, P., Whalen, S., Keren, B., McCarrier, J., ... Gecz, J. (2020). Missense variant contribution to USP9X-female syndrome. *NPJ Genomic Medicine*, 5(1), 53. <https://doi.org/10.1038/s41525-020-00162-9>
- Julien, P., Minana, B., Baeza-Centurion, P., Valcarcel, J. et Lehner, B. (2016). The complete local genotype-phenotype landscape for the alternative splicing of a human exon. *Nature Communications*, 7, 11558. <https://doi.org/10.1038/ncomms11558>
- Jung, H. H., Danek, A., Walker, R. H., Frey, B. M. et Peikert, K. (1993). McLeod Neuroacanthocytosis Syndrome. Dans M. P. Adam, G. M. Mirzaa, R. A. Pagon, S. E. Wallace, L. J. Bean, K. W. Gripp et A. Amemiya (dir.), *GeneReviews((R))*. University of Washington, Seattle.
- Kalscheuer, V. M., Freude, K., Musante, L., Jensen, L. R., Yntema, H. G., Gecz, J., Sefiani, A., Hoffmann, K., Moser, B., Haas, S., Gurok, U., Haesler, S., Aranda, B., Nshedjan, A., Tzschach, A., Hartmann, N., Roloff, T.-C., Shoichet, S., Hagens, O., ... Ropers, H.-H. (2003). Mutations in the polyglutamine binding protein 1 gene cause X-linked mental retardation. *Nature Genetics*, 35(4), 313-315. <https://doi.org/10.1038/ng1264>
- Kang, H., Han, K. A., Won, S. Y., Kim, H. M., Lee, Y.-H., Ko, J. et Um, J. W. (2016). Slitrk Missense Mutations Associated with Neuropsychiatric Disorders Distinctively Impair Slitrk Trafficking and Synapse Formation. *Frontiers in Molecular Neuroscience*, 9, 104. <https://doi.org/10.3389/fnmol.2016.00104>
- Kaplanis, J., Samocho, K. E., Wiel, L., Zhang, Z., Arvai, K. J., Eberhardt, R. Y., Gallone, G., Lelieveld, S. H., Martin, H. C., McRae, J. F., Short, P. J., Torene, R. I., de Boer, E., Danecek, P., Gardner, E. J., Huang, N., Lord, J., Martincorena, I., Pfundt, R., ... Retterer, K. (2020). Evidence for 28 genetic disorders discovered by combining healthcare and research data. *Nature*, 586(7831), 757-762. <https://doi.org/10.1038/s41586-020-2832-5>
- Karam, A., Delvallee, C., Estrada-Cuzcano, A., Geoffroy, V., Lamouche, J.-B., Leuvrey, A.-S., Nourisson, E., Tarabeux, J., Stoetzel, C., Scheidecker, S., Porter, L. F., Genin, E., Redon, R., Sandron, F., Boland, A., Deleuze, J.-F., Le May, N., Dollfus, H. et Muller, J. (2023). WGS Revealed Novel BBS5 Pathogenic Variants, Missed by WES, Causing Ciliary Structure and Function Defects. *International Journal of Molecular Sciences*, 24(10), 8729. <https://doi.org/10.3390/ijms24108729>
- Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., Collins, R. L., Laricchia, K. M., Ganna, A., Birnbaum, D. P., Gauthier, L. D., Brand, H., Solomonson, M., Watts, N. A., Rhodes, D., Singer-Berk, M., Seaby, E. G., Kosmicki, J. A., Walters, R. K., ... MacArthur, D. G. (2019). Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *bioRxiv*. <https://doi.org/10.1101/531210>
- Kato, M., Das, S., Petras, K., Kitamura, K., Morohashi, K.-I., Abuelo, D. N., Barr, M., Bonneau, D., Brady, A. F., Carpenter, N. J., Ciperio, K. L., Frisone, F., Fukuda, T., Guerrini, R., Iida, E., Itoh, M., Lewanda, A. F., Nanba, Y., Oka, A., ... Dobyns, W. B. (2004). Mutations of ARX are associated with striking pleiotropy and consistent genotype-phenotype correlation. *Human Mutation*, 23(2), 147-159. <https://doi.org/10.1002/humu.10310>
- Kaur, S. et Christodoulou, J. (1993). MECP2 Disorders. Dans M. P. Adam, G. M. Mirzaa, R. A. Pagon, S. E. Wallace, L. J. Bean, K. W. Gripp et A. Amemiya (dir.), *GeneReviews((R))*. University of Washington, Seattle.
- Kazazian, H. H. J., Wong, C., Youssoufian, H., Scott, A. F., Phillips, D. G. et Antonarakis, S. E. (1988). Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature*, 332(6160), 164-166. <https://doi.org/10.1038/332164a0>
- Khan, M., Cornelis, S. S., Sangermano, R., Post, I. J. M., Groesbeek, A. J., Amsu, J., Gilissen, C., Garanto, A., Collin, R. W. J. et Cremers, F. P. M. (2020). In or Out? New Insights on Exon Recognition through Splice-Site Interdependency. *International Journal of Molecular Sciences*, 21(7), 2300. <https://doi.org/10.3390/ijms21072300>
- King, B., Toth, K., Hodapp, R. M. et Dykens, E. (2009). Intellectual disability. *Comprehensive textbook of psychiatry*, 3444-3474.
- Kircher, M., Witten, D. M., Jain, P., O’Roak, B. J., Cooper, G. M. et Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics*, 46(3), 310-315. <https://doi.org/10.1038/ng.2892>
- Kitamura, K., Yanazawa, M., Sugiyama, N., Miura, H., Iizuka-Kogo, A., Kusaka, M., Omichi, K., Suzuki, R., Kato-Fukui, Y., Kamiirisa, K., Matsuo, M., Kamijo, S., Kasahara, M., Yoshioka, H., Ogata, T., Fukuda, T., Kondo, I., Kato, M., Dobyns, W. B., ... Morohashi, K. (2002). Mutation of ARX causes abnormal development of forebrain

- and testes in mice and X-linked lissencephaly with abnormal genitalia in humans. *Nature Genetics*, 32(3), 359-369. <https://doi.org/10.1038/ng1009>
- Kleefstra, T., van Zelst-Stams, W. A., Nillesen, W. M., Cormier-Daire, V., Houge, G., Foulds, N., van Dooren, M., Willemsen, M. H., Pfundt, R., Turner, A., Wilson, M., McGaughan, J., Rauch, A., Zenker, M., Adam, M. P., Innes, M., Davies, C., Lopez, A. G.-M., Casalone, R., ... Brunner, H. G. (2009). Further clinical and molecular delineation of the 9q subtelomeric deletion syndrome supports a major contribution of EHMT1 haploinsufficiency to the core phenotype. *Journal of Medical Genetics*, 46(9), 598-606. <https://doi.org/10.1136/jmg.2008.062950>
- Knebelmann, B., Breillat, C., Forestier, L., Arrondel, C., Jacassier, D., Giatras, I., Drouot, L., Deschênes, G., Grünfeld, J. P., Broyer, M., Gubler, M. C. et Antignac, C. (1996). Spectrum of mutations in the COL4A5 collagen gene in X-linked Alport syndrome. *American Journal of Human Genetics*, 59(6), 1221-1232.
- Knopp, C., Rudnik-Schoneborn, S., Eggermann, T., Bergmann, C., Begemann, M., Schoner, K., Zerres, K. et Ortiz Bruchle, N. (2015). Syndromic ciliopathies: From single gene to multi gene analysis by SNP arrays and next generation sequencing. *Molecular and Cellular Probes*, 29(5), 299-307. <https://doi.org/10.1016/j.mcp.2015.05.008>
- Knudsen, G. P. S., Pedersen, J., Klingenberg, O., Lygren, I. et Orstavik, K. H. (2007). Increased skewing of X chromosome inactivation with age in both blood and buccal cells. *Cytogenetic and Genome Research*, 116(1-2), 24-28. <https://doi.org/10.1159/000097414>
- Ko, J. (2012). The leucine-rich repeat superfamily of synaptic adhesion molecules: LRRTMs and Slitrks. *Molecules and Cells*, 34(4), 335-340. <https://doi.org/10.1007/s10059-012-0113-3>
- Koboldt, D. C., Chen, K., Wylie, T., Larson, D. E., McLellan, M. D., Mardis, E. R., Weinstock, G. M., Wilson, R. K. et Ding, L. (2009). VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics (Oxford, England)*, 25(17), 2283-2285. <https://doi.org/10.1093/bioinformatics/btp373>
- Kochinke, K., Zweier, C., Nijhof, B., Fenckova, M., Cizek, P., Honti, F., Keerthikumar, S., Oortveld, M. A. W., Kleefstra, T., Kramer, J. M., Webber, C., Huynen, M. A. et Schenck, A. (2016). Systematic Phenomics Analysis Deconvolutes Genes Mutated in Intellectual Disability into Biologically Coherent Modules. *American Journal of Human Genetics*, 98(1), 149-164. <https://doi.org/10.1016/j.ajhg.2015.11.024>
- Koide, R., Kobayashi, S., Shimohata, T., Ikeuchi, T., Maruyama, M., Saito, M., Yamada, M., Takahashi, H. et Tsuji, S. (1999). A neurological disease caused by an expanded CAG trinucleotide repeat in the TATA-binding protein gene: a new polyglutamine disease? *Human Molecular Genetics*, 8(11), 2047-2053. <https://doi.org/10.1093/hmg/8.11.2047>
- Kolehmainen, J., Black, G. C. M., Saarinen, A., Chandler, K., Clayton-Smith, J., Traskelin, A.-L., Perveen, R., Kivitie-Kallio, S., Norio, R., Warburg, M., Fryns, J.-P., de la Chapelle, A. et Lehesjoki, A.-E. (2003). Cohen syndrome is caused by mutations in a novel gene, COH1, encoding a transmembrane protein with a presumed role in vesicle-mediated sorting and intracellular protein transport. *American Journal of Human Genetics*, 72(6), 1359-1369. <https://doi.org/10.1086/375454>
- Korbel, J. O., Urban, A. E., Affourtit, J. P., Godwin, B., Grubert, F., Simons, J. F., Kim, P. M., Palejev, D., Carriero, N. J., Du, L., Taillon, B. E., Chen, Z., Tanzer, A., Saunders, A. C. E., Chi, J., Yang, F., Carter, N. P., Hurles, M. E., Weissman, S. M., ... Snyder, M. (2007). Paired-end mapping reveals extensive structural variation in the human genome. *Science (New York, N.Y.)*, 318(5849), 420-426. <https://doi.org/10.1126/science.1149504>
- Krumm, N. et Hoffman, N. (2020). Practical estimation of cloud storage costs for clinical genomic data. *Practical Laboratory Medicine*, 21, e00168. <https://doi.org/10.1016/j.plabm.2020.e00168>
- Kumar, P., Henikoff, S. et Ng, P. C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature Protocols*, 4(7), 1073-1081. <https://doi.org/10.1038/nprot.2009.86>
- Kumar, R., Ha, T., Pham, D., Shaw, M., Mangelsdorf, M., Friend, K. L., Hobson, L., Turner, G., Boyle, J., Field, M., Hackett, A., Corbett, M. et Gecz, J. (2016). A non-coding variant in the 5' UTR of DLG3 attenuates protein translation to cause non-syndromic intellectual disability. *European Journal of Human Genetics : EJHG*, 24(11), 1612-1616. <https://doi.org/10.1038/ejhg.2016.46>
- Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M. J., Amin, V., Whitaker, J. W., Schultz, M. D., Ward, L. D., Sarkar, A., Quon, G., Sandstrom, R. S., Eaton, M. L., Wu, Y.-C., ... Kellis, M. (2015). Integrative analysis of 111 reference human epigenomes. *Nature*, 518(7539), 317-330. <https://doi.org/10.1038/nature14248>

- La Spada, A. R., Wilson, E. M., Lubahn, D. B., Harding, A. E. et Fischbeck, K. H. (1991). Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature*, 352(6330), 77-79. <https://doi.org/10.1038/352077a0>
- Lam, E. T., Hastie, A., Lin, C., Ehrlich, D., Das, S. K., Austin, M. D., Deshpande, P., Cao, H., Nagarajan, N., Xiao, M. et Kwok, P.-Y. (2012). Genome mapping on nanochannel arrays for structural variation analysis and sequence assembly. *Nature Biotechnology*, 30(8), 771-776. <https://doi.org/10.1038/nbt.2303>
- Lamoril, J., Ameziane, N., Deybach, J.-C., Bouizegarène, P. et Bogard, M. (2008). Les techniques de séquençage de l'ADN : une révolution en marche. Première partie. *Immuno-analyse & Biologie Spécialisée*, 23(5), 260-279. <https://doi.org/10.1016/j.immbio.2008.07.016>
- Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczy, J., LeVine, R., McEwan, P., ... Szustakowki, J. (2001). Initial sequencing and analysis of the human genome. *Nature*, 409(6822), 860-921. <https://doi.org/10.1038/35057062>
- Landrum, M. J., Chitipiralla, S., Brown, G. R., Chen, C., Gu, B., Hart, J., Hoffman, D., Jang, W., Kaur, K., Liu, C., Lyoshin, V., Maddipatla, Z., Maiti, R., Mitchell, J., O'Leary, N., Riley, G. R., Shi, W., Zhou, G., Schneider, V., ... Kattman, B. L. (2020). ClinVar: improvements to accessing data. *Nucleic Acids Research*, 48(D1), D835-D844. <https://doi.org/10.1093/nar/gkz972>
- Landrum, M. J., Lee, J. M., Benson, M., Brown, G. R., Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., Jang, W., Karapetyan, K., Katz, K., Liu, C., Maddipatla, Z., Malheiro, A., McDaniel, K., Ovetsky, M., Riley, G., Zhou, G., ... Maglott, D. R. (2018). ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Research*, 46(D1), D1062-D1067. <https://doi.org/10.1093/nar/gkx1153>
- Laumonnier, F., Holbert, S., Ronce, N., Faravelli, F., Lenzner, S., Schwartz, C. E., Lespinasse, J., Van Esch, H., Lacombe, D., Goizet, C., Phan-Dinh Tuy, F., van Bokhoven, H., Fryns, J.-P., Chelly, J., Ropers, H.-H., Moraine, C., Hamel, B. C. J. et Briault, S. (2005). Mutations in PHF8 are associated with X linked mental retardation and cleft lip/cleft palate. *Journal of Medical Genetics*, 42(10), 780-786. <https://doi.org/10.1136/jmg.2004.029439>
- Laumonnier, Frederic, Bonnet-Brilhault, F., Gomot, M., Blanc, R., David, A., Moizard, M.-P., Raynaud, M., Ronce, N., Lemonnier, E., Calvas, P., Laudier, B., Chelly, J., Fryns, J.-P., Ropers, H.-H., Hamel, B. C. J., Andres, C., Barthelemy, C., Moraine, C. et Briault, S. (2004). X-linked mental retardation and autism are associated with a mutation in the NLGN4 gene, a member of the neuroligin family. *American Journal of Human Genetics*, 74(3), 552-557. <https://doi.org/10.1086/382137>
- Layer, R. M., Chiang, C., Quinlan, A. R. et Hall, I. M. (2014). LUMPY: a probabilistic framework for structural variant discovery. *Genome Biology*, 15(6), R84. <https://doi.org/10.1186/gb-2014-15-6-r84>
- Leitao, E., Schroder, C., Parenti, I., Dalle, C., Rastetter, A., Kuhnel, T., Kuechler, A., Kaya, S., Gerard, B., Schaefer, E., Nava, C., Drouot, N., Engel, C., Piard, J., Duban-Bedu, B., Villard, L., Stegmann, A. P. A., Vanhoutte, E. K., Verdonschot, J. A. J., ... Depienne, C. (2022). Systematic analysis and prediction of genes associated with monogenic disorders on human chromosome X. *Nature Communications*, 13(1), 6570. <https://doi.org/10.1038/s41467-022-34264-y>
- Lejeune, C., Robert-Viard, C., Meunier-Beillard, N., Borel, M. A., Gourvès, L., Staraci, S., Soilly, A.-L., Guillemin, F., Seror, V., Achit, H., Bouctot, M., Asensio, M.-L., Briffaut, A.-S., Delmas, C., Bruel, A.-L., Benoit, A., Simon, A., Gerard, B., Hadj Abdallah, H., ... Dollfus, H. (2022). The Economic, Medical and Psychosocial Consequences of Whole Genome Sequencing for the Genetic Diagnosis of Patients With Intellectual Disability: The DEFIDIAG Study Protocol. *Frontiers in Genetics*, 13, 852472. <https://doi.org/10.3389/fgene.2022.852472>
- Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., O'Donnell-Luria, A. H., Ware, J. S., Hill, A. J., Cummings, B. B., Tukiainen, T., Birnbaum, D. P., Kosmicki, J. A., Duncan, L. E., Estrada, K., Zhao, F., Zou, J., Pierce-Hoffman, E., Berghout, J., ... MacArthur, D. G. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, 536(7616), 285-291. <https://doi.org/10.1038/nature19057>
- Lelieveld, S. H., Spielmann, M., Mundlos, S., Veltman, J. A. et Gilissen, C. (2015). Comparison of Exome and Genome Sequencing Technologies for the Complete Capture of Protein-Coding Regions. *Human Mutation*, 36(8), 815-822. <https://doi.org/10.1002/humu.22813>
- Lelieveld, S. H., Veltman, J. A. et Gilissen, C. (2016). Novel bioinformatic developments for exome sequencing. *Human Genetics*, 135(6), 603-614. <https://doi.org/10.1007/s00439-016-1658-6>
- Leman, R., Parfait, B., Vidaud, D., Girodon, E., Pacot, L., Le Gac, G., Ka, C., Ferec, C., Fichou, Y., Quesnelle, C., Aucouturier, C., Muller, E., Vaur, D., Castera, L., Boulouard, F., Ricou, A., Tubeuf, H., Soukariéh, O., Gaildrat, P., ... Krieger, S. (2022). SPiP: Splicing Prediction Pipeline, a machine learning tool for massive

- detection of exonic and intronic variant effects on mRNA splicing. *Human Mutation*, 43(12), 2308-2323. <https://doi.org/10.1002/humu.24491>
- LESCH, M. et NYHAN, W. L. (1964). A FAMILIAL DISORDER OF URIC ACID METABOLISM AND CENTRAL NERVOUS SYSTEM FUNCTION. *The American Journal of Medicine*, 36, 561-570. [https://doi.org/10.1016/0002-9343\(64\)90104-4](https://doi.org/10.1016/0002-9343(64)90104-4)
- Lesieur-Sebellin, M., Till, M., Khau Van Kien, P., Herve, B., Bourgon, N., Dupont, C., Tabet, A.-C., Barrois, M., Coussement, A., Loeuillet, L., Mousty, E., Ea, V., El Assal, A., Mary, L., Jaillard, S., Beneteau, C., Le Vaillant, C., Coutton, C., Devillard, F., ... Malan, V. (2022). Terminal 6q deletions cause brain malformations, a phenotype mimicking heterozygous DLL1 pathogenic variants: A multicenter retrospective case series. *Prenatal Diagnosis*, 42(1), 118-135. <https://doi.org/10.1002/pd.6074>
- Levy, J., Cogan, G., Maruani, A., Maillard, A., Dupont, C., Drunat, S., Rachid, M., Atzori, P., Delorme, R., Jeyarajah, S., Isidor, B., Pichon, O., Moradkhani, K., Verloes, A. et Tabet, A.-C. (2022). Rare and de novo duplications containing TCF20 are associated with a neurodevelopmental disorder. *Clinical Genetics*, 101(3), 364-370. <https://doi.org/10.1111/cge.14099>
- Li, D., Strong, A., Shen, K. M., Cassiman, D., Van Dyck, M., Linhares, N. D., Valadares, E. R., Wang, T., Pena, S. D. J., Jaeken, J., Vergano, S., Zackai, E., Hing, A., Chow, P., Ganguly, A., Scholz, T., Bierhals, T., Philipp, D., Hakonarson, H. et Bhoj, E. (2021). De novo loss-of-function variants in X-linked MED12 are associated with Hardikar syndrome in females. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, 23(4), 637-644. <https://doi.org/10.1038/s41436-020-01031-7>
- Li, H. et Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)*, 25(14), 1754-1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Lim, K. H., Ferraris, L., Filloux, M. E., Raphael, B. J. et Fairbrother, W. G. (2011). Using positional distribution to identify splicing elements and predict pre-mRNA processing defects in human genes. *Proceedings of the National Academy of Sciences of the United States of America*, 108(27), 11093-11098. <https://doi.org/10.1073/pnas.1101135108>
- Liu, Xiaoming, Li, C. et Boerwinkle, E. (2017). The performance of deleteriousness prediction scores for rare non-protein-changing single nucleotide variants in human genes. *Journal of Medical Genetics*, 54(2), 134-144. <https://doi.org/10.1136/jmedgenet-2016-104369>
- Liu, Xuanyu, Ma, Y., Yin, K., Li, W., Chen, W., Zhang, Y., Zhu, C., Li, T., Han, B., Liu, X., Wang, S. et Zhou, Z. (2019). Long non-coding and coding RNA profiling using strand-specific RNA-seq in human hypertrophic cardiomyopathy. *Scientific Data*, 6(1), 90. <https://doi.org/10.1038/s41597-019-0094-6>
- Liu, Y., McDermott, S., Lawson, A. et Aelion, C. M. (2010). The relationship between mental retardation and developmental delays in children and the levels of arsenic, mercury and lead in soil samples taken near their mother's residence during pregnancy. *International Journal of Hygiene and Environmental Health*, 213(2), 116-123. <https://doi.org/10.1016/j.ijheh.2009.12.004>
- Long, P., May, M. M., James, V. M., Grannò, S., Johnson, J. P., Tarpey, P., Stevenson, R. E., Harvey, K., Schwartz, C. E. et Harvey, R. J. (2016). Missense Mutation R338W in ARHGEF9 in a Family with X-linked Intellectual Disability with Variable Macrocephaly and Macro-Orchidism. *Frontiers in Molecular Neuroscience*, 8. <https://doi.org/10.3389/fnmol.2015.00083>
- Lopez-Bigas, N., Audit, B., Ouzounis, C., Parra, G. et Guigo, R. (2005). Are splicing mutations the most frequent cause of hereditary disease? *FEBS Letters*, 579(9), 1900-1903. <https://doi.org/10.1016/j.febslet.2005.02.047>
- Lord, J. et Baralle, D. (2021). Splicing in the Diagnosis of Rare Disease: Advances and Challenges. *Frontiers in Genetics*, 12, 689892. <https://doi.org/10.3389/fgene.2021.689892>
- Lower, K. M., Turner, G., Kerr, B. A., Mathews, K. D., Shaw, M. A., Gedeon, A. K., Schelley, S., Hoyme, H. E., White, S. M., Delatycki, M. B., Lampe, A. K., Clayton-Smith, J., Stewart, H., van Ravenswaay, C. M. A., de Vries, B. B. A., Cox, B., Grompe, M., Ross, S., Thomas, P., ... Gecz, J. (2002). Mutations in PHF6 are associated with Borjeson-Forssman-Lehmann syndrome. *Nature Genetics*, 32(4), 661-665. <https://doi.org/10.1038/ng1040>
- Lubs, H. A. (1969). A marker X chromosome. *American Journal of Human Genetics*, 21(3), 231-244.
- Lubs, H., Abidi, F., Bier, J. A., Abuelo, D., Ouzts, L., Voeller, K., Fennell, E., Stevenson, R. E., Schwartz, C. E. et Arena, F. (1999). XLMR syndrome characterized by multiple respiratory infections, hypertelorism, severe CNS deterioration and early death localizes to distal Xq28. *American Journal of Medical Genetics*, 85(3), 243-248. [https://doi.org/10.1002/\(sici\)1096-8628\(19990730\)85:3<243::aid-ajmg11>3.0.co;2-e](https://doi.org/10.1002/(sici)1096-8628(19990730)85:3<243::aid-ajmg11>3.0.co;2-e)

- Lubs, Herbert A., Stevenson, R. E. et Schwartz, C. E. (2012). Fragile X and X-linked intellectual disability: four decades of discovery. *American Journal of Human Genetics*, 90(4), 579-590. <https://doi.org/10.1016/j.ajhg.2012.02.018>
- Lupianez, D. G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J. M., Laxova, R., Santos-Simarro, F., Gilbert-Dussardier, B., Wittler, L., Borschiwer, M., Haas, S. A., Osterwalder, M., Franke, M., Timmermann, B., Hecht, J., ... Mundlos, S. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell*, 161(5), 1012-1025. <https://doi.org/10.1016/j.cell.2015.04.004>
- Lupski, J. R., Reid, J. G., Gonzaga-Jauregui, C., Rio Deiros, D., Chen, D. C. Y., Nazareth, L., Bainbridge, M., Dinh, H., Jing, C., Wheeler, D. A., McGuire, A. L., Zhang, F., Stankiewicz, P., Halperin, J. J., Yang, C., Gehman, C., Guo, D., Irikat, R. K., Tom, W., ... Gibbs, R. A. (2010). Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *The New England Journal of Medicine*, 362(13), 1181-1191. <https://doi.org/10.1056/NEJMoa0908094>
- LYON, M. F. (1961). Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature*, 190, 372-373. <https://doi.org/10.1038/190372a0>
- Machado, C. O. F., Griesi-Oliveira, K., Rosenberg, C., Kok, F., Martins, S., Rita Passos-Bueno, M. et Sertie, A. L. (2016). Collybistin binds and inhibits mTORC1 signaling: a potential novel mechanism contributing to intellectual disability and autism. *European Journal of Human Genetics*, 24(1), 59-65. <https://doi.org/10.1038/ejhg.2015.69>
- Mahmoud, M., Gobet, N., Cruz-Davalos, D. I., Mounier, N., Dessimoz, C. et Sedlazeck, F. J. (2019). Structural variant calling: the long and the short of it. *Genome Biology*, 20(1), 246. <https://doi.org/10.1186/s13059-019-1828-7>
- Maia, N., Ibarluzea, N., Misra-Isrie, M., Koboldt, D. C., Marques, I., Soares, G., Santos, R., Marcelis, C. L. M., Keski-Filppula, R., Guitart, M., Gabau Vila, E., Lehman, A., Hickey, S., Mori, M., Terhal, P., Valenzuela, I., Lasaranzasti, A., Cueto-González, A. M., Chhouk, B. H., ... de Brouwer, A. P. M. (2023). Missense MED12 variants in 22 males with intellectual disability: From nonspecific symptoms to complete syndromes. *American Journal of Medical Genetics. Part A*, 191(1), 135-143. <https://doi.org/10.1002/ajmg.a.63004>
- Maia, N., Nabais Sa, M. J., Melo-Pires, M., de Brouwer, A. P. M. et Jorge, P. (2021). Intellectual disability genomics: current state, pitfalls and future challenges. *BMC Genomics*, 22(1), 909. <https://doi.org/10.1186/s12864-021-08227-4>
- Mainali, A., Athey, T., Bahl, S., Hung, C., Caluseriu, O., Chan, A., Eaton, A., Ghai, S. J., Kannu, P., MacPherson, M., Niederhoffer, K. Y., Siriwardena, K. et Mercimek-Andrews, S. (2023). Diagnostic yield of clinical exome sequencing in adulthood in medical genetics clinics. *American Journal of Medical Genetics. Part A*, 191(2), 510-517. <https://doi.org/10.1002/ajmg.a.63053>
- Mak, A. C. Y., Lai, Y. Y. Y., Lam, E. T., Kwok, T.-P., Leung, A. K. Y., Poon, A., Mostovoy, Y., Hastie, A. R., Stedman, W., Anantharaman, T., Andrews, W., Zhou, X., Pang, A. W. C., Dai, H., Chu, C., Lin, C., Wu, J. J. K., Li, C. M. L., Li, J.-W., ... Kwok, P.-Y. (2016). Genome-Wide Structural Variation Detection by Genome Mapping on Nanochannel Arrays. *Genetics*, 202(1), 351-362. <https://doi.org/10.1534/genetics.115.183483>
- Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorf, L. A., Hunter, D. J., McCarthy, M. I., Ramos, E. M., Cardon, L. R., Chakravarti, A., Cho, J. H., Guttmacher, A. E., Kong, A., Kruglyak, L., Mardis, E., Rotimi, C. N., Slatkin, M., Valle, D., Whittemore, A. S., ... Visscher, P. M. (2009). Finding the missing heritability of complex diseases. *Nature*, 461(7265), 747-753. <https://doi.org/10.1038/nature08494>
- Mantere, T., Neveling, K., Pebrel-Richard, C., Benoist, M., van der Zande, G., Kater-Baats, E., Baatout, I., van Beek, R., Yammine, T., Oorsprong, M., Hsoumi, F., Olde-Weghuis, D., Majdali, W., Vermeulen, S., Pauper, M., Lebbar, A., Stevens-Kroef, M., Sanlaville, D., Dupont, J. M., ... El Khattabi, L. (2021). Optical genome mapping enables constitutional chromosomal aberration detection. *American Journal of Human Genetics*, 108(8), 1409-1422. <https://doi.org/10.1016/j.ajhg.2021.05.012>
- Mardis, E. R. (2011). A decade's perspective on DNA sequencing technology. *Nature*, 470(7333), 198-203. <https://doi.org/10.1038/nature09796>
- Martin, H. C., Gardner, E. J., Samocha, K. E., Kaplanis, J., Akawi, N., Sifrim, A., Eberhardt, R. Y., Tavares, A. L. T., Neville, M. D. C., Niemi, M. E. K., Gallone, G., McRae, J., Wright, C. F., FitzPatrick, D. R., Firth, H. V. et Hurles, M. E. (2021). The contribution of X-linked coding variation to severe developmental disorders. *Nature Communications*, 12(1), 627. <https://doi.org/10.1038/s41467-020-20852-3>
- Martin, J. P. et Bell, J. (1943). A PEDIGREE OF MENTAL DEFECT SHOWING SEX-LINKAGE. *Journal of Neurology and Psychiatry*, 6(3-4), 154-157. <https://doi.org/10.1136/jnnp.6.3-4.154>

- Mattioli, F., Isidor, B., Abdul-Rahman, O., Gunter, A., Huang, L., Kumar, R., Beaulieu, C., Gecz, J., Innes, M., Mandel, J.-L. et Piton, A. (2019). Clinical and functional characterization of recurrent missense variants implicated in *THOC6* -related intellectual disability. *Human Molecular Genetics*, 28(6), 952-960. <https://doi.org/10.1093/hmg/ddy391>
- Mattioli, F., Piton, A., Gérard, B., Superti-Furga, A., Mandel, J.-L. et Unger, S. (2016). Novel de novo mutations in ZBTB20 in Primrose syndrome with congenital hypothyroidism. *American Journal of Medical Genetics. Part A*, 170(6), 1626-1629. <https://doi.org/10.1002/ajmg.a.37645>
- Mattioli, F., Schaefer, E., Magee, A., Mark, P., Mancini, G. M., Dieterich, K., Von Allmen, G., Alders, M., Coutton, C., van Slegtenhorst, M., Vieville, G., Engelen, M., Cobben, J. M., Juusola, J., Pujol, A., Mandel, J.-L. et Piton, A. (2017). Mutations in Histone Acetylase Modifier BRPF1 Cause an Autosomal-Dominant Form of Intellectual Disability with Associated Ptosis. *American Journal of Human Genetics*, 100(1), 105-116. <https://doi.org/10.1016/j.ajhg.2016.11.010>
- Maxam, A. M. et Gilbert, W. (1977). A new method for sequencing DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 74(2), 560-564. <https://doi.org/10.1073/pnas.74.2.560>
- Mayo, S., Monfort, S., Rosello, M., Orellana, C., Oltra, S., Armstrong, J., Catala, V. et Martinez, F. (2011). De novo interstitial triplication of MECP2 in a girl with neurodevelopmental disorder and random X chromosome inactivation. *Cytogenetic and Genome Research*, 135(2), 93-101. <https://doi.org/10.1159/000330917>
- McDermott, H., Sherlaw-Sturrock, C., Baptista, J., Hartles-Spencer, L. et Naik, S. (2022). Rapid exome sequencing in critically ill children impacts acute and long- term management of patients and their families: A retrospective regional evaluation. *European Journal of Medical Genetics*, 65(9), 104571. <https://doi.org/10.1016/j.ejmg.2022.104571>
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. et DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next- generation DNA sequencing data. *Genome Research*, 20(9), 1297-1303. <https://doi.org/10.1101/gr.107524.110>
- Meienberg, J., Bruggmann, R., Oexle, K. et Matyas, G. (2016). Clinical sequencing: is WGS the better WES? *Human Genetics*, 135(3), 359-362. <https://doi.org/10.1007/s00439-015-1631-9>
- Metzker, M. L. (2010). Sequencing technologies - the next generation. *Nature Reviews. Genetics*, 11(1), 31-46. <https://doi.org/10.1038/nrg2626>
- Migeon, B. R. (2016). An overview of X inactivation based on species differences. *Seminars in Cell & Developmental Biology*, 56, 111-116. <https://doi.org/10.1016/j.semcd.2016.01.024>
- Mignot, C., McMahon, A. C., Bar, C., Campeau, P. M., Davidson, C., Buratti, J., Nava, C., Jacquemont, M.-L., Tallot, M., Milh, M., Edery, P., Marzin, P., Barcia, G., Barnerias, C., Besmond, C., Bienvenu, T., Bruel, A.-L., Brunga, L., Ceulemans, B., ... Depienne, C. (2019). IQSEC2-related encephalopathy in males and females: a comparative study including 37 novel patients. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, 21(4), 837-849. <https://doi.org/10.1038/s41436-018-0268-1>
- Miguet, M., Faivre, L., Amiel, J., Nizon, M., Touraine, R., Prieur, F., Pasquier, L., Lefebvre, M., Thevenon, J., Dubourg, C., Julia, S., Sarret, C., Remerand, G., Francannet, C., Laffargue, F., Boespflug-Tanguy, O., David, A., Isidor, B., Vigneron, J., ... El Chehadeh, S. (2018). Further delineation of the *MECP2* duplication syndrome phenotype in 59 French male patients, with a particular focus on morphological and neurological features. *Journal of Medical Genetics*, 55(6), 359-371. <https://doi.org/10.1136/jmedgenet-2017-104956>
- Mills, R. E., Bennett, E. A., Iskow, R. C. et Devine, S. E. (2007). Which transposable elements are active in the human genome? *Trends in Genetics : TIG*, 23(4), 183-191. <https://doi.org/10.1016/j.tig.2007.02.006>
- Mlika, A., du Mazaubrun, C. et Rumeau-Rouquette, C. (1993). [Prevalence of severe mental retardation and trisomy 21 in 3 generations: 1972, 1976, and 1981]. *Revue D'épidemiologie Et De Santé Publique*, 41(1), 44-52.
- Moey, C., Hinze, S. J., Brueton, L., Morton, J., McMullan, D. J., Kamien, B., Barnett, C. P., Brunetti-Pierri, N., Nicholl, J., Gecz, J. et Shoubridge, C. (2016). Xp11.2 microduplications including IQSEC2, TSPYL2 and KDM5C genes in patients with neurodevelopmental disorders. *European Journal of Human Genetics : EJHG*, 24(3), 373-380. <https://doi.org/10.1038/ejhg.2015.123>
- Monies, D., Goljan, E., Assoum, M., Albrecan, M., Binhumaid, F., Subhani, S., Boureggah, A., Hashem, M., Abdulwahab, F., Abuyousef, O., Temsah, M. H., Alshime, F., Kelaher, J., Abouelhoda, M., Meyer, B. F. et Alkuraya, F. S. (2023). The clinical utility of rapid exome sequencing in a consanguineous population. *Genome Medicine*, 15(1), 44. <https://doi.org/10.1186/s13073-023-01192-5>
- Monk, M. et Harper, M. I. (1979). Sequential X chromosome inactivation coupled with cellular differentiation in early mouse embryos. *Nature*, 281(5729), 311-313. <https://doi.org/10.1038/281311a0>

- Monroe, G. R., Frederix, G. W., Savelberg, S. M. C., de Vries, T. I., Duran, K. J., van der Smagt, J. J., Terhal, P. A., van Hasselt, P. M., Kroes, H. Y., Verhoeven-Duif, N. M., Nijman, I. J., Carbo, E. C., van Gassen, K. L., Knoers, N. V., Hövels, A. M., van Haelst, M. M., Visser, G. et van Haafden, G. (2016). Effectiveness of whole-exome sequencing and costs of the traditional diagnostic trajectory in children with intellectual disability. *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, 18(9), 949-956. <https://doi.org/10.1038/gim.2015.200>
- Montillot, C., Skutunova, E., Ayushma, Dubied, M., Lahmar, A., Nguyen, S., Peerally, B., Prin, F., Duffourd, Y., Thauvin-Robinet, C., Duplomb, L., Wang, H., Ansar, M., Faivre, L., Navarro, N., Minocha, S., Collins, S. C. et Yalcin, B. (2023). Characterization of Vps13b-mutant mice reveals neuroanatomical and behavioral phenotypes with females less affected. *Neurobiology of Disease*, 185, 106259. <https://doi.org/10.1016/j.nbd.2023.106259>
- Moortgat, S., Berland, S., Aukrust, I., Maystadt, I., Baker, L., Benoit, V., Caro-Llopis, A., Cooper, N. S., Debray, F.-G., Faivre, L., Gardeitchik, T., Haukanes, B. I., Houge, G., Kivuva, E., Martinez, F., Mehta, S. G., Nassogne, M.-C., Powell-Hamilton, N., Pfundt, R., ... Newbury-Ecob, R. A. (2018). HUWE1 variants cause dominant X-linked intellectual disability: a clinical study of 21 patients. *European Journal of Human Genetics : EJHG*, 26(1), 64-74. <https://doi.org/10.1038/s41431-017-0038-6>
- Moortgat, S., Lederer, D., Deprez, M., Buzatu, M., Clapuyt, P., Boulanger, S., Benoit, V., Mary, S., Guichet, A., Ziegler, A., Colin, E., Bonneau, D. et Maystadt, I. (2018). Expanding the phenotypic spectrum associated with OPHN1 mutations: Report of 17 individuals with intellectual disability but no cerebellar hypoplasia. *European Journal of Medical Genetics*, 61(8), 442-450. <https://doi.org/10.1016/j.ejmg.2018.03.002>
- Mowat, D. R., Wilson, M. J. et Goossens, M. (2003). Mowat-Wilson syndrome. *Journal of Medical Genetics*, 40(5), 305-310. <https://doi.org/10.1136/jmg.40.5.305>
- Mueller, W. F., Larsen, L. S. Z., Garibaldi, A., Hatfield, G. W. et Hertel, K. J. (2015). The Silent Sway of Splicing by Synonymous Substitutions. *The Journal of Biological Chemistry*, 290(46), 27700-27711. <https://doi.org/10.1074/jbc.M115.684035>
- Muller, J., Stoetzel, C., Vincent, M. C., Leitch, C. C., Laurier, V., Danse, J. M., Helle, S., Marion, V., Bennouna-Greene, V., Vicaire, S., Megarbane, A., Kaplan, J., Drouin-Garraud, V., Hamdani, M., Sigaudy, S., Francannet, C., Roume, J., Bitoun, P., Goldenberg, A., ... Dollfus, H. (2010). Identification of 28 novel mutations in the Bardet-Biedl syndrome genes: the burden of private mutations in an extensively heterogeneous disease. *Human Genetics*, 127(5), 583-593. <https://doi.org/10.1007/s00439-010-0804-9>
- Musante, L. et Ropers, H. H. (2014). Genetics of recessive cognitive disorders. *Trends in Genetics: TIG*, 30(1), 32-39. <https://doi.org/10.1016/j.tig.2013.09.008>
- Muthusamy, B., Bellad, A., Girimaji, S. C. et Pandey, A. (2021). Shukla-Vernon Syndrome: A Second Family with a Novel Variant in the BCORL1 Gene. *Genes*, 12(3), 452. <https://doi.org/10.3390/genes12030452>
- Najm, J., Horn, D., Wimplinger, I., Golden, J. A., Chizhikov, V. V., Sudi, J., Christian, S. L., Ullmann, R., Kuechler, A., Haas, C. A., Flubacher, A., Charnas, L. R., Uyanik, G., Frank, U., Klopocki, E., Dobyns, W. B. et Kutsche, K. (2008). Mutations of CASK cause an X-linked brain malformation phenotype with microcephaly and hypoplasia of the brainstem and cerebellum. *Nature Genetics*, 40(9), 1065-1067. <https://doi.org/10.1038/ng.194>
- Nashabat, M., Al Qahtani, X. S., Almakdub, S., Altwajiri, W., Ba-Armah, D. M., Hundallah, K., Al Hashem, A., Al Tala, S., Maddirevula, S., Alkuraya, F. S., Tabarki, B. et Alfadhel, M. (2019). The landscape of early infantile epileptic encephalopathy in a consanguineous population. *Seizure*, 69, 154-172. <https://doi.org/10.1016/j.seizure.2019.04.018>
- Navarro-Cobos, M. J., Balaton, B. P. et Brown, C. J. (2020). Genes that escape from X-chromosome inactivation: Potential contributors to Klinefelter syndrome. *American Journal of Medical Genetics. Part C, Seminars in Medical Genetics*, 184(2), 226-238. <https://doi.org/10.1002/ajmg.c.31800>
- Neri, G., Schwartz, C. E., Lubs, H. A. et Stevenson, R. E. (2018). X-linked intellectual disability update 2017. *American Journal of Medical Genetics. Part A*, 176(6), 1375-1388. <https://doi.org/10.1002/ajmg.a.38710>
- Neveling, K., Mantere, T., Vermeulen, S., Oorsprong, M., van Beek, R., Kater-Baats, E., Pauper, M., van der Zande, G., Smeets, D., Weghuis, D. O., Stevens-Kroef, M. J. P. L. et Hoischen, A. (2021). Next-generation cytogenetics: Comprehensive assessment of 52 hematological malignancy genomes by optical genome mapping. *American Journal of Human Genetics*, 108(8), 1423-1435. <https://doi.org/10.1016/j.ajhg.2021.06.001>
- Ng, P. C. et Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Research*, 31(13), 3812-3814. <https://doi.org/10.1093/nar/gkg509>

- Ng, S. B., Bigham, A. W., Buckingham, K. J., Hannibal, M. C., McMillin, M. J., Gildersleeve, H. I., Beck, A. E., Tabor, H. K., Cooper, G. M., Mefford, H. C., Lee, C., Turner, E. H., Smith, J. D., Rieder, M. J., Yoshiura, K.-I., Matsumoto, N., Ohta, T., Niikawa, N., Nickerson, D. A., ... Shendure, J. (2010). Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nature Genetics*, *42*(9), 790-793. <https://doi.org/10.1038/ng.646>
- Ng, S. B., Buckingham, K. J., Lee, C., Bigham, A. W., Tabor, H. K., Dent, K. M., Huff, C. D., Shannon, P. T., Jabs, E. W., Nickerson, D. A., Shendure, J. et Bamshad, M. J. (2010). Exome sequencing identifies the cause of a mendelian disorder. *Nature Genetics*, *42*(1), 30-35. <https://doi.org/10.1038/ng.499>
- Ng, S. B., Turner, E. H., Robertson, P. D., Flygare, S. D., Bigham, A. W., Lee, C., Shaffer, T., Wong, M., Bhattacharjee, A., Eichler, E. E., Bamshad, M., Nickerson, D. A. et Shendure, J. (2009). Targeted capture and massively parallel sequencing of 12 human exomes. *Nature*, *461*(7261), 272-276. <https://doi.org/10.1038/nature08250>
- Nielsen, K. B., Tommerup, N., Poulsen, H. et Mikkelsen, M. (1981). X-linked mental retardation with fragile X. A pedigree showing transmission by apparently unaffected males and partial expression in female carriers. *Human Genetics*, *59*(1), 23-25. <https://doi.org/10.1007/BF00278849>
- Nimwegen, K. V. (2017). *Health Technology Assessment of next-generation sequencing*. <https://www.semanticscholar.org/paper/Health-Technology-Assessment-of-next-generation-Nimwegen/4ebf57481b42b15888be63eab090eeba3c1a5e9f>
- Nimwegen, K. van, Vissers, L., Willemsen, M., Schieving, J., Veltman, J., Wilt, G. van D. et Grutters, J. P. (2016). The Cost-Effectiveness of whole-Exome Sequencing in Complex Paediatric Neurology. *Value in Health*, *19*(7), A695. <https://doi.org/10.1016/j.jval.2016.09.1998>
- Niranjan, T. S., Skinner, C., May, M., Turner, T., Rose, R., Stevenson, R., Schwartz, C. E. et Wang, T. (2015). Affected kindred analysis of human X chromosome exomes to identify novel X-linked intellectual disability genes. *PLoS One*, *10*(2), e0116454. <https://doi.org/10.1371/journal.pone.0116454>
- Novara, F., Simonati, A., Sicca, F., Battini, R., Fiori, S., Contaldo, A., Criscuolo, L., Zuffardi, O. et Ciccone, R. (2014). MECP2 duplication phenotype in symptomatic females: report of three further cases. *Molecular Cytogenetics*, *7*(1), 10. <https://doi.org/10.1186/1755-8166-7-10>
- Nurk, S., Koren, S., Rhie, A., Rautiainen, M., Bizakadze, A. V., Mikheenko, A., Vollger, M. R., Altemose, N., Uralsky, L., Gershman, A., Aganezov, S., Hoyt, S. J., Diekhans, M., Logsdon, G. A., Alonge, M., Antonarakis, S. E., Borchers, M., Bouffard, G. G., Brooks, S. Y., ... Phillippy, A. M. (2022). The complete sequence of a human genome. *Science (New York, N.Y.)*, *376*(6588), 44-53. <https://doi.org/10.1126/science.abj6987>
- Oberle, I., Rousseau, F., Heitz, D., Kretz, C., Devys, D., Hanauer, A., Boue, J., Bertheas, M. F. et Mandel, J. L. (1991). Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. *Science (New York, N.Y.)*, *252*(5009), 1097-1102. <https://doi.org/10.1126/science.252.5009.1097>
- OHNO, S., KAPLAN, W. D. et KINOSITA, R. (1959). Formation of the sex chromatin by a single X-chromosome in liver cells of *Rattus norvegicus*. *Experimental Cell Research*, *18*, 415-418. [https://doi.org/10.1016/0014-4827\(59\)90031-x](https://doi.org/10.1016/0014-4827(59)90031-x)
- Oliver, G. R., Hart, S. N. et Klee, E. W. (2015). Bioinformatics for clinical next generation sequencing. *Clinical Chemistry*, *61*(1), 124-135. <https://doi.org/10.1373/clinchem.2014.224360>
- O'Neill, J. P. (2004). Mutation carrier testing in Lesch-Nyhan syndrome families: HPRT mutant frequency and mutation analysis with peripheral blood T lymphocytes. *Genetic Testing*, *8*(1), 51-64. <https://doi.org/10.1089/109065704323016030>
- Ontario Health (Quality). (2020). Genome-Wide Sequencing for Unexplained Developmental Disabilities or Multiple Congenital Anomalies: A Health Technology Assessment. *Ontario Health Technology Assessment Series*, *20*(11), 1-178.
- Ozsolak, F. et Milos, P. M. (2011). RNA sequencing: advances, challenges and opportunities. *Nature Reviews. Genetics*, *12*(2), 87-98. <https://doi.org/10.1038/nrg2934>
- Palmer, E. E., Stuhlmann, T., Weinert, S., Haan, E., Van Esch, H., Holvoet, M., Boyle, J., Leffler, M., Raynaud, M., Moraine, C., van Bokhoven, H., Kleefstra, T., Kahrizi, K., Najmabadi, H., Ropers, H.-H., Delgado, M. R., Sirsi, D., Golla, S., Sommer, A., ... Kalscheuer, V. M. (2018). De novo and inherited mutations in the X-linked gene CLCN4 are associated with syndromic intellectual disability and behavior and seizure disorders in males and females. *Molecular Psychiatry*, *23*(2), 222-230. <https://doi.org/10.1038/mp.2016.135>
- Palmer, Elizabeth Emma, Sachdev, R., Macintosh, R., Melo, U. S., Mundlos, S., Righetti, S., Kandula, T., Minoche, A. E., Puttick, C., Gayevskiy, V., Hesson, L., Idrisoglu, S., Shoubridge, C., Thai, M. H. N., Davis, R. L., Drew, A. P., Sampaio, H., Andrews, P. I., Lawson, J., ... Kirk, E. (2021). Diagnostic Yield of Whole Genome



- Sequencing After Nondiagnostic Exome Sequencing or Gene Panel in Developmental and Epileptic Encephalopathies. *Neurology*, 96(13), e1770-e1782. <https://doi.org/10.1212/WNL.0000000000011655>
- Paradowska-Stolarz, A. M. (2014). Wolf-Hirschhorn syndrome (WHS) - literature review on the features of the syndrome. *Advances in Clinical and Experimental Medicine : Official Organ Wroclaw Medical University*, 23(3), 485-489. <https://doi.org/10.17219/acem/24111>
- Parisi, M. et Glass, I. (1993). Joubert Syndrome. Dans M. P. Adam, G. M. Mirzaa, R. A. Pagon, S. E. Wallace, L. J. Bean, K. W. Gripp et A. Amemiya (dir.), *GeneReviews(R)*. University of Washington, Seattle.
- Parrini, E., Ramazzotti, A., Dobyns, W. B., Mei, D., Moro, F., Veggiotti, P., Marini, C., Brilstra, E. H., Dalla Bernardina, B., Goodwin, L., Bodell, A., Jones, M. C., Nangeroni, M., Palmeri, S., Said, E., Sander, J. W., Striano, P., Takahashi, Y., Van Maldergem, L., ... Guerrini, R. (2006). Periventricular heterotopia: phenotypic heterogeneity and correlation with Filamin A mutations. *Brain : A Journal of Neurology*, 129(Pt 7), 1892-1906. <https://doi.org/10.1093/brain/awl125>
- Pereira, P. M., Schneider, A., Pannetier, S., Heron, D. et Hanauer, A. (2010). Coffin-Lowry syndrome. *European Journal of Human Genetics : EJHG*, 18(6), 627-633. <https://doi.org/10.1038/ejhg.2009.189>
- Pertea, M., Lin, X. et Salzberg, S. L. (2001). GeneSplicer: a new computational method for splice site prediction. *Nucleic Acids Research*, 29(5), 1185-1190. <https://doi.org/10.1093/nar/29.5.1185>
- Peters, S. U., Fu, C., Suter, B., Marsh, E., Benke, T. A., Skinner, S. A., Lieberman, D. N., Standridge, S., Jones, M., Beisang, A., Feyma, T., Heydeman, P., Ryther, R., Kaufmann, W. E., Glaze, D. G., Neul, J. L. et Percy, A. K. (2019). Characterizing the phenotypic effect of Xq28 duplication size in MECP2 duplication syndrome. *Clinical Genetics*, 95(5), 575-581. <https://doi.org/10.1111/cge.13521>
- Pfundt, R., Del Rosario, M., Vissers, L. E. L. M., Kwint, M. P., Janssen, I. M., de Leeuw, N., Yntema, H. G., Nelen, M. R., Lugtenberg, D., Kamsteeg, E.-J., Wieskamp, N., Stegmann, A. P. A., Stevens, S. J. C., Rodenburg, R. J. T., Simons, A., Mensenkamp, A. R., Rinne, T., Gilissen, C., Scheffer, H., ... Hehir-Kwa, J. Y. (2017). Detection of clinically relevant copy-number variants by exome sequencing in a large cohort of genetic disorders. *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, 19(6), 667-675. <https://doi.org/10.1038/gim.2016.163>
- Philips, A. K., Siren, A., Avela, K., Somer, M., Peippo, M., Ahvenainen, M., Doagu, F., Arvio, M., Kaariainen, H., Van Esch, H., Froyen, G., Haas, S. A., Hu, H., Kalscheuer, V. M. et Jarvela, I. (2014). X-exome sequencing in Finnish families with intellectual disability--four novel mutations and two novel syndromic phenotypes. *Orphanet Journal of Rare Diseases*, 9, 49. <https://doi.org/10.1186/1750-1172-9-49>
- Piton, A, Gauthier, J., Hamdan, F. F., Lafrenière, R. G., Yang, Y., Henrion, E., Laurent, S., Noreau, A., Thibodeau, P., Karemera, L., Spiegelman, D., Kuku, F., Duguay, J., Destroismaisons, L., Jolivet, P., Côté, M., Lachapelle, K., Diallo, O., Raymond, A., ... Rouleau, G. A. (2011). Systematic resequencing of X-chromosome synaptic genes in autism spectrum disorder and schizophrenia. *Molecular Psychiatry*, 16(8), 867-880. <https://doi.org/10.1038/mp.2010.54>
- Piton, Amélie, Michaud, J. L., Peng, H., Aradhya, S., Gauthier, J., Mottron, L., Champagne, N., Lafrenière, R. G., Hamdan, F. F., S2D team, Joobor, R., Fombonne, E., Marineau, C., Cossette, P., Dubé, M.-P., Haghghi, P., Drapeau, P., Barker, P. A., Carbonetto, S. et Rouleau, G. A. (2008). Mutations in the calcium-related gene IL1RAPL1 are associated with autism. *Human Molecular Genetics*, 17(24), 3965-3974. <https://doi.org/10.1093/hmg/ddn300>
- Piton, Amélie, Poquet, H., Redin, C., Masurel, A., Lauer, J., Muller, J., Thevenon, J., Herenger, Y., Chancenotte, S., Bonnet, M., Pinoit, J.-M., Huet, F., Thauvin-Robinet, C., Jaeger, A.-S., Le Gras, S., Jost, B., Gérard, B., Peoc'h, K., Launay, J.-M., ... Mandel, J.-L. (2014). 20 ans après: a second mutation in MAOA identified by targeted high-throughput sequencing in a family with altered behavior and cognition. *European Journal of Human Genetics: EJHG*, 22(6), 776-783. <https://doi.org/10.1038/ejhg.2013.243>
- Piton, Amélie, Redin, C. et Mandel, J.-L. (2013). XLID-Causing Mutations and Associated Genes Challenged in Light of Data From Large-Scale Human Exome Sequencing. *The American Journal of Human Genetics*, 93(2), 368-383. <https://doi.org/10.1016/j.ajhg.2013.06.013>
- Pober, B. R. (2010). Williams-Beuren syndrome. *The New England Journal of Medicine*, 362(3), 239-252. <https://doi.org/10.1056/NEJMra0903074>
- Pollard, K. S., Hubisz, M. J., Rosenbloom, K. R. et Siepel, A. (2010). Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Research*, 20(1), 110-121. <https://doi.org/10.1101/gr.097857.109>
- Pronicka, E., Piekutowska-Abramczuk, D., Ciara, E., Trubicka, J., Rokicki, D., Karkucińska-Więckowska, A., Pajdowska, M., Jurkiewicz, E., Halat, P., Kosińska, J., Pollak, A., Rydzanicz, M., Stawinski, P., Pronicki, M., Krajewska-Walasek, M. et Płoski, R. (2016). New perspective in diagnostics of mitochondrial disorders:

- two years' experience with whole-exome sequencing at a national paediatric centre. *Journal of Translational Medicine*, 14(1), 174. <https://doi.org/10.1186/s12967-016-0930-9>
- Prontera, P., Ottaviani, V., Rogaia, D., Isidori, I., Mencarelli, A., Malerba, N., Cocciadiferro, D., Rolph, P., Stangoni, G., Vulto-van Silfhout, A. et Merla, G. (2016). A novel MED12 mutation: Evidence for a fourth phenotype. *American Journal of Medical Genetics. Part A*, 170(9), 2377-2382. <https://doi.org/10.1002/ajmg.a.37805>
- Proud, V. K., Levine, C. et Carpenter, N. J. (1992). New X-linked syndrome with seizures, acquired micrencephaly, and agenesis of the corpus callosum. *American Journal of Medical Genetics*, 43(1-2), 458-466. <https://doi.org/10.1002/ajmg.1320430169>
- Provenzano, A., La Barbera, A., Lai, F., Perra, A., Farina, A., Cariati, E., Zuffardi, O. et Giglio, S. (2022). Non-Invasive Detection of a De Novo Frameshift Variant of STAG2 in a Female Fetus: Escape Genes Influence the Manifestation of X-Linked Diseases in Females. *Journal of Clinical Medicine*, 11(14), 4182. <https://doi.org/10.3390/jcm11144182>
- Puig, J. G., Mateos, F. A., Torres, R. J. et Buno, A. S. (1998). Purine metabolism in female heterozygotes for hypoxanthine-guanine phosphoribosyltransferase deficiency. *European Journal of Clinical Investigation*, 28(11), 950-957. <https://doi.org/10.1046/j.1365-2362.1998.00392.x>
- Quartier, A., Courraud, J., Thi Ha, T., McGillivray, G., Isidor, B., Rose, K., Drouot, N., Savidan, M., Feger, C., Jagline, H., Chelly, J., Shaw, M., Laumonnier, F., Gecz, J., Mandel, J. et Piton, A. (2019). Novel mutations in NLGN3 causing autism spectrum disorder and cognitive impairment. *Human Mutation*, 40(11), 2021-2032. <https://doi.org/10.1002/humu.23836>
- Quartier, A., Poquet, H., Gilbert-Dussardier, B., Rossi, M., Casteleyn, A.-S., Portes, V. D., Feger, C., Nourisson, E., Kuentz, P., Redin, C., Thevenon, J., Mosca-Boidron, A.-L., Callier, P., Muller, J., Lesca, G., Huet, F., Geoffroy, V., El Chehadeh, S., Jung, M., ... Piton, A. (2017). Intragenic FMR1 disease-causing variants: a significant mutational mechanism leading to Fragile-X syndrome. *European Journal of Human Genetics*, 25(4), 423-431. <https://doi.org/10.1038/ejhg.2016.204>
- Rauch, A., Wieczorek, D., Graf, E., Wieland, T., Endeley, S., Schwarzmayr, T., Albrecht, B., Bartholdi, D., Beygo, J., Di Donato, N., Dufke, A., Cremer, K., Hempel, M., Horn, D., Hoyer, J., Joset, P., Ropke, A., Moog, U., Riess, A., ... Strom, T. M. (2012). Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet (London, England)*, 380(9854), 1674-1682. [https://doi.org/10.1016/S0140-6736\(12\)61480-9](https://doi.org/10.1016/S0140-6736(12)61480-9)
- Rausch, T., Zichner, T., Schlattl, A., Stutz, A. M., Benes, V. et Korbel, J. O. (2012). DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics (Oxford, England)*, 28(18), i333-i339. <https://doi.org/10.1093/bioinformatics/bts378>
- Redin, C., Brand, H., Collins, R. L., Kammin, T., Mitchell, E., Hodge, J. C., Hanscom, C., Pillalamarri, V., Seabra, C. M., Abbott, M.-A., Abdul-Rahman, O. A., Aberg, E., Adley, R., Alcaraz-Estrada, S. L., Alkuraya, F. S., An, Y., Anderson, M.-A., Antolik, C., Anyane-Yeboah, K., ... Talkowski, M. E. (2017). The genomic landscape of balanced cytogenetic abnormalities associated with human congenital anomalies. *Nature Genetics*, 49(1), 36-45. <https://doi.org/10.1038/ng.3720>
- 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., ... Piton, A. (2014). Efficient strategy for the molecular diagnosis of intellectual disability using targeted high-throughput sequencing. *Journal of Medical Genetics*, 51(11), 724-736. <https://doi.org/10.1136/jmedgenet-2014-102554>
- Reese, M. G., Eeckman, F. H., Kulp, D. et Haussler, D. (1997). Improved splice site detection in Genie. *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology*, 4(3), 311-323. <https://doi.org/10.1089/cmb.1997.4.311>
- RENPENNING, H., GERRARD, J. W., ZALESKI, W. A. et TABATA, T. (1962). Familial sex-linked mental retardation. *Canadian Medical Association Journal*, 87, 954-956.
- Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J. et Kircher, M. (2019). CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Research*, 47(D1), D886-D894. <https://doi.org/10.1093/nar/gky1016>
- Rhie, A., Nurk, S., Cechova, M., Hoyt, S. J., Taylor, D. J., Altemose, N., Hook, P. W., Koren, S., Rautiainen, M., Alexandrov, I. A., Allen, J., Asri, M., Bzikadze, A. V., Chen, N.-C., Chin, C.-S., Diekhans, M., Flicek, P., Formenti, G., Fungtammasan, A., ... Phillippy, A. M. (2023). The complete sequence of a human Y chromosome. *Nature*, 621(7978), 344-354. <https://doi.org/10.1038/s41586-023-06457-y>
- Rice, G. I., Forte, G. M. A., Szykiewicz, M., Chase, D. S., Aeby, A., Abdel-Hamid, M. S., Ackroyd, S., Allcock, R., Bailey, K. M., Balottin, U., Barnerias, C., Bernard, G., Bodemer, C., Botella, M. P., Cereda, C., Chandler, K.

- E., Dabydeen, L., Dale, R. C., De Laet, C., ... Crow, Y. J. (2013). Assessment of interferon-related biomarkers in Aicardi-Goutieres syndrome associated with mutations in TREX1, RNASEH2A, RNASEH2B, RNASEH2C, SAMHD1, and ADAR: a case-control study. *The Lancet. Neurology*, *12*(12), 1159-1169. [https://doi.org/10.1016/S1474-4422\(13\)70258-8](https://doi.org/10.1016/S1474-4422(13)70258-8)
- Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K. et Rehm, H. L. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, *17*(5), 405-424. <https://doi.org/10.1038/gim.2015.30>
- Ritter, A. L., Gold, J., Hayashi, H., Ackermann, A. M., Hanke, S., Skraban, C., Cuddapah, S., Bhoj, E., Li, D., Kuroda, Y., Wen, J., Takeda, R., Bibb, A., El Chehadah, S., Piton, A., Ohl, J., Kukolich, M. K., Nagasaki, K., Kato, K., ... Izumi, K. (2022). Expanding the phenotypic spectrum of ARCN1-related syndrome. *Genetics in Medicine*, *24*(6), 1227-1237. <https://doi.org/10.1016/j.gim.2022.02.005>
- Robinson, J. T., Thorvaldsdottir, H., Turner, D. et Mesirov, J. P. (2023). igv.js: an embeddable JavaScript implementation of the Integrative Genomics Viewer (IGV). *Bioinformatics (Oxford, England)*, *39*(1), btac830. <https://doi.org/10.1093/bioinformatics/btac830>
- Robinson, J. T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G. et Mesirov, J. P. (2011). Integrative genomics viewer. *Nature Biotechnology*, *29*(1), 24-26. <https://doi.org/10.1038/nbt.1754>
- Roeleveld, N., Zielhuis, G. A. et Gabreels, F. (1997). The prevalence of mental retardation: a critical review of recent literature. *Developmental Medicine and Child Neurology*, *39*(2), 125-132. <https://doi.org/10.1111/j.1469-8749.1997.tb07395.x>
- Rojano, E., Seoane, P., Ranea, J. A. G. et Perkins, J. R. (2019). Regulatory variants: from detection to predicting impact. *Briefings in Bioinformatics*, *20*(5), 1639-1654. <https://doi.org/10.1093/bib/bby039>
- Rooney, K. et Sadikovic, B. (2022). DNA Methylation Episignatures in Neurodevelopmental Disorders Associated with Large Structural Copy Number Variants: Clinical Implications. *International Journal of Molecular Sciences*, *23*(14), 7862. <https://doi.org/10.3390/ijms23147862>
- Ropers, H.-H. et Hamel, B. C. J. (2005). X-linked mental retardation. *Nature Reviews. Genetics*, *6*(1), 46-57. <https://doi.org/10.1038/nrg1501>
- Ross, M. T., Grafham, D. V., Coffey, A. J., Scherer, S., McLay, K., Muzny, D., Platzer, M., Howell, G. R., Burrows, C., Bird, C. P., Frankish, A., Lovell, F. L., Howe, K. L., Ashurst, J. L., Fulton, R. S., Sudbrak, R., Wen, G., Jones, M. C., Hurler, M. E., ... Bentley, D. R. (2005). The DNA sequence of the human X chromosome. *Nature*, *434*(7031), 325-337. <https://doi.org/10.1038/nature03440>
- Rudolph, C. D., Rudolph, A. M., Lister, G. E., First, L. R. et Gershon, A. A. (2011). Dans *Rudolph's Pediatrics*, 22e (vol. 1-Book, Section). The McGraw-Hill Companies. [accesspediatrics.mhmedical.com/content.aspx?aid=1000080512](https://accesspediatrics.mhmedical.com/content.aspx?aid=1000080512)
- Sadee, W., Hartmann, K., Seweryn, M., Pietrzak, M., Handelman, S. K. et Rempala, G. A. (2014). Missing heritability of common diseases and treatments outside the protein-coding exome. *Human Genetics*, *133*(10), 1199-1215. <https://doi.org/10.1007/s00439-014-1476-7>
- San Antonio-Arce, V., Fenollar-Cortés, M., Oancea Ionescu, R., DeSantos-Moreno, T., Gallego-Merlo, J., Illana Cámara, F. J. et Cotarelo Pérez, M. C. (2016). MECP2 Duplications in Symptomatic Females: Report on 3 Patients Showing the Broad Phenotypic Spectrum. *Child Neurology Open*, *3*, 2329048X16630673. <https://doi.org/10.1177/2329048X16630673>
- Sanchis-Juan, A., Stephens, J., French, C. E., Gleadall, N., Megy, K., Penkett, C., Shamardina, O., Stirrups, K., Delon, I., Dewhurst, E., Dolling, H., Erwood, M., Grozeva, D., Stefanucci, L., Arno, G., Webster, A. R., Cole, T., Austin, T., Branco, R. G., ... Carss, K. J. (2018). Complex structural variants in Mendelian disorders: identification and breakpoint resolution using short- and long-read genome sequencing. *Genome Medicine*, *10*(1), 95. <https://doi.org/10.1186/s13073-018-0606-6>
- Sanger, F., Nicklen, S. et Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, *74*(12), 5463-5467. <https://doi.org/10.1073/pnas.74.12.5463>
- Sanlaville, D., Prieur, M., de Blois, M.-C., Genevieve, D., Lapierre, J.-M., Ozilou, C., Picq, M., Gosset, P., Morichon-Delvallez, N., Munnich, A., Cormier-Daire, V., Baujat, G., Romana, S., Vekemans, M. et Turleau, C. (2005). Functional disomy of the Xq28 chromosome region. *European Journal of Human Genetics: EJHG*, *13*(5), 579-585. <https://doi.org/10.1038/sj.ejhg.5201384>

- Santiago-Algarra, D., Dao, L. T. M., Pradel, L., Espana, A. et Spicuglia, S. (2017). Recent advances in high-throughput approaches to dissect enhancer function. *F1000Research*, 6, 939. <https://doi.org/10.12688/f1000research.11581.1>
- Santos-Reboucas, C. B., de Almeida, L. G., Belet, S., Dos Santos, S. R., Ribeiro, M. G., da Silva, A. F. A., Medina-Acosta, E., Dos Santos, J. M., Goncalves, A. P., Bahia, P. R. V., Pimentel, M. M. G. et Froyen, G. (2015). Novel microduplications at Xp11.22 including HUWE1: clinical and molecular insights into these genomic rearrangements associated with intellectual disability. *Journal of Human Genetics*, 60(4), 207-211. <https://doi.org/10.1038/jhg.2015.1>
- Saunders, C. J., Miller, N. A., Soden, S. E., Dinwiddie, D. L., Noll, A., Alnadi, N. A., Andraws, N., Patterson, M. L., Krivohlavek, L. A., Fellis, J., Humphray, S., Saffrey, P., Kingsbury, Z., Weir, J. C., Betley, J., Grocock, R. J., Margulies, E. H., Farrow, E. G., Artman, M., ... Kingsmore, S. F. (2012). Rapid whole-genome sequencing for genetic disease diagnosis in neonatal intensive care units. *Science Translational Medicine*, 4(154), 154ra135. <https://doi.org/10.1126/scitranslmed.3004041>
- Scala, M., Zonneveld-Huijssoon, E., Brienza, M., Mecarelli, O., Van Der Hout, A. H., Zambrelli, E., Turner, K., Zara, F., Peron, A., Vignoli, A. et Striano, P. (2021). De novo ARHGEF9 missense variants associated with neurodevelopmental disorder in females: expanding the genotypic and phenotypic spectrum of ARHGEF9 disease in females. *Neurogenetics*, 22(1), 87-94. <https://doi.org/10.1007/s10048-020-00622-5>
- Schalock, R. L. (2011). The evolving understanding of the construct of intellectual disability. *Journal of Intellectual & Developmental Disability*, 36(4), 223-233. <https://doi.org/10.3109/13668250.2011.624087>
- Schalock, R. L., Luckasson, R. A., Shogren, K. A., Borthwick-Duffy, S., Bradley, V., Buntinx, W. H. E., Coulter, D. L., Craig, E. M., Gomez, S. C., Lachapelle, Y., Reeve, A., Snell, M. E., Spreat, S., Tassé, M. J., Thompson, J. R., Verdugo, M. A., Wehmeyer, M. L. et Yeager, M. H. (2007). The renaming of mental retardation: understanding the change to the term intellectual disability. *Intellectual and Developmental Disabilities*, 45(2), 116-124. [https://doi.org/10.1352/1934-9556\(2007\)45\[116:TROMRU\]2.0.CO;2](https://doi.org/10.1352/1934-9556(2007)45[116:TROMRU]2.0.CO;2)
- Schirwani, S., Novelli, A., Digilio, M. C., Bourn, D., Wilson, V., Roberts, C., Dallapiccola, B. et Hobson, E. (2019). Duplications of GPC3 and GPC4 genes in symptomatic female carriers of Simpson-Golabi-Behmel syndrome type 1. *European Journal of Medical Genetics*, 62(4), 243-247. <https://doi.org/10.1016/j.ejmg.2018.07.022>
- Schluth-Bolard, C., Labalme, A., Cordier, M.-P., Till, M., Nadeau, G., Tevissen, H., Lesca, G., Boutry-Kryza, N., Rossignol, S., Rocas, D., Dubruc, E., Edery, P. et Sanlaville, D. (2013). Breakpoint mapping by next generation sequencing reveals causative gene disruption in patients carrying apparently balanced chromosome rearrangements with intellectual deficiency and/or congenital malformations. *Journal of Medical Genetics*, 50(3), 144-150. <https://doi.org/10.1136/jmedgenet-2012-101351>
- Schobers, G., Schieving, J. H., Yntema, H. G., Pennings, M., Pfundt, R., Derks, R., Hofste, T., de Wijs, I., Wieskamp, N., van den Heuvel, S., Galbany, J. C., Gilissen, C., Nelen, M., Brunner, H. G., Kleefstra, T., Kamsteeg, E.-J., Willemsen, M. A. A. P. et Vissers, L. E. L. M. (2022). Reanalysis of exome negative patients with rare disease: a pragmatic workflow for diagnostic applications. *Genome Medicine*, 14(1), 66. <https://doi.org/10.1186/s13073-022-01069-z>
- Schuurs-Hoeijmakers, J. H. M., Vulto-van Silfhout, A. T., Vissers, L. E. L. M., Van De Vondervoort, I. I. G. M., Van Bon, B. W. M., De Ligt, J., Gilissen, C., Hehir-Kwa, J. Y., Neveling, K., Del Rosario, M., Hira, G., Reitano, S., Vitello, A., Failla, P., Greco, D., Fichera, M., Galesi, O., Kleefstra, T., Grealley, M. T., ... De Brouwer, A. P. M. (2013). Identification of pathogenic gene variants in small families with intellectually disabled siblings by exome sequencing. *Journal of Medical Genetics*, 50(12), 802-811. <https://doi.org/10.1136/jmedgenet-2013-101644>
- Schwartz, C. E., Louie, R. J., Toutain, A., Skinner, C., Friez, M. J. et Stevenson, R. E. (2023). X-Linked intellectual disability update 2022. *American Journal of Medical Genetics. Part A*, 191(1), 144-159. <https://doi.org/10.1002/ajmg.a.63008>
- Schwarz, J. M., Rödelberger, C., Schuelke, M. et Seelow, D. (2010). MutationTaster evaluates disease-causing potential of sequence alterations. *Nature Methods*, 7(8), 575-576. <https://doi.org/10.1038/nmeth0810-575>
- Schwarze, K., Buchanan, J., Fermont, J. M., Dreau, H., Tilley, M. W., Taylor, J. M., Antoniou, P., Knight, S. J. L., Camps, C., Pentony, M. M., Kvikstad, E. M., Harris, S., Popitsch, N., Pagnamenta, A. T., Schuh, A., Taylor, J. C. et Wordsworth, S. (2020). The complete costs of genome sequencing: a microcosting study in cancer and rare diseases from a single center in the United Kingdom. *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, 22(1), 85-94. <https://doi.org/10.1038/s41436-019-0618-7>

- Scott Schwoerer, J., Laffin, J., Haun, J., Raca, G., Friez, M. J. et Giampietro, P. F. (2014). MECP2 duplication: possible cause of severe phenotype in females. *American Journal of Medical Genetics. Part A*, 164A(4), 1029-1034. <https://doi.org/10.1002/ajmg.a.36380>
- Seidman, L. J., Buka, S. L., Goldstein, J. M., Horton, N. J., Rieder, R. O. et Tsuang, M. T. (2000). The relationship of prenatal and perinatal complications to cognitive functioning at age 7 in the New England Cohorts of the National Collaborative Perinatal Project. *Schizophrenia Bulletin*, 26(2), 309-321. <https://doi.org/10.1093/oxfordjournals.schbul.a033455>
- Sen, S. K., Han, K., Wang, J., Lee, J., Wang, H., Callinan, P. A., Dyer, M., Cordaux, R., Liang, P. et Batzer, M. A. (2006). Human genomic deletions mediated by recombination between Alu elements. *American Journal of Human Genetics*, 79(1), 41-53. <https://doi.org/10.1086/504600>
- Shapiro, M. B. et Senapathy, P. (1987). RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Research*, 15(17), 7155-7174. <https://doi.org/10.1093/nar/15.17.7155>
- Shendure, J. et Lieberman Aiden, E. (2012). The expanding scope of DNA sequencing. *Nature Biotechnology*, 30(11), 1084-1094. <https://doi.org/10.1038/nbt.2421>
- Sheng, M. et Hoogenraad, C. C. (2007). The postsynaptic architecture of excitatory synapses: a more quantitative view. *Annual Review of Biochemistry*, 76, 823-847. <https://doi.org/10.1146/annurev.biochem.76.060805.160029>
- Shihab, H. A., Rogers, M. F., Gough, J., Mort, M., Cooper, D. N., Day, I. N. M., Gaunt, T. R. et Campbell, C. (2015). An integrative approach to predicting the functional effects of non-coding and coding sequence variation. *Bioinformatics (Oxford, England)*, 31(10), 1536-1543. <https://doi.org/10.1093/bioinformatics/btv009>
- Shimada, S., Okamoto, N., Ito, M., Arai, Y., Momosaki, K., Togawa, M., Maegaki, Y., Sugawara, M., Shimojima, K., Osawa, M. et Yamamoto, T. (2013). MECP2 duplication syndrome in both genders. *Brain & Development*, 35(5), 411-419. <https://doi.org/10.1016/j.braindev.2012.07.010>
- Shimojima, K., Sugawara, M., Shichiji, M., Mukaida, S., Takayama, R., Imai, K. et Yamamoto, T. (2011). Loss-of-function mutation of collybistin is responsible for X-linked mental retardation associated with epilepsy. *Journal of Human Genetics*, 56(8), 561-565. <https://doi.org/10.1038/jhg.2011.58>
- Shukla, A., Girisha, K. M., Somashekar, P. H., Nampoothiri, S., McClellan, R. et Vernon, H. J. (2019). Variants in the transcriptional corepressor *BCORL1* are associated with an X-linked disorder of intellectual disability, dysmorphic features, and behavioral abnormalities. *American Journal of Medical Genetics Part A*, *ajmg.a.61118*. <https://doi.org/10.1002/ajmg.a.61118>
- Siepel, A., Bejerano, G., Pedersen, J. S., Hinrichs, A. S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L. W., Richards, S., Weinstock, G. M., Wilson, R. K., Gibbs, R. A., Kent, W. J., Miller, W. et Haussler, D. (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Research*, 15(8), 1034-1050. <https://doi.org/10.1101/gr.3715005>
- Smedley, D., Schubach, M., Jacobsen, J. O. B., Kohler, S., Zemojtel, T., Spielmann, M., Jager, M., Hochheiser, H., Washington, N. L., McMurry, J. A., Haendel, M. A., Mungall, C. J., Lewis, S. E., Groza, T., Valentini, G. et Robinson, P. N. (2016). A Whole-Genome Analysis Framework for Effective Identification of Pathogenic Regulatory Variants in Mendelian Disease. *American Journal of Human Genetics*, 99(3), 595-606. <https://doi.org/10.1016/j.ajhg.2016.07.005>
- Smith, A. C., Neveling, K. et Kanagal-Shamanna, R. (2022). Optical genome mapping for structural variation analysis in hematologic malignancies. *American Journal of Hematology*, 97(7), 975-982. <https://doi.org/10.1002/ajh.26587>
- Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., Heiner, C., Kent, S. B. et Hood, L. E. (1986). Fluorescence detection in automated DNA sequence analysis. *Nature*, 321(6071), 674-679. <https://doi.org/10.1038/321674a0>
- Snel, B., Lehmann, G., Bork, P. et Huynen, M. A. (2000). STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. *Nucleic Acids Research*, 28(18), 3442-3444. <https://doi.org/10.1093/nar/28.18.3442>
- Snijders Blok, L., Madsen, E., Juusola, J., Gilissen, C., Baralle, D., Reijnders, M. R. F., Venselaar, H., Helsmoortel, C., Cho, M. T., Hoischen, A., Vissers, L. E. L. M., Koemans, T. S., Wissink-Lindhout, W., Eichler, E. E., Romano, C., Van Esch, H., Stumpel, C., Vreeburg, M., Smeets, E., ... Kleefstra, T. (2015). Mutations in DDX3X Are a Common Cause of Unexplained Intellectual Disability with Gender-Specific Effects on Wnt Signaling. *American Journal of Human Genetics*, 97(2), 343-352. <https://doi.org/10.1016/j.ajhg.2015.07.004>

- Sobreira, N., Schiettecatte, F., Valle, D. et Hamosh, A. (2015). GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Human Mutation*, 36(10), 928-930. <https://doi.org/10.1002/humu.22844>
- Soemedi, R., Cygan, K. J., Rhine, C. L., Wang, J., Bulacan, C., Yang, J., Bayrak-Toydemir, P., McDonald, J. et Fairbrother, W. G. (2017). Pathogenic variants that alter protein code often disrupt splicing. *Nature Genetics*, 49(6), 848-855. <https://doi.org/10.1038/ng.3837>
- Song, M., Giza, J., Proenca, C. C., Jing, D., Elliott, M., Dincheva, I., Shmelkov, S. V., Kim, J., Schreiner, R., Huang, S.-H., Castren, E., Prekeris, R., Hempstead, B. L., Chao, M. V., Dictenberg, J. B., Rafii, S., Chen, Z.-Y., Rodriguez-Boulant, E. et Lee, F. S. (2015). Slitrk5 Mediates BDNF-Dependent TrkB Receptor Trafficking and Signaling. *Developmental Cell*, 33(6), 690-702. <https://doi.org/10.1016/j.devcel.2015.04.009>
- Song, M., Mathews, C. A., Stewart, S. E., Shmelkov, S. V., Mezey, J. G., Rodriguez-Flores, J. L., Rasmussen, S. A., Britton, J. C., Oh, Y.-S., Walkup, J. T., Lee, F. S. et Glatt, C. E. (2017). Rare Synaptogenesis-Impairing Mutations in SLITRK5 Are Associated with Obsessive Compulsive Disorder. *PLoS One*, 12(1), e0169994. <https://doi.org/10.1371/journal.pone.0169994>
- Song, P. et Perkins, B. D. (2018). Developmental expression of the zebrafish Arf-like small GTPase paralogs arl13a and arl13b. *Gene Expression Patterns : GEP*, 29, 82-87. <https://doi.org/10.1016/j.gep.2018.07.002>
- Stamberger, H., Hammer, T. B., Gardella, E., Vlaskamp, D. R. M., Bertelsen, B., Mandelstam, S., de Lange, I., Zhang, J., Myers, C. T., Fenger, C., Afawi, Z., Almanza Fuerte, E. P., Andrade, D. M., Balci, Y., Ben Zeev, B., Bennett, M. F., Berkovic, S. F., Isidor, B., Bouman, A., ... Scheffer, I. E. (2021). NEXMIF encephalopathy: an X-linked disorder with male and female phenotypic patterns. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, 23(2), 363-373. <https://doi.org/10.1038/s41436-020-00988-9>
- Stark, Z., Schofield, D., Alam, K., Wilson, W., Mupfeki, N., Macciocca, I., Shrestha, R., White, S. M. et Gaff, C. (2017). Prospective comparison of the cost-effectiveness of clinical whole-exome sequencing with that of usual care overwhelmingly supports early use and reimbursement. *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, 19(8), 867-874. <https://doi.org/10.1038/gim.2016.221>
- Steffensen, A. Y., Dandanell, M., Jonson, L., Ejlersen, B., Gerdes, A.-M., Nielsen, F. C. et Hansen, T. vO. (2014). Functional characterization of BRCA1 gene variants by mini-gene splicing assay. *European Journal of Human Genetics : EJHG*, 22(12), 1362-1368. <https://doi.org/10.1038/ejhg.2014.40>
- Stenson, P. D., Mort, M., Ball, E. V., Evans, K., Hayden, M., Heywood, S., Hussain, M., Phillips, A. D. et Cooper, D. N. (2017). The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. *Human Genetics*, 136(6), 665-677. <https://doi.org/10.1007/s00439-017-1779-6>
- Sterne-Weiler, T., Howard, J., Mort, M., Cooper, D. N. et Sanford, J. R. (2011). Loss of exon identity is a common mechanism of human inherited disease. *Genome Research*, 21(10), 1563-1571. <https://doi.org/10.1101/gr.118638.110>
- Stevenson, R. E. et Schwartz, C. E. (2009). X-linked intellectual disability: unique vulnerability of the male genome. *Developmental Disabilities Research Reviews*, 15(4), 361-368. <https://doi.org/10.1002/ddrr.81>
- Stranger, B. E., Forrest, M. S., Dunning, M., Ingle, C. E., Beazley, C., Thorne, N., Redon, R., Bird, C. P., de Grassi, A., Lee, C., Tyler-Smith, C., Carter, N., Scherer, S. W., Tavaré, S., Deloukas, P., Hurles, M. E. et Dermitzakis, E. T. (2007). Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science (New York, N.Y.)*, 315(5813), 848-853. <https://doi.org/10.1126/science.1136678>
- Strømme, P., Mangelsdorf, M. E., Scheffer, I. E. et Géczy, J. (2002). Infantile spasms, dystonia, and other X-linked phenotypes caused by mutations in Aristaless related homeobox gene, ARX. *Brain & Development*, 24(5), 266-268. [https://doi.org/10.1016/s0387-7604\(02\)00079-7](https://doi.org/10.1016/s0387-7604(02)00079-7)
- Sudhof, T. C. (2018). Towards an Understanding of Synapse Formation. *Neuron*, 100(2), 276-293. <https://doi.org/10.1016/j.neuron.2018.09.040>
- Sudmant, P. H., Rausch, T., Gardner, E. J., Handsaker, R. E., Abyzov, A., Huddleston, J., Zhang, Y., Ye, K., Jun, G., Fritz, M. H.-Y., Konkel, M. K., Malhotra, A., Stutz, A. M., Shi, X., Casale, F. P., Chen, J., Hormozdiari, F., Dayama, G., Chen, K., ... Korb, J. O. (2015). An integrated map of structural variation in 2,504 human genomes. *Nature*, 526(7571), 75-81. <https://doi.org/10.1038/nature15394>
- Suhl, J. A., Muddashetty, R. S., Anderson, B. R., Ifrim, M. F., Visootsak, J., Bassell, G. J. et Warren, S. T. (2015). A 3' untranslated region variant in FMR1 eliminates neuronal activity-dependent translation of FMRP by disrupting binding of the RNA-binding protein HuR. *Proceedings of the National Academy of Sciences of the United States of America*, 112(47), E6553-6561. <https://doi.org/10.1073/pnas.1514260112>
- Sutherland, G. R. (1977). Marker X chromosomes and mental retardation. *The New England Journal of Medicine*, 296(24), 1415. <https://doi.org/10.1056/nejm197706162962423>

- Suvakov, M., Panda, A., Diesh, C., Holmes, I. et Abyzov, A. (2021). CNVpytor: a tool for copy number variation detection and analysis from read depth and allele imbalance in whole-genome sequencing. *GigaScience*, 10(11), giab074. <https://doi.org/10.1093/gigascience/giab074>
- Ta, D., Downs, J., Baynam, G., Wilson, A., Richmond, P. et Leonard, H. (2022). Medical Comorbidities in MECP2 Duplication Syndrome: Results from the International MECP2 Duplication Database. *Children (Basel, Switzerland)*, 9(5), 633. <https://doi.org/10.3390/children9050633>
- Takahashi, H. et Craig, A. M. (2013). Protein tyrosine phosphatases PTPdelta, PTPsigma, and LAR: presynaptic hubs for synapse organization. *Trends in Neurosciences*, 36(9), 522-534. <https://doi.org/10.1016/j.tins.2013.06.002>
- Takahashi, H., Katayama, K.-I., Sohya, K., Miyamoto, H., Prasad, T., Matsumoto, Y., Ota, M., Yasuda, H., Tsumoto, T., Aruga, J. et Craig, A. M. (2012). Selective control of inhibitory synapse development by Slitrk3-PTPdelta trans-synaptic interaction. *Nature Neuroscience*, 15(3), 389-398, S1-2. <https://doi.org/10.1038/nn.3040>
- Tan, Y.-Q., Tan, Y.-Q. et Cheng, D.-H. (2020). Whole-genome mate-pair sequencing of apparently balanced chromosome rearrangements reveals complex structural variations: two case studies. *Molecular Cytogenetics*, 13, 15. <https://doi.org/10.1186/s13039-020-00487-1>
- Tarpey, P., Parnau, J., Blow, M., Woffendin, H., Bignell, G., Cox, C., Cox, J., Davies, H., Edkins, S., Holden, S., Korny, A., Mallya, U., Moon, J., O'Meara, S., Parker, A., Stephens, P., Stevens, C., Teague, J., Donnelly, A., ... Raymond, F. L. (2004). Mutations in the DLG3 gene cause nonsyndromic X-linked mental retardation. *American Journal of Human Genetics*, 75(2), 318-324. <https://doi.org/10.1086/422703>
- Tarpey, P. S., Raymond, F. L., O'Meara, S., Edkins, S., Teague, J., Butler, A., Dicks, E., Stevens, C., Tofts, C., Avis, T., Barthorpe, S., Buck, G., Cole, J., Gray, K., Halliday, K., Harrison, R., Hills, K., Jenkinson, A., Jones, D., ... Partington, M. (2007). Mutations in CUL4B, which encodes a ubiquitin E3 ligase subunit, cause an X-linked mental retardation syndrome associated with aggressive outbursts, seizures, relative macrocephaly, central obesity, hypogonadism, pes cavus, and tremor. *American Journal of Human Genetics*, 80(2), 345-352. <https://doi.org/10.1086/511134>
- Tarpey, P. S., Smith, R., Pleasance, E., Whibley, A., Edkins, S., Hardy, C., O'Meara, S., Latimer, C., Dicks, E., Menzies, A., Stephens, P., Blow, M., Greenman, C., Xue, Y., Tyler-Smith, C., Thompson, D., Gray, K., Andrews, J., Barthorpe, S., ... Stratton, M. R. (2009). A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation. *Nature Genetics*, 41(5), 535-543. <https://doi.org/10.1038/ng.367>
- Tarpey, P. S., Stevens, C., Teague, J., Edkins, S., O'Meara, S., Avis, T., Barthorpe, S., Buck, G., Butler, A., Cole, J., Dicks, E., Gray, K., Halliday, K., Harrison, R., Hills, K., Hinton, J., Jones, D., Menzies, A., Mironenko, T., ... Raymond, F. L. (2006). Mutations in the gene encoding the Sigma 2 subunit of the adaptor protein 1 complex, AP1S2, cause X-linked mental retardation. *American Journal of Human Genetics*, 79(6), 1119-1124. <https://doi.org/10.1086/510137>
- Tatton-Brown, K., Seal, S., Ruark, E., Harmer, J., Ramsay, E., Del Vecchio Duarte, S., Zachariou, A., Hanks, S., O'Brien, E., Aksglaede, L., Baralle, D., Dabir, T., Gener, B., Goudie, D., Homfray, T., Kumar, A., Pilz, D. T., Selicorni, A., Temple, I. K., ... Rahman, N. (2014). Mutations in the DNA methyltransferase gene DNMT3A cause an overgrowth syndrome with intellectual disability. *Nature Genetics*, 46(4), 385-388. <https://doi.org/10.1038/ng.2917>
- Tavtigian, S. V., Greenblatt, M. S., Lesueur, F. et Byrnes, G. B. (2008). In silico analysis of missense substitutions using sequence-alignment based methods. *Human Mutation*, 29(11), 1327-1336. <https://doi.org/10.1002/humu.20892>
- Telomere-to-Telomere*. (s. d.). Genome.gov. <https://www.genome.gov/about-genomics/telomere-to-telomere>
- The Cost of Sequencing a Human Genome*. (s. d.). Genome.gov. <https://www.genome.gov/about-genomics/factsheets/Sequencing-Human-Genome-cost>
- The Human Genome Project*. (s. d.). Genome.gov. <https://www.genome.gov/human-genome-project>
- Théry, J. C., Krieger, S., Gaildrat, P., Révillion, F., Buisine, M.-P., Killian, A., Duponchel, C., Rousselin, A., Vaur, D., Peyrat, J.-P., Berthet, P., Frébourg, T., Martins, A., Hardouin, A. et Tosi, M. (2011). Contribution of bioinformatics predictions and functional splicing assays to the interpretation of unclassified variants of the BRCA genes. *European Journal of Human Genetics: EJHG*, 19(10), 1052-1058. <https://doi.org/10.1038/ejhg.2011.100>
- Theunissen, T. E. J., Nguyen, M., Kamps, R., Hendrickx, A. T., Sallevelt, S. C. E. H., Gottschalk, R. W. H., Calis, C. M., Stassen, A. P. M., de Koning, B., Mulder-Den Hartog, E. N. M., Schoonderwoerd, K., Fuchs, S. A., Hilhorst-Hofstee, Y., de Visser, M., Vanoevelen, J., Szklarczyk, R., Gerards, M., de Coo, I. F. M., Hellebrekers, D. M. E. I. et Smeets, H. J. M. (2018). Whole Exome Sequencing Is the Preferred Strategy to Identify the Genetic

- Defect in Patients With a Probable or Possible Mitochondrial Cause. *Frontiers in Genetics*, 9. <https://www.frontiersin.org/articles/10.3389/fgene.2018.00400>
- Thevenon, J., Duffourd, Y., Masurel-Paulet, A., Lefebvre, M., Feillet, F., El Chehadeh-Djebbar, S., St-Onge, J., Steinmetz, A., Huet, F., Chouchane, M., Darmency-Stamboul, V., Callier, P., Thauvin-Robinet, C., Faivre, L. et Riviere, J. B. (2016). Diagnostic odyssey in severe neurodevelopmental disorders: toward clinical whole-exome sequencing as a first-line diagnostic test. *Clinical Genetics*, 89(6), 700-707. <https://doi.org/10.1111/cge.12732>
- Thorvaldsdottir, H., Robinson, J. T. et Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, 14(2), 178-192. <https://doi.org/10.1093/bib/bbs017>
- Thung, D. T., de Ligt, J., Vissers, L. E. M., Steehouwer, M., Kroon, M., de Vries, P., Slagboom, E. P., Ye, K., Veltman, J. A. et Hehir-Kwa, J. Y. (2014). Mobster: accurate detection of mobile element insertions in next generation sequencing data. *Genome Biology*, 15(10), 488. <https://doi.org/10.1186/s13059-014-0488-x>
- Torene, R. I., Galens, K., Liu, S., Arvai, K., Borroto, C., Scuffins, J., Zhang, Z., Friedman, B., Sroka, H., Heeley, J., Beaver, E., Clarke, L., Neil, S., Walia, J., Hull, D., Juusola, J. et Retterer, K. (2020). Mobile element insertion detection in 89,874 clinical exomes. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, 22(5), 974-978. <https://doi.org/10.1038/s41436-020-0749-x>
- Trask, B. J. (2002). Human cytogenetics: 46 chromosomes, 46 years and counting. *Nature Reviews. Genetics*, 3(10), 769-778. <https://doi.org/10.1038/nrg905>
- Trivisano, M., Santarone, M. E., Micalizzi, A., Ferretti, A., Dentici, M. L., Novelli, A., Vigevano, F. et Specchio, N. (2020). GRIA3 missense mutation is cause of an x-linked developmental and epileptic encephalopathy. *Seizure*, 82, 1-6. <https://doi.org/10.1016/j.seizure.2020.08.032>
- Tukiainen, T., Villani, A.-C., Yen, A., Rivas, M. A., Marshall, J. L., Satija, R., Aguirre, M., Gauthier, L., Fleharty, M., Kirby, A., Cummings, B. B., Castel, S. E., Karczewski, K. J., Aguet, F., Byrnes, A., Lappalainen, T., Regev, A., Ardlie, K. G., Hacohen, N. et MacArthur, D. G. (2017). Landscape of X chromosome inactivation across human tissues. *Nature*, 550(7675), 244-248. <https://doi.org/10.1038/nature24265>
- Tzschach, A., Grasshoff, U., Beck-Woedl, S., Dufke, C., Bauer, C., Kehrer, M., Evers, C., Moog, U., Oehl-Jaschkowitz, B., Di Donato, N., Maiwald, R., Jung, C., Kuechler, A., Schulz, S., Meinecke, P., Spranger, S., Kohlhase, J., Seidel, J., Reif, S., ... Bauer, P. (2015). Next-generation sequencing in X-linked intellectual disability. *European Journal of Human Genetics : EJHG*, 23(11), 1513-1518. <https://doi.org/10.1038/ejhg.2015.5>
- Um, J. W., Kim, K. H., Park, B. S., Choi, Y., Kim, D., Kim, C. Y., Kim, S. J., Kim, M., Ko, J. S., Lee, S.-G., Choi, G., Nam, J., Heo, W. D., Kim, E., Lee, J.-O., Ko, J. et Kim, H. M. (2014). Structural basis for LAR-RPTP/Slitrk complex-mediated synaptic adhesion. *Nature Communications*, 5, 5423. <https://doi.org/10.1038/ncomms6423>
- Um, J. W. et Ko, J. (2013). LAR-RPTPs: synaptic adhesion molecules that shape synapse development. *Trends in Cell Biology*, 23(10), 465-475. <https://doi.org/10.1016/j.tcb.2013.07.004>
- Van der Auwera, G. A. et O'Connor, B. D. (Brian D. (2020). *Genomics in the cloud : using Docker, GATK, and WDL in Terra* (First edition.). O'Reilly Media.
- Van Esch, H., Bauters, M., Ignatius, J., Jansen, M., Raynaud, M., Hollanders, K., Lugtenberg, D., Bienvenu, T., Jensen, L. R., Gecz, J., Moraine, C., Marynen, P., Fryns, J.-P. et Froyen, G. (2005). Duplication of the MECP2 region is a frequent cause of severe mental retardation and progressive neurological symptoms in males. *American Journal of Human Genetics*, 77(3), 442-453. <https://doi.org/10.1086/444549>
- van Bokhoven, H. (2011). Genetic and epigenetic networks in intellectual disabilities. *Annual Review of Genetics*, 45, 81-104. <https://doi.org/10.1146/annurev-genet-110410-132512>
- Verloes, A., Lesenfants, S., Misson, J. P., Galand, A. et Koulischer, L. (1997). Microcephaly, muscular build, rhizomelia, and cataracts: description of a possible recessive syndrome and some comments on the use of electronic databases in syndromology. *American Journal of Medical Genetics*, 68(4), 455-460; discussion 461. [https://doi.org/10.1002/\(sici\)1096-8628\(19970211\)68:4<455::aid-ajmg16>3.0.co;2-r](https://doi.org/10.1002/(sici)1096-8628(19970211)68:4<455::aid-ajmg16>3.0.co;2-r)
- Vicoso, B. et Bachtrog, D. (2009). Progress and prospects toward our understanding of the evolution of dosage compensation. *Chromosome Research : An International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology*, 17(5), 585-602. <https://doi.org/10.1007/s10577-009-9053-y>
- Vincent, A., Heitz, D., Petit, C., Kretz, C., Oberle, I. et Mandel, J. L. (1991). Abnormal pattern detected in fragile-X patients by pulsed-field gel electrophoresis. *Nature*, 349(6310), 624-626. <https://doi.org/10.1038/349624a0>
- Vissers, L. E. L. M., de Ligt, J., Gilissen, C., Janssen, I., Steehouwer, M., de Vries, P., van Lier, B., Arts, P., Wieskamp, N., del Rosario, M., van Bon, B. W. M., Hoischen, A., de Vries, B. B. A., Brunner, H. G. et Veltman, J. A.



- (2010). A de novo paradigm for mental retardation. *Nature Genetics*, 42(12), 1109-1112. <https://doi.org/10.1038/ng.712>
- Vissers, L. E. L. M., de Vries, B. B. A., Osoegawa, K., Janssen, I. M., Feuth, T., Choy, C. O., Straatman, H., van der Vliet, W., Huys, E. H. L. P. G., van Rijk, A., Smeets, D., van Ravenswaaij-Arts, C. M. A., Knoers, N. V., van der Burgt, I., de Jong, P. J., Brunner, H. G., van Kessel, A. G., Schoenmakers, E. F. P. M. et Veltman, J. A. (2003). Array-based comparative genomic hybridization for the genomewide detection of submicroscopic chromosomal abnormalities. *American Journal of Human Genetics*, 73(6), 1261-1270. <https://doi.org/10.1086/379977>
- Vissers, L. E. L. M., Gilissen, C. et Veltman, J. A. (2016). Genetic studies in intellectual disability and related disorders. *Nature Reviews. Genetics*, 17(1), 9-18. <https://doi.org/10.1038/nrg3999>
- Wallerstein, R., Sugalski, R., Cohn, L., Jawetz, R. et Friez, M. (2008). Expansion of the ARX spectrum. *Clinical Neurology and Neurosurgery*, 110(6), 631-634. <https://doi.org/10.1016/j.clineuro.2008.03.007>
- Wang, J., Song, L., Grover, D., Azrak, S., Batzer, M. A. et Liang, P. (2006). dbRIP: a highly integrated database of retrotransposon insertion polymorphisms in humans. *Human Mutation*, 27(4), 323-329. <https://doi.org/10.1002/humu.20307>
- Wang, J.-Y., Zhou, P., Wang, J., Tang, B., Su, T., Liu, X.-R., Li, B.-M., Meng, H., Shi, Y.-W., Yi, Y.-H., He, N. et Liao, W.-P. (2018). ARHGEF9 mutations in epileptic encephalopathy/intellectual disability: toward understanding the mechanism underlying phenotypic variation. *Neurogenetics*, 19(1), 9-16. <https://doi.org/10.1007/s10048-017-0528-2>
- Wang, Q., Chen, P., Liu, J., Lou, J., Liu, Y. et Yuan, H. (2020). Xp11.22 duplications in four unrelated Chinese families: delineating the genotype-phenotype relationship for HSD17B10 and FGD1. *BMC Medical Genomics*, 13(1), 66. <https://doi.org/10.1186/s12920-020-0728-8>
- Wanke, K. A., Devanna, P. et Vernes, S. C. (2018). Understanding Neurodevelopmental Disorders: The Promise of Regulatory Variation in the 3'UTRome. *Biological Psychiatry*, 83(7), 548-557. <https://doi.org/10.1016/j.biopsych.2017.11.006>
- Wassif, C. A., Maslen, C., Kachilele-Linjewile, S., Lin, D., Linck, L. M., Connor, W. E., Steiner, R. D. et Porter, F. D. (1998). Mutations in the human sterol delta7-reductase gene at 11q12-13 cause Smith-Lemli-Opitz syndrome. *American Journal of Human Genetics*, 63(1), 55-62. <https://doi.org/10.1086/301936>
- Watson, J. D. (1990). The Human Genome Project: Past, Present, and Future. *Science*, 248(4951), 44-49.
- WATSON, J. D. et CRICK, F. H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, 171(4356), 737-738. <https://doi.org/10.1038/171737a0>
- Weaving, L. S., Christodoulou, J., Williamson, S. L., Friend, K. L., McKenzie, O. L. D., Archer, H., Evans, J., Clarke, A., Pelka, G. J., Tam, P. P. L., Watson, C., Lahooti, H., Ellaway, C. J., Bennetts, B., Leonard, H. et Gecz, J. (2004). Mutations of CDKL5 cause a severe neurodevelopmental disorder with infantile spasms and mental retardation. *American Journal of Human Genetics*, 75(6), 1079-1093. <https://doi.org/10.1086/426462>
- Weissbach, S., Sys, S., Hewel, C., Todorov, H., Schweiger, S., Winter, J., Pfenninger, M., Torkamani, A., Evans, D., Burger, J., Everschor-Sitte, K., May-Simera, H. L. et Gerber, S. (2021). Reliability of genomic variants across different next-generation sequencing platforms and bioinformatic processing pipelines. *BMC Genomics*, 22(1), 62. <https://doi.org/10.1186/s12864-020-07362-8>
- Weller, S. et Gartner, J. (2001). Genetic and clinical aspects of X-linked hydrocephalus (L1 disease): Mutations in the L1CAM gene. *Human Mutation*, 18(1), 1-12. <https://doi.org/10.1002/humu.1144>
- Whiffin, N., Karczewski, K. J., Zhang, X., Chothani, S., Smith, M. J., Evans, D. G., Roberts, A. M., Quaife, N. M., Schafer, S., Rackham, O., Alfoldi, J., O'Donnell-Luria, A. H., Francioli, L. C., Cook, S. A., Barton, P. J. R., MacArthur, D. G. et Ware, J. S. (2020). Characterising the loss-of-function impact of 5' untranslated region variants in 15,708 individuals. *Nature Communications*, 11(1), 2523. <https://doi.org/10.1038/s41467-019-10717-9>
- Wijngaard, R., Demidov, G., O'Gorman, L., Corominas-Galbany, J., Yaldiz, B., Steyaert, W., de Boer, E., Vissers, L. E. L. M., Kamsteeg, E.-J., Pfundt, R., Swinkels, H., den Ouden, A., te Paske, I. B. A. W., de Voer, R. M., Faivre, L., Denommé-Pichon, A.-S., Duffourd, Y., Vitobello, A., Chevarin, M., ... Gilissen, C. (2023). Mobile element insertions in rare diseases: a comparative benchmark and reanalysis of 60,000 exome samples. *European Journal of Human Genetics*, 1-9. <https://doi.org/10.1038/s41431-023-01478-7>
- Willems, T., Gymrek, M., Highnam, G., Mittelman, D. et Erlich, Y. (2014). The landscape of human STR variation. *Genome Research*, 24(11), 1894-1904. <https://doi.org/10.1101/gr.177774.114>
- Wimmer, K., Roca, X., Beiglbock, H., Callens, T., Etzler, J., Rao, A. R., Krainer, A. R., Fonatsch, C. et Messiaen, L. (2007). Extensive in silico analysis of NF1 splicing defects uncovers determinants for splicing outcome upon 5' splice-site disruption. *Human Mutation*, 28(6), 599-612. <https://doi.org/10.1002/humu.20493>

- Wolf, N. I., van Spaendonk, R. M., Hobson, G. M. et Kamholz, J. (1993). PLP1 Disorders. Dans M. P. Adam, G. M. Mirzaa, R. A. Pagon, S. E. Wallace, L. J. Bean, K. W. Gripp et A. Amemiya (dir.), *GeneReviews*(R). University of Washington, Seattle.
- Wolking, S., May, P., Mei, D., Moller, R. S., Balestrini, S., Helbig, K. L., Altuzarra, C. D., Chatron, N., Kaiwar, C., Stohr, K., Widdess-Walsh, P., Mendelsohn, B. A., Numis, A., Cilio, M. R., Van Paesschen, W., Svendsen, L. L., Oates, S., Hughes, E., Goyal, S., ... Schubert, J. (2019). Clinical spectrum of STX1B-related epileptic disorders. *Neurology*, *92*(11), e1238-e1249. <https://doi.org/10.1212/WNL.0000000000007089>
- Won, S. Y., Lee, P. et Kim, H. M. (2019). Synaptic organizer: Slitrks and type IIa receptor protein tyrosine phosphatases. *Current Opinion in Structural Biology*, *54*, 95-103. <https://doi.org/10.1016/j.sbi.2019.01.010>
- Wright, C. F., Fitzgerald, T. W., Jones, W. D., Clayton, S., McRae, J. F., van Kogelenberg, M., King, D. A., Ambridge, K., Barrett, D. M., Bayzatinova, T., Bevan, A. P., Bragin, E., Chatzimichali, E. A., Gribble, S., Jones, P., Krishnappa, N., Mason, L. E., Miller, R., Morley, K. I., ... Firth, H. V. (2015). Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *Lancet (London, England)*, *385*(9975), 1305-1314. [https://doi.org/10.1016/S0140-6736\(14\)61705-0](https://doi.org/10.1016/S0140-6736(14)61705-0)
- Wright, C. F., Quaipe, N. M., Ramos-Hernandez, L., Danecek, P., Ferla, M. P., Samocha, K. E., Kaplanis, J., Gardner, E. J., Eberhardt, R. Y., Chao, K. R., Karczewski, K. J., Morales, J., Gallone, G., Balasubramanian, M., Banka, S., Gompertz, L., Kerr, B., Kirby, A., Lynch, S. A., ... Whiffin, N. (2021). Non-coding region variants upstream of MEF2C cause severe developmental disorder through three distinct loss-of-function mechanisms. *American Journal of Human Genetics*, *108*(6), 1083-1094. <https://doi.org/10.1016/j.ajhg.2021.04.025>
- Xue, Y., Ankala, A., Wilcox, W. R. et Hegde, M. R. (2015). Solving the molecular diagnostic testing conundrum for Mendelian disorders in the era of next-generation sequencing: single-gene, gene panel, or exome/genome sequencing. *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, *17*(6), 444-451. <https://doi.org/10.1038/gim.2014.122>
- Yao, R., Zhang, Y., Liu, J., Wang, J., Xu, Y., Li, N., Wang, J. et Yu, T. (2020). Clinical and Molecular Characterization of Three Novel ARHGEF9 Mutations in Patients with Developmental Delay and Epilepsy. *Journal of Molecular Neuroscience*, *70*(6), 908-915. <https://doi.org/10.1007/s12031-019-01465-y>
- Yauy, K., de Leeuw, N., Yntema, H. G., Pfundt, R. et Gilissen, C. (2020). Accurate detection of clinically relevant uniparental disomy from exome sequencing data. *Genetics in Medicine: Official Journal of the American College of Medical Genetics*, *22*(4), 803-808. <https://doi.org/10.1038/s41436-019-0704-x>
- Yeo, G. et Burge, C. B. (2004). Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology*, *11*(2-3), 377-394. <https://doi.org/10.1089/1066527041410418>
- Yoshikawa, H., Fujiyama, A., Nakai, K., Inazawa, J. et Matsubara, K. (1998). Detection and Isolation of a Novel Human Gene Located on Xp11.2-p11.4 That Escapes X-Inactivation Using a Two-Dimensional DNA Mapping Method. *Genomics*, *49*(2), 237-246. <https://doi.org/10.1006/geno.1998.5246>
- Zhang, F. et Lupski, J. R. (2015). Non-coding genetic variants in human disease. *Human Molecular Genetics*, *24*(R1), R102-110. <https://doi.org/10.1093/hmg/ddv259>
- Zhao, Z., Chen, X., Dowbaj, A. M., Sljukic, A., Bratlie, K., Lin, L., Fong, E. L. S., Balachander, G. M., Chen, Z., Soragni, A., Huch, M., Zeng, Y. A., Wang, Q. et Yu, H. (2022). Organoids. *Nature Reviews. Methods Primers*, *2*, 94. <https://doi.org/10.1038/s43586-022-00174-y>
- Zhou, C., Sun, Y., Yan, R., Liu, Y., Zuo, E., Gu, C., Han, L., Wei, Y., Hu, X., Zeng, R., Li, Y., Zhou, H., Guo, F. et Yang, H. (2019). Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. *Nature*, *571*(7764), 275-278. <https://doi.org/10.1038/s41586-019-1314-0>
- Ziats, C. A., Schwartz, C. E., Gecz, J., Shaw, M., Field, M. J., Stevenson, R. E. et Neri, G. (2020). X-linked intellectual disability: Phenotypic expression in carrier females. *Clinical Genetics*, *97*(3), 418-425. <https://doi.org/10.1111/cge.13667>
- Zurek, B., Ellwanger, K., Vissers, L. E. L. M., Schule, R., Synofzik, M., Topf, A., de Voer, R. M., Laurie, S., Matalonga, L., Gilissen, C., Ossowski, S., 't Hoen, P. A. C., Vitobello, A., Schulze-Hentrich, J. M., Riess, O., Brunner, H. G., Brookes, A. J., Rath, A., Bonne, G., ... Graessner, H. (2021). Solve-RD: systematic pan-European data sharing and collaborative analysis to solve rare diseases. *European Journal of Human Genetics: EJHG*, *29*(9), 1325-1331. <https://doi.org/10.1038/s41431-021-00859-0>



# Annexes

## Annexe 1 : Niveaux de sévérité du trouble du développement intellectuel – DSM-5

Niveau de sévérité	Domaine conceptuel	Domaine social	Domaine pratique
Léger	<p>Chez les enfants d'âge préscolaire, il peut ne pas y avoir de différence évidente au plan intellectuel. Pour les enfants d'âge scolaire et les adultes, il existe des difficultés à acquérir des compétences scolaires telles que la lecture, l'écriture, le calcul, l'apprentissage de l'heure, la valeur de l'argent, avec besoin d'aide dans un ou plusieurs domaines pour satisfaire aux attentes en rapport avec l'âge. Chez l'adulte, l'abstraction, les fonctions exécutives (c.-à-d. planification, élaboration de stratégies, classement par priorité, flexibilité cognitive), la mémoire à court terme tout autant que l'utilisation des compétences scolaires (p. ex. lecture, gestion de l'argent) sont altérées. Il existe une approche plutôt concrète des problèmes et des solutions par rapport aux adultes du même âge.</p>	<p>Par rapport aux adultes du même âge, le sujet est immature dans ses interactions sociales. Par exemple, il peut avoir des difficultés à percevoir avec acuité les codes sociaux. La communication, la conversation, le langage sont plus concrets ou immatures que ce qui est attendu pour l'âge. Il peut y avoir des difficultés à contrôler l'émotion et le comportement de façon appropriée à l'âge ; ces difficultés sont remarquées par les autres dans la vie sociale. Le sujet a une compréhension limitée des risques dans les situations sociales ; son jugement y est immature, et il court le risque d'être manipulé par les autres (crédulité).</p>	<p>Le sujet peut agir de manière appropriée à son âge pour les soins personnels. Il nécessite cependant, plus que ses pairs, une assistance pour les tâches plus complexes de la vie quotidienne. À l'âge adulte, les aides concernent surtout les achats alimentaires, les transports, la prise en charge des enfants et de la maison, la préparation de repas équilibrés, la gestion des comptes et de l'argent. L'aptitude aux loisirs est peu différente des sujets du même âge, bien que l'appréciation des aspects relatifs au bien-être et à l'organisation durant les distractions nécessite une aide. À l'âge adulte, le sujet peut réussir à trouver un emploi en milieu normal mais dans des fonctions qui ne mettent pas en avant les compétences intellectuelles. Ces personnes ont généralement besoin d'aide pour prendre des décisions médicales et légales, et pour pouvoir mettre à profit avec compétence une formation professionnelle. Un soutien est habituellement nécessaire pour élever une famille.</p>

Suite

Niveau de sévérité	Domaine conceptuel	Domaine social	Domaine pratique
Moyen	<p>Tout au long du développement, les capacités intellectuelles du sujet restent largement en deçà de celles de ses pairs. Pour les enfants non encore scolarisés, le langage et les compétences préscolaires se développent lentement. Pour les enfants scolarisés, les acquisitions en lecture, écriture, calcul, la compréhension de l'heure et la gestion de l'argent progressent lentement au fil des années de scolarité mais sont manifestement limitées par rapport aux autres élèves. Chez les adultes, le développement des capacités intellectuelles reste manifestement à un niveau élémentaire, et une aide est nécessaire pour toute application des apprentissages scolaires dans le monde du travail ou la vie personnelle. Une assistance au long cours est requise pour mener à bien des tâches conceptuelles du quotidien, et il peut s'avérer nécessaire que d'autres en assurent la pleine responsabilité à la place du sujet.</p>	<p>Le sujet, au cours de son développement, montre de grandes différences par rapport aux autres dans la communication et les comportements sociaux. Le langage parlé reste d'évidence le premier moyen de communication mais à un niveau de complexité nettement inférieur à celui des pairs. La capacité de lier des relations est manifeste avec la famille et des amis ; le sujet peut même au cours de sa vie arriver à établir des relations amicales durables voire des relations amoureuses à l'âge adulte. Cependant, les sujets peuvent ne pas percevoir ou interpréter avec finesse les codes sociaux. Le jugement social et les capacités décisionnelles sont limités et des aidants doivent assister la personne dans les décisions importantes de la vie. Les relations amicales avec des pairs non handicapés sont souvent affectées par une communication et une sociabilité limitées. Une aide soutenue, tant au niveau social que relationnel, est nécessaire pour réussir dans le monde du travail.</p>	<p>Le sujet arrivé à l'âge adulte peut assurer ses besoins personnels pour ce qui est de la nourriture, de l'habillement, de l'élimination sphinctérienne, de la toilette, bien qu'une période prolongée d'éducation pour accéder à l'autonomie dans ces domaines soit nécessaire et que des rappels soient parfois indispensables. De même, la participation à toutes les tâches domestiques peut être acquise à l'âge adulte, bien qu'une période prolongée d'éducation soit nécessaire et que des aides suivies soient typiquement indispensables pour accéder à un niveau de performance adulte. Un travail autonome dans des emplois requérant des aptitudes intellectuelles et de communication limitées peut être exercé mais un soutien considérable de la part des collègues de travail, de l'encadrement et des autres est nécessaire pour satisfaire aux attentes sociales, aux difficultés du travail et aux exigences annexes telles que les horaires, les transports, les soins, la gestion de l'argent. La personne peut accéder à des activités de loisirs variées. Cela bien sûr requiert une aide complémentaire et des possibilités d'accès à l'apprentissage sur une période plus longue. Pour une minorité significative, un comportement inadapté est la cause de problèmes sociaux.</p>

Suite

Niveau de sévérité	Domaine conceptuel	Domaine social	Domaine pratique
Grave	<p>L'acquisition des compétences conceptuelles est limitée. Le sujet a habituellement peu de compréhension du langage écrit et des notions impliquant des nombres, des quantités, le temps et l'argent. Les aidants doivent fournir une aide substantielle pour résoudre les problèmes tout au long de la vie.</p>	<p>Le langage parlé est assez limité en termes de vocabulaire et de grammaire. Le discours peut se résumer à des mots ou phrases simples et être complété par des moyens de suppléance. Le discours et la communication sont centrés sur « l'ici et le maintenant » des événements quotidiens. Le langage est plus utilisé à des fins de communication sociale qu'à de l'explication. Les sujets comprennent un discours simple et la communication gestuelle. Les relations avec les membres de la famille et des proches sont une source de plaisir et d'aide.</p>	<p>Le sujet a besoin d'aide pour toute activité du quotidien, ce qui inclut les repas, l'habillage, la toilette, l'élimination. Il nécessite une surveillance de tous les instants. La personne ne peut pas prendre de décisions responsables concernant son bien-être ou celui des autres. À l'âge adulte, la participation à des tâches ménagères, aux distractions et au travail requiert une aide et une assistance permanentes. L'acquisition de compétences en tout domaine nécessite un enseignement prolongé et une aide constante. Un comportement inadapté, incluant l'auto-agressivité, est présent chez une minorité significative de sujets.</p>
Profond	<p>Les compétences intellectuelles sont essentiellement centrées sur le monde physique plutôt que sur le monde symbolique. Le sujet peut utiliser des objets de façon appropriée pour prendre soin de lui, travailler ou se distraire. Quelques compétences visuospatiales, comme assortir et trier des objets selon leurs caractéristiques physiques, peuvent être acquises. Cependant, des déficits sensori-moteurs associés peuvent interdire l'utilisation des objets.</p>	<p>Le sujet a une compréhension très limitée de la communication symbolique, qu'elle soit orale ou gestuelle. Il peut comprendre des instructions ou des gestes simples. La personne exprime très largement ses désirs et ses émotions dans la communication non verbale et non symbolique. Elle trouve du plaisir dans les relations avec les membres de sa famille qu'elle connaît bien, les soignants, les proches, et amorce ou répond aux interactions sociales par des signes gestuels ou émotionnels. Des déficits sensori-moteurs associés peuvent interdire un grand nombre d'activités sociales.</p>	<p>Le sujet est dépendant des autres pour tous les aspects du soin quotidien, de sa santé et de sa sécurité bien qu'il puisse aussi être capable de participer à quelques-unes de ces activités. Les individus indemnes d'atteintes physiques graves peuvent aider à certaines tâches domestiques du quotidien, comme servir à table. Des actions simples utilisant des objets peuvent servir de base de participation à des activités professionnelles qui nécessitent néanmoins de hauts niveaux d'assistance soutenue. Les activités de loisirs comprennent le plaisir à écouter de la musique, regarder des films, se promener, participer à des activités aquatiques, toujours avec un soutien extérieur. Des déficits physiques et sensoriels associés sont de fréquentes entraves à la participation (au-delà d'observer), à ces activités domestiques, de loisirs ou professionnelles. Un comportement inadapté est présent chez une minorité significative.</p>

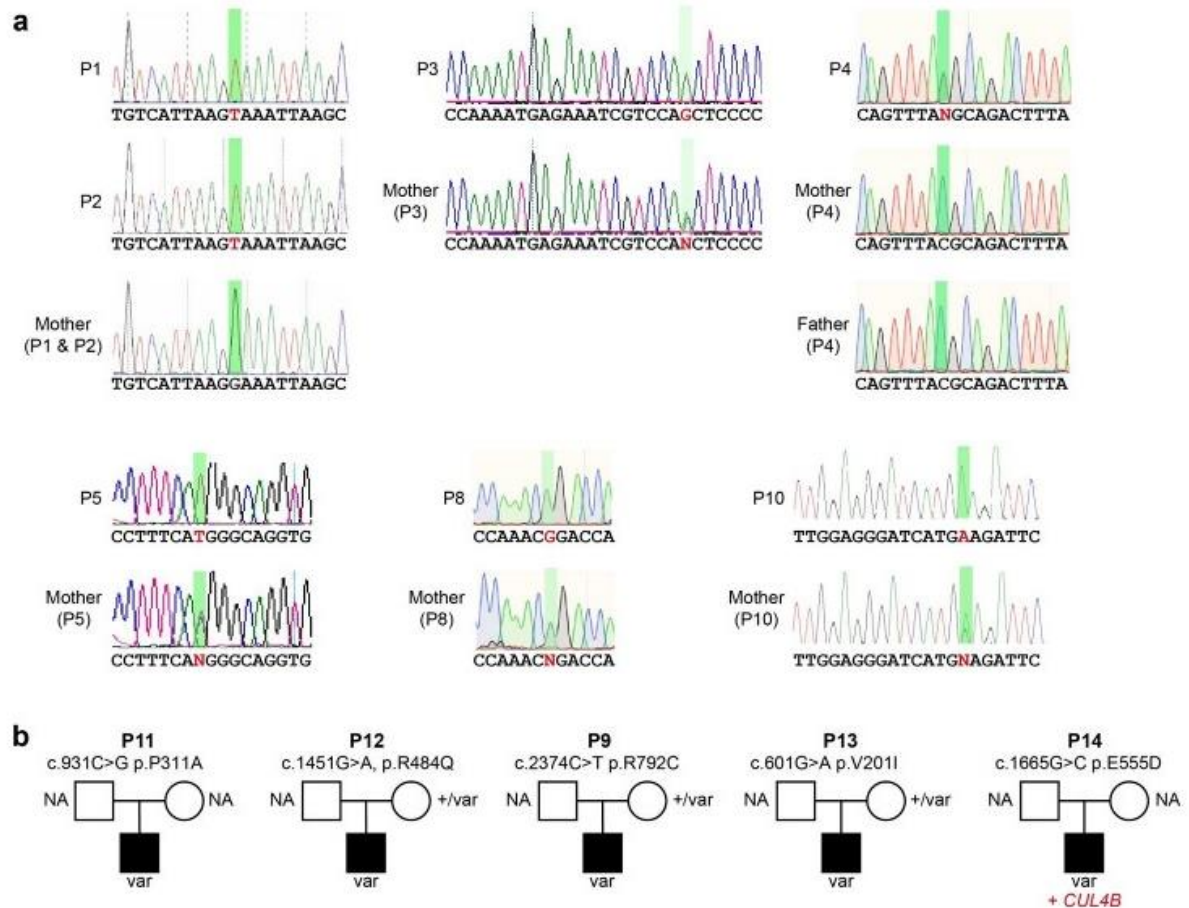
Annexe 2 : Les variations du gène *SLITRK2* dans les troubles du neurodéveloppement (*Supplemental data*)

**Supplementary Information for Chehadeh, Han, Kim, Jang *et al.*, “SLITRK2 variants associated with neurodevelopmental disorders impair excitatory synaptic function and cognition in mice”**

**Supplementary Figures 1 to 20**

**Supplementary Tables 1, 2, 3 and 4**

**Supplementary Notes**



**Supplementary Figure 1. Additional variants identified in *SLITRK2* in individuals with NDD.**

**a** Sanger sequencing traces of P1, P2, P3, P4, P5, P8 and P10 with *SLITRK2* variants. Mutated residues are highlighted in red.

**b** Pedigrees of families with additional rare variants of *SLITRK2*. These variants were not considered to be probably disease-causing because all, except for the E555D variant, have been reported in males from the gnomAD population. The E555D variant was reported in an individual having another molecular diagnosis (a pathogenic variant in *CUL4B*) sufficient to explain his phenotype.



**a**

	L74S	V201I	E210K	
SLITRK1	YHLFLHGNSLRLFPNEF	RLKTLFYEELEIQIPG-IAEILLEDNFW		
SLITRK2	YQLFLNGNLSRLYPNEF	RLKVMFFAGLEHIGG-IAEILQLEENFW		
SLITRK3	FKLYLQRNSMRKLYTNF	RLKVL YRGLDHIGRsLMLQLEENFW		
SLITRK4	YHLNFQNNFNILYFNTE	RIQKLPYIGLEHIGR-VVQLLEDNFW		
SLITRK5	YHLLSGNLNRLYPNEF	RLKLLPYVGLLQHMdk-VVQLQLEENFW		
SLITRK6	FQLSLLNNGTMLHTNDF	QLQTLFVYVGLHIGR-ILDQLLEDNKH		

	P311A T312A	S323N	P374R
SLITRK1	QIKIRPTAAIATGSSRNKPLANS-----		NVSSLADLRLKLSNVQEL
SLITRK2	RPP-KMRNRGPRVTVSKDR--CFGPIIMVYQ		KFTNISDLQKPKTSPPKL
SLITRK3	KQPRTPRPPSISQAL----YPGFNQPIAPVQ		GFNNISELRLPLNAKKL
SLITRK4	NPS-----KISGIVAGKALSNRNLSQIVSYQ		NIQSMSELRLPLNAKKL
SLITRK5	RQPNKPRVGRSRQP-SKDLGYSNYGPIAYQ		KIESIAELQKPYNPKKM
SLITRK6	-----IPYI		NIESLSDLRPPQNPRLK

	R426C	E461*	R484Q
SLITRK1	NTFFKLLDLSRLYMSNY	LNVEYNAIQILPPTFNA	LIILNNLLSLPVDVFA
SLITRK2	GAFTNLSLRLYLNGNY	LYLEYNVIRIKPLTFDA	LLFLNNLLSLPFDNIFG
SLITRK3	GAFINLNLKSLFLNGND	LYFEFNVIRIQPAAFSL	LLFLNNLLSLPTDAFA
SLITRK4	DVFNLTNLSRLYLNGNQ	LYLEYNLIRIKSAGTFDS	LLYLNLLSLPVIYFS
SLITRK5	RAFGLDNLRLYLNGNR	LFLQYNLIRIQSGTFDF	LLFLNNLLQAMPFSGVFS
SLITRK6	GSEFMNLRQLKYLNGNH	LYLEYNVIRIKPLPTFN	VLYLNLLQLVLPHFIS

	V511M	E555D	V589I	R792C
SLITRK1	HNNYFMYLAVAGVLDQLT	RLGSEVLMSDLKCEFPVN	QLYARIS-PTLTSHSK--	-----
SLITRK2	RNNHFSHLVAVGVLDQLP	HANSFVIINAVTCEPAK	DSPNLSDGTILSMN----	QFAPSYESRQSQNQ---DR
SLITRK3	RKNYFLYLVAVGVLEHLN	TISSVSVVGVLDVCRSPEN	EMLHVAPAGESPAQ---P	LFPFGGMQYPDQQDAR
SLITRK4	RNNKFMYLEAVGVLDQLQ	KLSDGIIVKILKCEFPVQ	KLNKPSAPPTSPA---P	---GFEIRYPEKQDCK
SLITRK5	RSNHFTSLVAVGVLDQLK	QLKVGVLVAVVICKAPFK	DYSDVVVSTPTPSIQQV	TFSPNYDLRPHQYGDGR
SLITRK6	KTNQFTHLVAVGVLDQLD	KLSKNTVTDLICTSPGH	GLVNNFSPMTQTSYLM	QLQPDMEAHYPGAHEELK

**b**

	L74S	V201I	E210K	P311A T312A	S323N
Human	QLFLNGNLSRLYPNE	LKVMFFAGLEHIGGIMIQLEENP	RPPKMRNRGPRVTVSKDRQSGPIIMVY		
Gorilla	QLFLNGNLSRLYPNE	LKVMFFAGLEHIGGIMIQLEENP	RPPKMRNRGPRVTVSKDRQSGPIIMVY		
Macaque	QLFLNGNLSRLYPNE	LKVMFFAGLEHIGGIMIQLEENP	RPPKMRNRGPRVTVSKDRQSGPIIMVY		
Bovine	QLFLNGNLSRLYPNE	LKLMFFAGLEHIGGIMIQLEENP	RPPKMRNRGPRVTVSKDRQSGPIIMVY		
Sheep	QLFLNGNLSRLYPNE	LKLMFFAGLEHIGGIMIQLEENP	RPPKMRNRGPRVTVSKDRQSGPIIMVY		
Mouse	QLFLNGNLSRLYPNE	LKVMFFAGLEHIGGIMIQLEENP	RPPKMRNRGPRVTVSKDRQSGPIIMVY		
Rat	QLFLNGNLSRLYPNE	LKVMFFAGLEHIGGIMIQLEENP	RPPKMRNRGPRVTVSKDRQSGPIIMVY		
Dog	QLFLNGNLSRLYPNE	LKVMFFAGLEHIGGIMIQLEENP	RPPKMRNRGPRVTVSKDRQSGPIIMVY		
Zebrafish	QLFLNGNLSRLKISFNE	LKTLFFAGLEHIGGIMIQLEENP	RPPKMRNRGPRVTVSKDRQSGPIIMVY		
Chicken	QLFLNGNLSRLKISFNE	LKMMFFAGLEHIGGIMIQLEENP	RPPKMRNRGPRVTVSKDRQSGPIIMVY		

	P374R	R426C	E461*	R484Q
Human	NISDLQPKTSPPKKLY	AFTNLSLRLYLNGN	YLEYNVIRIKPLTFD	LFLNNLLSLPDNIF
Gorilla	NISDLQPKTSPPKKLY	AFTNLSLRLYLNGN	YLEYNVIRIKPLTFD	LFLNNLLSLPDNIF
Macaque	NISDLQPKTSPPKKLY	AFTNLSLRLYLNGN	YLEYNVIRIKPLTFD	LFLNNLLSLPDNIF
Bovine	NISDLQPKTSPPKKLY	AFTNLSLRLYLNGN	YLEYNVIRIKPLTFD	LFLNNLLSLPDNIF
Sheep	NISDLQPKTSPPKKLY	AFTNLSLRLYLNGN	YLEYNVIRIKPLTFD	LFLNNLLSLPDNIF
Mouse	NISDLQPKTSPPKKLY	AFTNLSLRLYLNGN	YLEYNVIRIKPLTFD	LFLNNLLSLPDNIF
Rat	NISDLQPKTSPPKKLY	AFTNLSLRLYLNGN	YLEYNVIRIKPLTFD	LFLNNLLSLPDNIF
Dog	NISDLQPKTSPPKKLY	AFTNLSLRLYLNGN	YLEYNVIRIKPLTFD	LFLNNLLSLPDNIF
Zebrafish	NISELQPKTSPPKKLH	AFENLSLRLYLNGN	YLEYNVIRIKPLTFN	LFLNNLLSLPDNVF
Chicken	NISELQPKTSPPKKLY	AFTNLSLRLYLNGN	YLEYNVIRIKPLTFD	LFLNNLLSLPDNVF

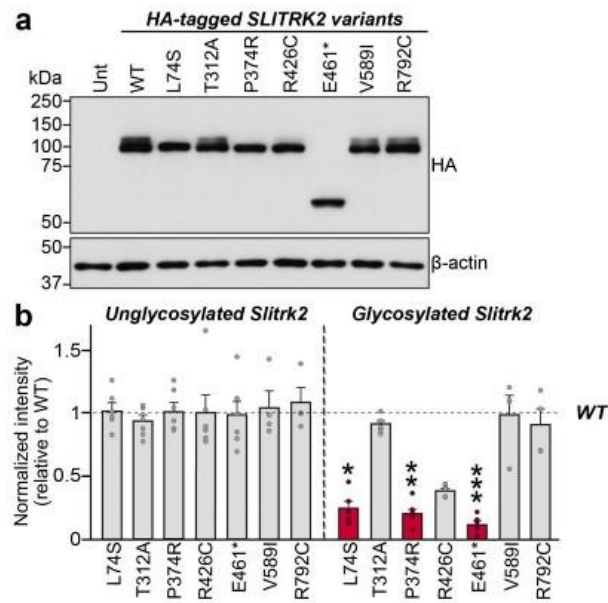
  

	V511M	E555D	V589I	R792C
Human	NNHFSHLVAVGVLDQL	ANSFVIINAVTCEPA	LS---DGTILSMNHNT	APSYE-SRQSQNQDRIN
Gorilla	NNHFSHLVAVGVLDQL	ANSFVIINAVTCEPA	LS---DGTILSMNHNT	APSYE-SRQSQNQDRIN
Macaque	NNHFSHLVAVGVLDQL	ANSFVIINAVTCEPA	LS---DGTILSMNHNT	APSYE-SRQSQNQDRIN
Bovine	NNHFSHLVAVGVLDQL	ANSFVIINAVTCEPA	LS---DGTILSMNHNT	TFSYE-SRQSQNQDRIN
Sheep	NNHFSHLVAVGVLDQL	ANSFVIINAVTCEPA	LS---DGTILSMNHNT	TFSYE-SRQSQNQDRIN
Mouse	NNHFSHLVAVGVLDQL	ANSFVIINAVTCEPA	LS---DGTILSMNHNT	APSYE-SRQSQNQDRIN
Rat	NNHFSHLVAVGVLDQL	ANSFVIINAVTCEPA	LS---DGTILSMNHNT	APSYE-SRQSQNQDRIN
Dog	NNHFSHLVAVGVLDQL	ANSFVIINAVTCEPA	LS---DGTILSMNHNT	APSYE-SRQSQNQDRIN
Zebrafish	NNHFSHLVAVGVLDQL	SSTSVVNVAVTCDSPS	VSVTLKPPVAVSSSTA	IFPYETSRQSQNQDRIN
Chicken	NNHFSHLVAVGVLDQL	SNPFPVNVAVTCEPST	LS---ESSLPMNQNT	APSYE-SRQSQNQDRIN

**Supplementary Figure 2. Alignment and conservation across different species of SLITRK2 residues mutated in human patients.**

**a** Alignment of human SLITRK1-6 amino acids surrounding SLITRK2 mutated residues found in human patients. The target mutated residues are indicated in yellow letters on a black background.

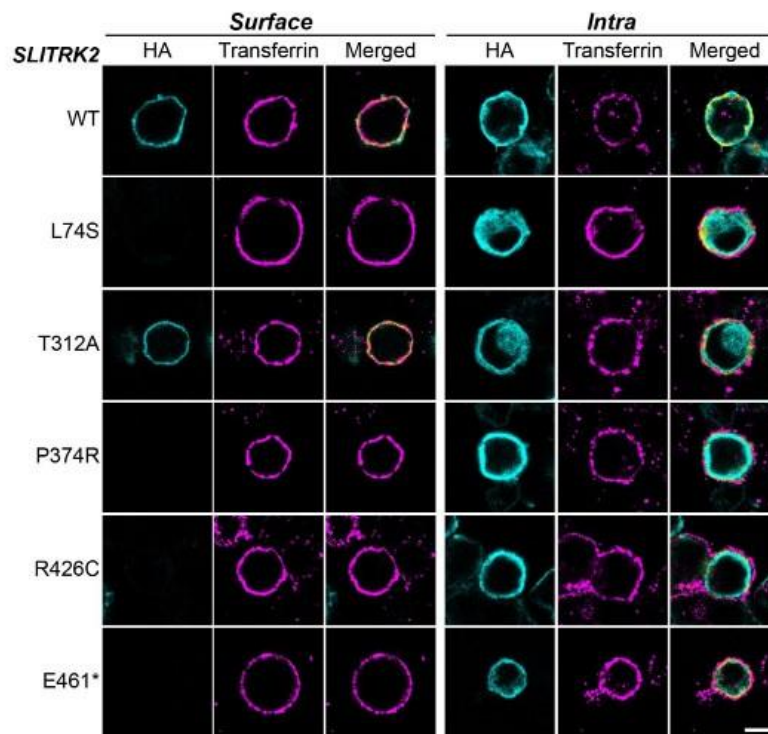
**b** Similarity or identity of mutated residues investigated in the current study across different species. The following GenBank accession numbers were utilized for sequence alignment: SLITRK1/human, NP\_001268432.1; SLITRK2/human, NP\_001137475.1; SLITRK3/human, NP\_001305739.1; SLITRK4/human, NP\_001171678.1; SLITRK5/human, NP\_001371538.1; SLITRK6/human, NP\_115605.2; Slitrk2/gorilla, XP\_030861981.1; Slitrk2/macaque, XP\_024649663.1; Slitrk2/bovine, XP\_005227636.1; Slitrk2/sheep, XP\_011962603.2; Slitrk2/mouse, XP\_006528079.1; Slitrk2/rat, XP\_006257688.1; Slitrk2/dog, XP\_038306961.1; Slitrk2/zebrafish, XP\_001332767.2; and Slitrk2/chicken, XP\_040526579.1.



**Supplementary Figure 3. Abnormally altered glycosylation of recombinant SLITRK2 mutant proteins.**

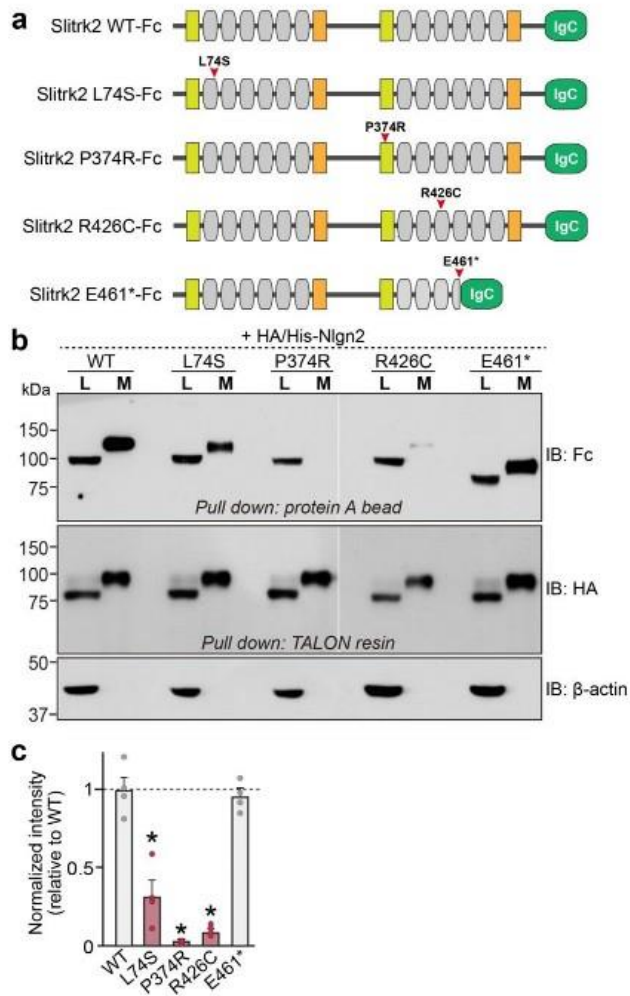
**a** Representative immunoblots from HEK293T cells transfected with the indicated WT or mutant forms of SLITRK2. Samples containing equal amounts of protein were resolved by SDS-PAGE and immunoblotted using anti-HA;  $\beta$ -actin was used for normalization. Molecular mass markers are labeled in kilodaltons.

**b** Quantification of unglycosylated and glycosylated species of SLITRK2 WT protein or the indicated mutant forms. Data are mean  $\pm$  SEMs ('n' denotes the number of independent experiments; WT, n = 10; L74S, n = 6; T312A, n = 6; P374R, n = 6; R426C, n = 6; E461\*, n = 6; V589I, n = 4; and R792C, n = 4; \* $p$  < 0.05, \*\* $p$  < 0.01, \*\*\* $p$  < 0.001; ANOVA with a non-parametric Kruskal-Wallis test). See **Source Data** for raw data values and **Supplementary Table 4** for statistical details.



**Supplementary Figure 4. Monitoring of surface expression of a subset of *SLITRK2* variants in HEK293T cells by immunofluorescence analysis.**

Surface expression analysis of HEK293T cells expressing the indicated WT or mutant forms of *SLITRK2*. Transfected HEK293T cells were immunostained with mouse anti-HA antibodies (cyan) and rabbit anti-transferrin antibodies (magenta) under non-permeabilized (surface) or permeabilized (intra) conditions. Scale bar, 10  $\mu\text{m}$  (applies to all images). The experiments were independently repeated three times.

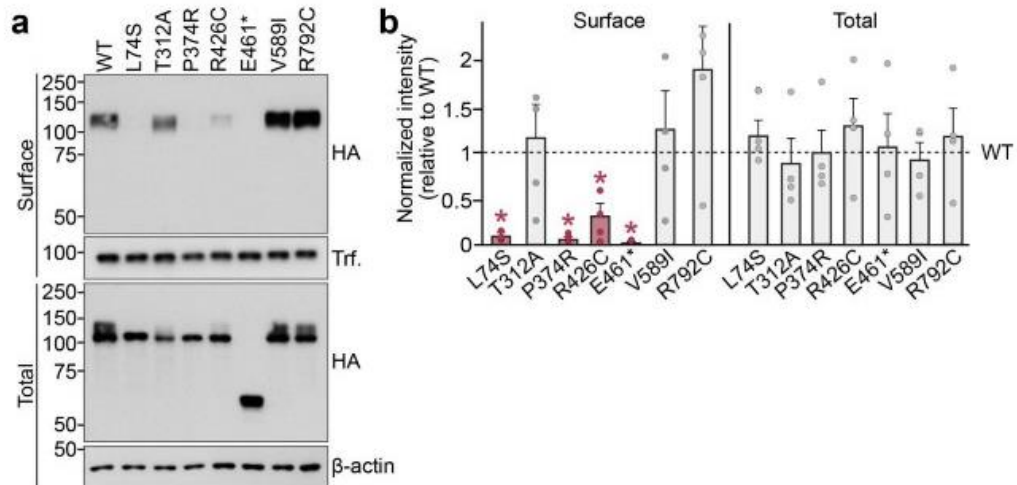


**Supplementary Figure 5. Impaired secretion of recombinant SLITRK2 mutant proteins.**

**a** Schematic diagrams depicting Ig-fused Slitrk2 WT and its variants.

**b** Representative images of western blots from lysates (L) or media (M) of HEK293T cells transfected with Ig-fused SLITRK2 and HA/His-tagged Nlgn2 vectors and immunoblotted with human Fc or HA antibodies.

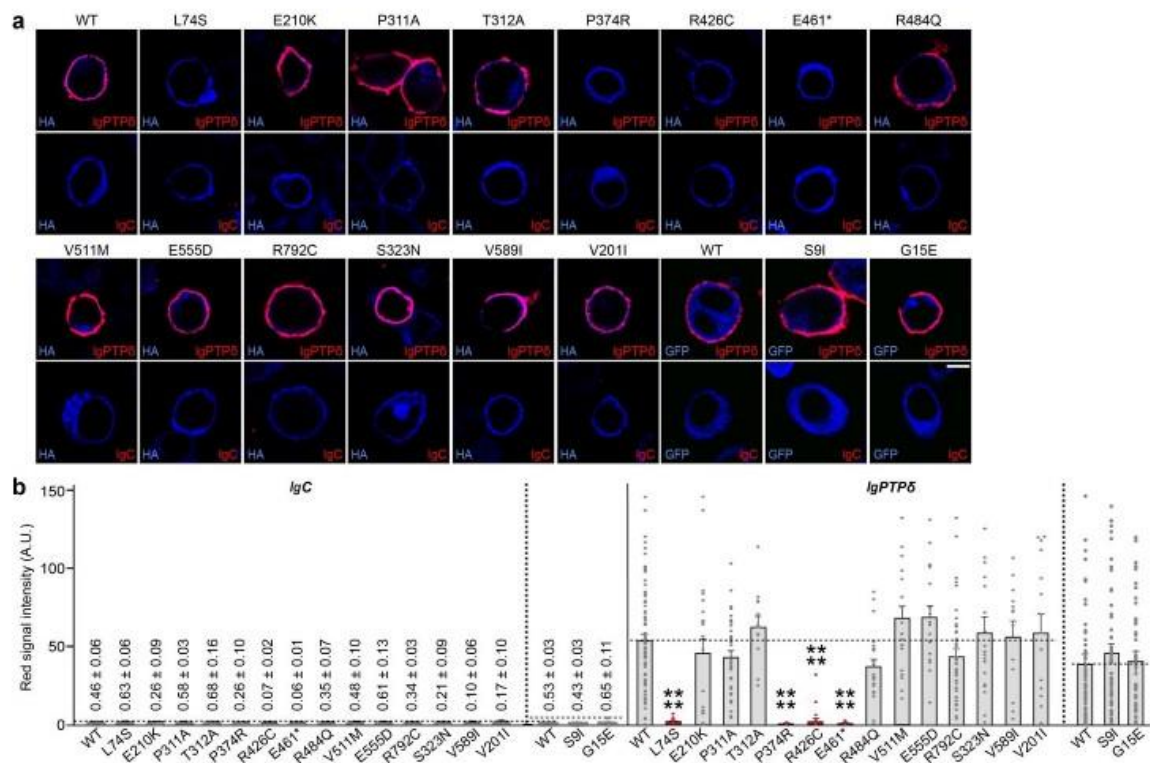
**c** Quantification of secreted levels of the indicated SLITRK2-Fc recombinant proteins in HEK293T cells. Data are mean  $\pm$  SEMs ('n' denotes the number of independent experiments; WT, n = 4; L74S, n = 4, P374R, n = 4; R426C, n = 4; and E461\*, n = 4; \* $p$  < 0.05; two-tailed Mann-Whitney U test). See **Source Data** for raw data values and **Supplementary Table 4** for statistical details.



**Supplementary Figure 6. Monitoring of surface expression of a subset of SLITRK2 variants in HEK293T cells by biotinylation analysis.**

**a** Surface exposure of SLITRK2 variants on transfected HEK293T cells, analyzed by immunoblotting of affinity-purified, surface-biotinylated SLITRK2 proteins. Biotinylated cell surface proteins and total lysate proteins were assessed by immunoblot with anti-HA antibodies.

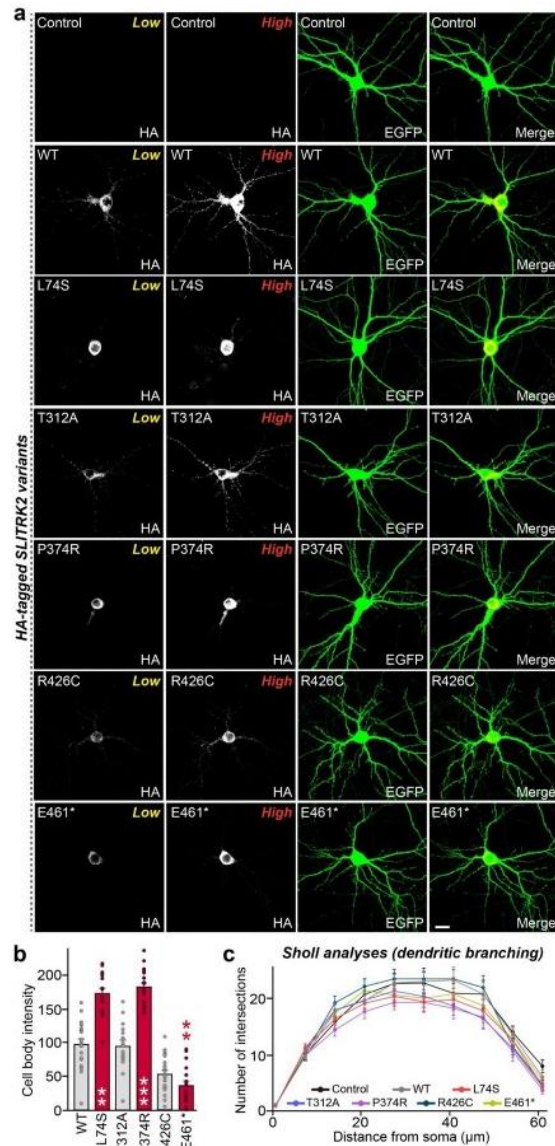
**b** Quantification of surface or total levels of SLITRK2 WT or the indicated mutant forms. Data are means  $\pm$  SEMs ('n' denotes the number of independent experiments; n = 4 for all experimental groups; \*p < 0.05; two-tailed Mann-Whitney U test). See **Source Data** for raw data values and **Supplementary Table 4** for statistical details.



**Supplementary Figure 7. Impaired PTPδ-binding of a subset of SLITRK2 variants.**

**a** Representative images of cell-surface-binding assays. HEK293T cells expressing the indicated WT or variant forms of SLITRK2 were incubated with 10 μg/ml control IgC or Ig-PTPδ, and then analyzed by immunofluorescence imaging of Ig-fusion proteins (red) and HA antibodies (blue; pseudocolor). Scale bar, 10 μm (applies to all images).

**b** Quantification of cell surface binding in **a**. Data are mean ± SEMs ('n' denotes the number of cells from three independent experiments; IgC: WT, n = 34; L74S, n = 21; E210K, n = 10; P311A, n = 15; T312A, n = 8; P374R, n = 13; R426C, n = 7; E461\*, n = 8; R484Q, n = 8; V511M, n = 9; E555D, n = 9; R792C, n = 14; S323N, n = 13; V589I, n = 11; V201I, n = 13; WT, n = 28; S9I, n = 29; and G15E, n = 30; Ig-PTPδ: WT, n = 74; L74S, n = 43; E210K, n = 22; P311A, n = 34; T312A, n = 12; P374R, n = 15; R426C, n = 20; E461\*, n = 17; R484Q, n = 24; V511M, n = 21; E555D, n = 20; R792C, n = 42; S323N, n = 19; V589I, n = 13; V201I, n = 15; WT, n = 51; S9I, n = 54; and G15E, n = 47; \*\*\*\**p* < 0.0001; ANOVA with non-parametric Kruskal-Wallis test). See **Source Data** for raw data values and **Supplementary Table 4** for statistical details.

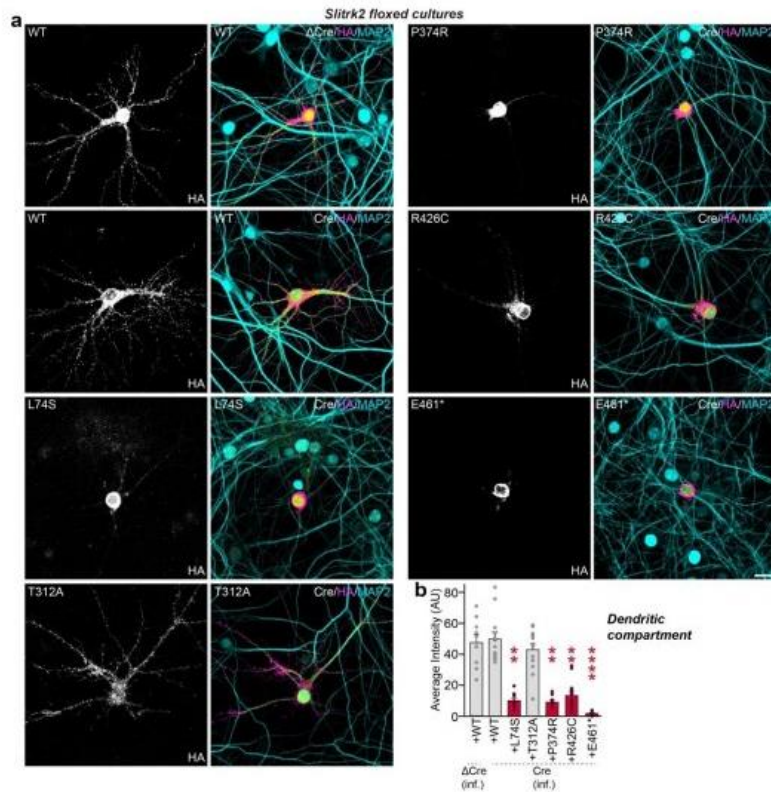


**Supplementary Figure 8. Impaired dendritic targeting of a subset of SLITRK2 variants in cultured hippocampal neurons.**

**a** Representative images of hippocampal neurons transfected at DIV10 with the indicated WT or mutant forms of SLITRK2 together with EGFP vector. Transfected neurons were double-immunostained at DIV14 with antibodies against HA (gray) and EGFP (green). Images were captured by confocal microscopy using different exposure times (low or high). Scale bar, 20 μm (applies to all images).

**b** Quantification of the average intensity of HA immunoreactivity in the soma of transfected neurons using images acquired at low exposure times. Data are shown as means ± SEMs ('n' denotes the number of neurons from three independent experiments; WT, n = 24; L74S, n = 19; T312A, n = 18; P374R, n = 22; R426C, n = 28; and E461\*, n = 22; \*\*p < 0.01, \*\*\*p < 0.001; ANOVA with non-parametric Kruskal-Wallis test).

**c** Quantification of dendritic branching in cultured neurons by Sholl analysis demonstrating no alterations in neurons expressing the indicated SLITRK2 variants. Data are shown as means ± SEM ('n' denotes the number of neurons from three independent experiments; Control, n = 14; WT, n = 24; L74S, n = 25; T312A, n = 18; P374R, n = 25; R426C, n = 21; and E461\*, n = 22). See **Source Data** for raw data values and **Supplementary Table 4** for statistical details.

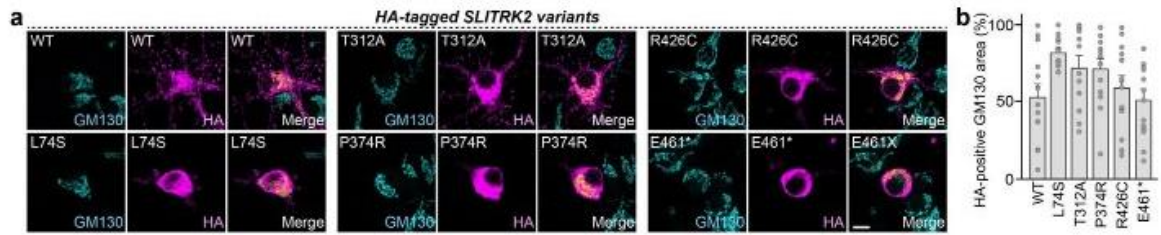


**Supplementary Figure 9. Impaired dendritic targeting of a subset of SLITRK2 variants in *Slitrk2*-cKO hippocampal neurons.**

**a** Representative images of cultured hippocampal neurons (DIV14) from *Slitrk2*<sup>fl/fl</sup> mice infected with lentiviruses expressing  $\Delta$ Cre or Cre at DIV5 and transfected with vectors expressing SLITRK2 WT or the indicated mutant forms at DIV8–9. Neurons were immunostained using antibodies against HA (magenta), and MAP2 (cyan [non-nuclear fluorescence signal]). Scale bar, 20  $\mu$ m (applies to all images).

**b** Dendritic targeting of SLITRK2 WT or the indicated mutant forms in *Slitrk2*-cKO hippocampal cultured neurons was quantified by measuring average intensity of HA immunofluorescence in primary dendrites. Data are shown as means  $\pm$  SEMs ('n' denotes the number of neurons from three independent batches;  $\Delta$ Cre+WT, n = 10; Cre+WT, n = 13; Cre+L74S, n = 12, Cre+T312A, n = 13, Cre+P374R, n = 10; Cre+R426C, n = 16; and Cre+E461\*, n = 11; \*\* $p$  < 0.01, \*\*\*\* $p$  < 0.0001; ANOVA with non-parametric Kruskal-Wallis test). See **Source Data** for raw data values and **Supplementary Table 4** for statistical details.

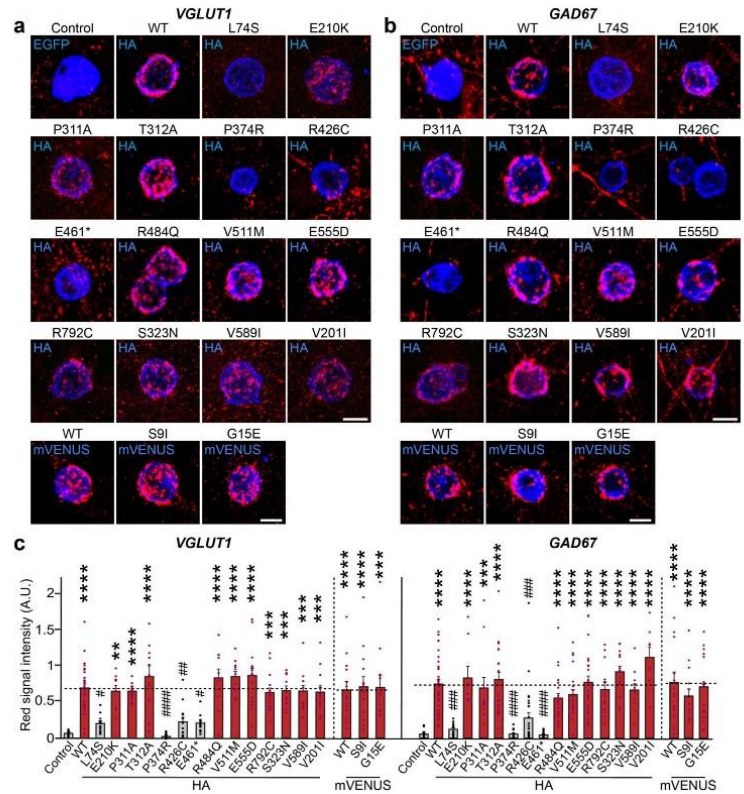




**Supplementary Figure 10. Analysis of the distribution of a subset of recombinant SLITRK2 variants in the *cis*-Golgi compartment of cultured neurons.**

**a** Representative images of hippocampal cultured neurons transfected with SLITRK2 WT or the indicated mutant forms at DIV10. Transfected neurons were double-immunostained with antibodies against the *cis*-Golgi marker GM130 (cyan) and HA (magenta) at DIV14. Scale bar, 20  $\mu$ m (applies to all images).

**b** Quantification of HA immunoreactive-positive GM130 puncta area. Data are means  $\pm$  SEMs ('n' denotes the number of neurons from three independent experiments; WT, n = 14; L74S, n = 16, T312A, n = 12; P374R, n = 14; R426C, n = 15; and E461\*, n = 14). See **Source Data** for raw data values and **Supplementary Table 4** for statistical details.



**Supplementary Figure 11. Impaired synaptogenic activity of a subset of SLITRK2 variants in cultured neurons.**

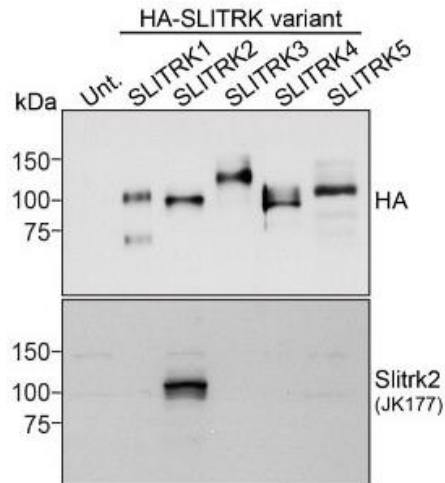
**a, b** Representative images of the heterologous synapse-formation activities of WT SLITRK2 and the indicated point mutants in cultured hippocampal neurons. Neurons were cocultured with HEK293T cells transfected with the indicated WT or variant forms of SLITRK2 at DIV10. Neurons were then immunostained at DIV12 with antibodies against EGFP or HA (blue) and VGLUT1 (**a**; red) or GAD67 (**b**; red). Scale bar, 10  $\mu$ m (applies to all images).

**c** Synapse-formation activity was quantified by measuring the ratio of VGLUT1 or GAD67 staining intensity (red) to HA/EGFP intensity (blue). All data are shown as means  $\pm$  SEMs ('n' denotes the number of cells from three independent batches; VGLUT1: Control, n = 17; WT, n = 28; L74S, n = 13; E210K, n = 10; P311A, n = 15; T312A, n = 11;

P374R, n = 15; R426C, n = 15; E461\*, n = 12; R484Q, n = 11; V511M, n = 11; E555D, n = 15; R792C, n = 16; S323N, n = 13; V589I, n = 12; V201I, n = 15; WT, n = 14; S9I, n = 14; and G15E, n = 12; GAD67: Control, n = 29; WT, n = 36; L74S, n = 16;

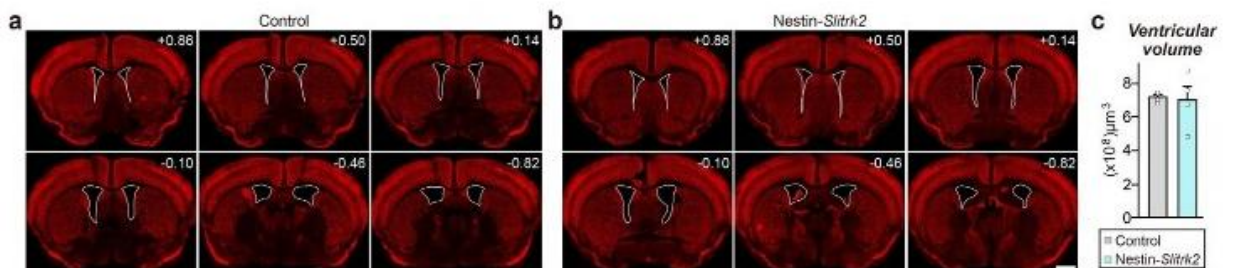
E210K, n = 8; P311A, n = 12; T312A, n = 21; P374R, n = 15; R426C, n = 27; E461\*, n = 21; R484Q, n = 22; V511M, n = 21; E555D, n = 23; R792C, n = 13; S323N, n = 14; V589I, n = 11; V201I, n = 9; WT, n = 16; S9I, n = 13; and G15E, n = 18; \*\* $p$  < 0.01, \*\*\* $p$  < 0.001, \*\*\*\* $p$  < 0.0001; Control vs. experimental group; # $p$  < 0.05, ## $p$  < 0.01, ### $p$  < 0.001, #### $p$  < 0.0001 WT vs.

experimental group; ANOVA with a non-parametric Kruskal-Wallis test). See **Source Data** for raw data values and **Supplementary Table 4** for statistical details.



**Supplementary Figure 12. Cross-reactivity of the in-house SLITRK2 antibody.**

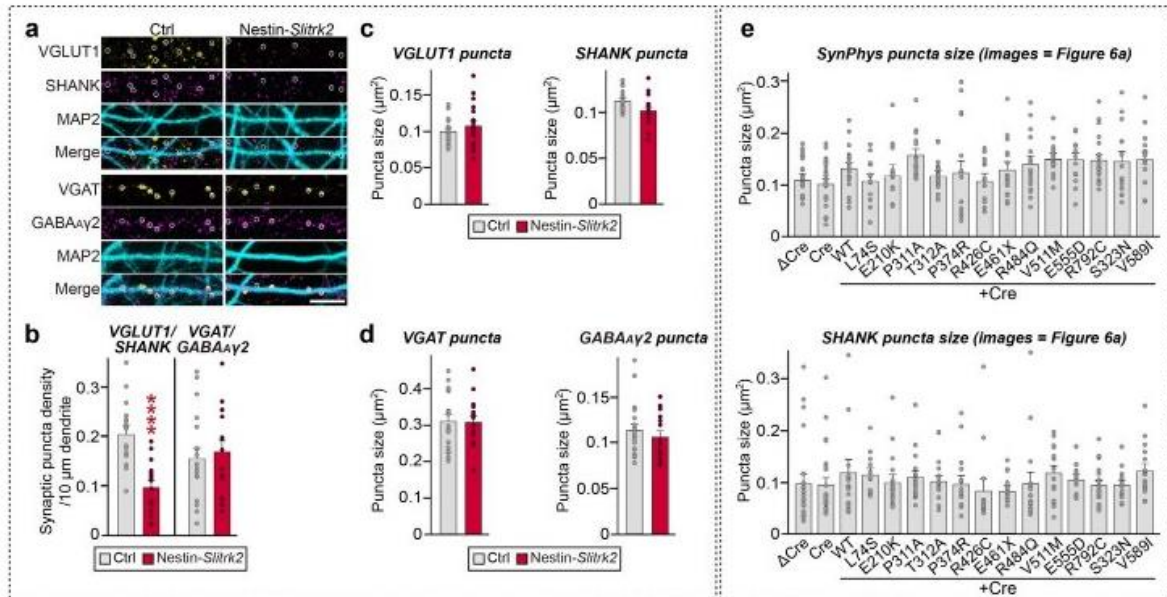
Lysates of untransfected (Unt.) or transfected HEK293T cells expressing the indicated SLITRK expression plasmids were immunoblotted with the anti-Slitrk2 antibody JK177. Expression of HA-tagged SLITRK plasmids was verified by immunoblotting with anti-HA antibodies. The experiments were independently repeated three times.



**Supplementary Figure 13. Ventricular volumes are comparable between control and *Slitrk2*-cKO mice.**

**a, b** Representative images of NeuN-stained coronal sections of Control (**a**) and Nestin-Slitrk2 (**b**) mice. Scale bar, 1 mm (applies to all images).

**c** Summary data showing quantification of ventricular volume. Data are means  $\pm$  SEMs (Ctrl, n = 4 mice; Nestin-Slitrk2, n = 4 mice). See **Source Data** for raw data values and **Supplementary Table 4** for statistical details.



**Supplementary Figure 14. Impaired excitatory synapse maintenance in cultured hippocampal neurons from Nestin-Slitrk2 mice.**

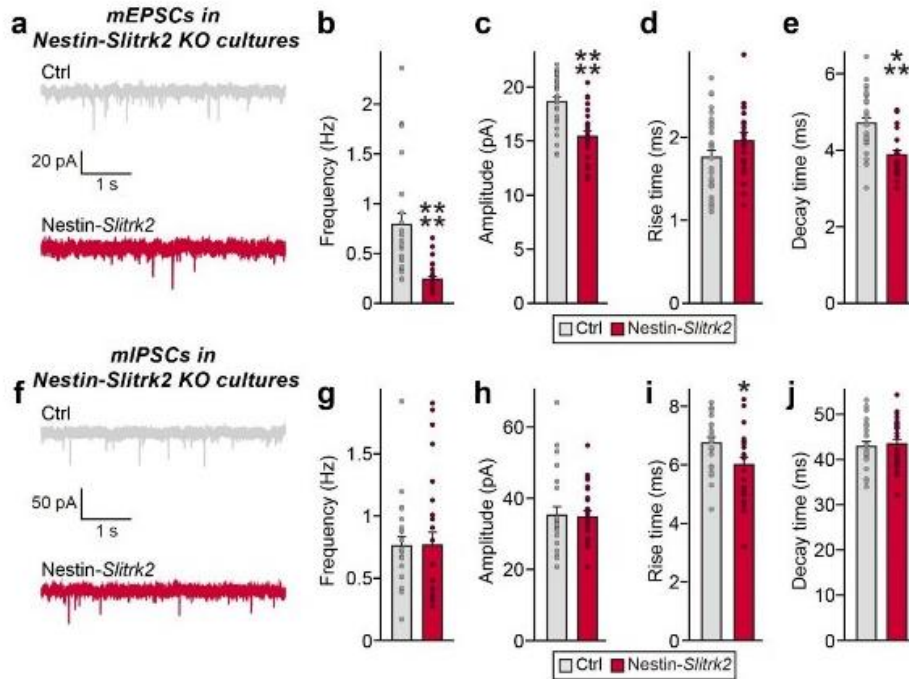
**a** Representative images from Nestin-Slitrk2 cultured hippocampal neurons using antibodies against MAP2 (cyan), VGLUT1 (yellow), SHANK (magenta), VGAT (yellow) and GABA<sub>A</sub>γ2 (magenta) at DIV14. Scale bar, 10 μm (applies to all images).

**b** Summary data showing quantification of excitatory and inhibitory synaptic puncta density. Data are means ± SEMs ('n' denotes the number of neurons from three independent batches; VGLUT1/SHANK/Ctrl, n = 16; VGLUT1/SHANK/Nestin-Slitrk2; n = 19; VGAT/GABA<sub>A</sub>γ2/Ctrl, n = 19; and VGAT/GABA<sub>A</sub>γ2/Nestin-Slitrk2, n = 15; \*\*\*\*p < 0.0001, two-tailed Mann-Whitney U test).

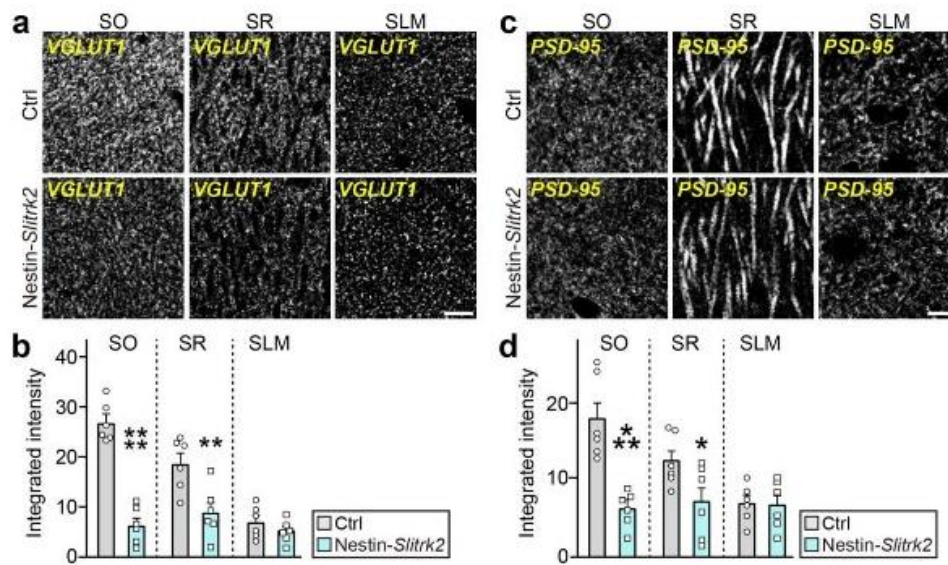
**c, d** Summary data showing quantification of excitatory (**c**) and inhibitory (**d**) synaptic puncta size. Data are means ± SEMs ('n' denotes the number of neurons from three independent batches; VGLUT1/SHANK/Ctrl, n = 16; VGLUT1/SHANK/Nestin-Slitrk2; n = 19; VGAT/GABA<sub>A</sub>γ2/Ctrl, n = 19; and VGAT/GABA<sub>A</sub>γ2/Nestin-Slitrk2, n = 15; two-tailed Mann-Whitney U test).

**e** Summary data showing quantification of excitatory synaptic puncta size in cultured Slitrk2-floxed hippocampal neurons transfected with the indicated SLITRK2 variants. Data are means ± SEMs. See **Figure 6** for representative images and quantification of excitatory synaptic puncta density. See **Source Data** for raw data values and

**Supplementary Table 4** for statistical details.



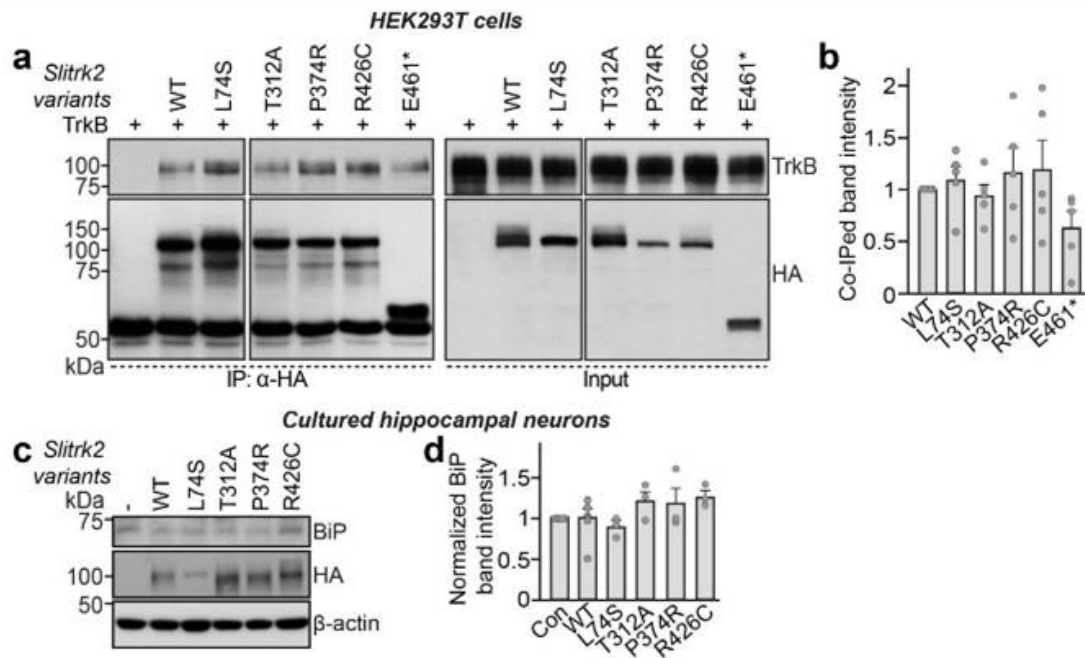
**Supplementary Figure 15. Impaired excitatory synaptic transmission in cultured hippocampal neurons from Nestin-Slitrk2 mice.** **a–e** Representative mEPSC traces (**a**) and quantification of frequency (**b**), amplitudes (**c**), rise time (**d**), and decay time (**e**) of mEPSCs recorded in cultured hippocampal neurons from Nestin-Slitrk2 mice. Data are shown as means  $\pm$  SEMs ('n' denotes the number of neurons from three independent experiments; Control, n = 26; Nestin-Slitrk2, n = 24; \*\*\* $p$  < 0.001, \*\*\*\* $p$  < 0.0001; two-tailed Mann-Whitney  $U$  test). **f–j** Representative mIPSC traces (**f**) and quantification of frequency (**g**), amplitude (**h**), rise time (**i**), and decay time (**j**) of mIPSCs recorded in cultured hippocampal neurons from Nestin-Slitrk2 mice. Data are shown as means  $\pm$  SEMs ('n' denotes the number of neurons from three independent experiments; Control, n = 23; Nestin-Slitrk2, n = 25; \* $p$  < 0.05; two-tailed Mann-Whitney  $U$  test). See **Source Data** for raw data values and **Supplementary Table 4** for statistical details.



**Supplementary Figure 16. Impaired excitatory synapse development in the hippocampal CA1 region of Nestin-Slitrk2 mice.**

**a, c** Representative images of hippocampal CA1 layers (SO, SR and SLM) in control and Nestin-Slitrk2 mice, analyzed by assessing the excitatory synaptic markers VGLUT1 (**a**) or PSD-95 (**c**). Scale bar, 20  $\mu$ m (applies to all images).

**b, d** Quantification of the integrated intensity of VGLUT1-positive (**b**) and PSD-95-positive (**d**) synaptic puncta. Data are shown as means  $\pm$  SEMs (n = 6 mice; \* $p$  < 0.05, \*\* $p$  < 0.01, \*\*\* $p$  < 0.001, \*\*\*\* $p$  < 0.0001; two-tailed unpaired t test). See **Source Data** for raw data values and **Supplementary Table 4** for statistical details.



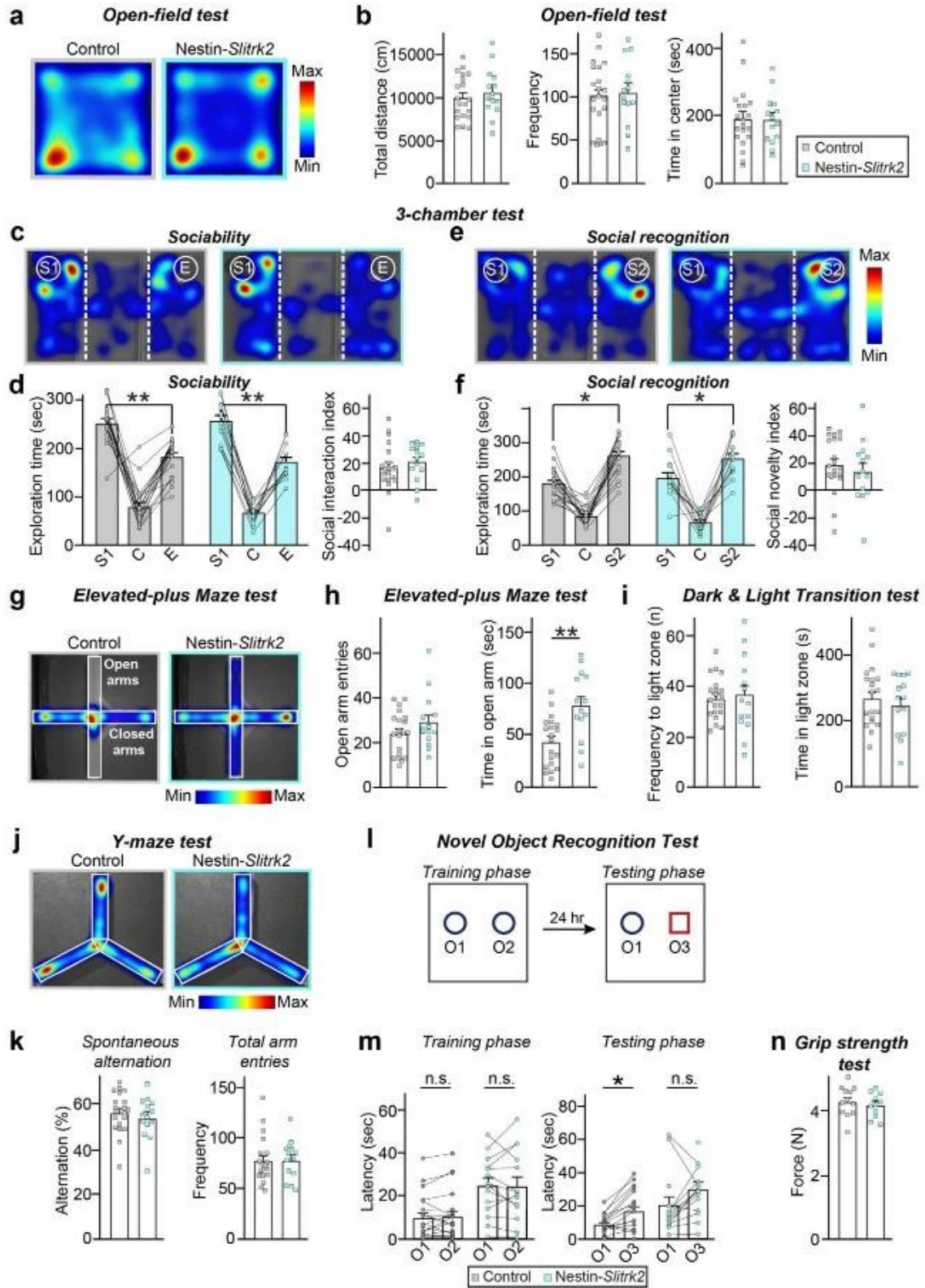
**Supplementary Figure 17. No changes in TrkB binding and unfolded protein responses in cultured neurons expressing disease-associated SLITRK2 variants.**

**a** Representative immunoblots showing co-immunoprecipitation of TrkB with SLITRK2 WT or the indicated SLITRK2 variants. HEK293T cells were transfected with HA-tagged SLITRK2 WT or its variants alone or together with untagged TrkB, after which coimmunoprecipitation of TrkB2 with SLITRK2 was assayed. Input, 5%.

**b** Quantification of coimmunoprecipitated TrkB in **a**, normalized to WT. Data are means  $\pm$  SEMs ( $n = 5$  independent experiments).

**c** Cultured cortical neurons were infected with lentiviruses expressing SLITRK2 WT or the indicated SLITRK2 variants. Levels of BiP were analyzed by semi-quantitative immunoblotting.

**d** Quantification of BiP levels from **c**. Data are means  $\pm$  SEMs (WT,  $n = 6$ ; L74S,  $n = 6$ ; T312A,  $n = 3$ ; P374R,  $n = 3$ ; R426C,  $n = 3$ ; and E461\*,  $n =$  independent experiments). See **Source Data** for raw data values and **Supplementary Table 4** for statistical details.





**Supplementary Figure 18. *Slitrk2*-cKO mice display anxiolytic behavior, with normal locomotion and social recognition.**

**a, b** Analysis of locomotor activity by open field test in Control and Nestin-*Slitrk2* mice. **(a)** Representative heat maps of the time spent in the open chamber. **(b)** Number of entries into the center zone, time spent in the center zone, and total distance moved. Data are presented as means  $\pm$  SEMs ('n' denotes the number of mice; Control, n = 20; Nestin-*Slitrk2*, n = 14).

**c, d** Analysis of sociability in Control and Nestin-*Slitrk2* mice by 3-chamber test. **(c)** Representative heat maps of the time spent in each side chamber containing a novel mouse (S1) or an empty wire cup (E), or time spent in the center chamber. **(d)** Summary graphs showing time in chamber and preference index. The preference index obtained from exploration time represents the numerical difference between the time spent exploring or sniffing the two targets (S1/stranger vs. E/empty wire cup) divided by total exploration time  $\times$  100. Data are presented as means  $\pm$  SEMs ('n' denotes the number of mice; Control, n = 20; Nestin-*Slitrk2*, n = 14; <sup>\*\*\*</sup> $p < 0.001$ ; Tukey's multiple comparisons and Mann Whitney U test).

**e, f** Analysis of social novelty recognition in Control and Nestin-*Slitrk2* mice by 3-chamber test. **(e)** Representative heat maps of the time spent in each side chamber containing the previous mouse (S1) or a novel mouse (S2), or time spent in the center chamber. **(f)** Summary graphs showing time in chamber and preference index. The preference index obtained from exploration time represents the numerical difference between the time spent exploring or sniffing the two targets (S2/new stranger vs. S1/previous stranger) divided by total exploration time  $\times$  100. Data are presented as means  $\pm$  SEMs ('n' denotes the number of mice; Control, n = 20; Nestin-*Slitrk2*, n = 14; <sup>\*</sup> $p < 0.05$ ; Tukey's multiple comparisons and Mann Whitney U test).

**g, h** Analysis of anxiety/exploration-related behavior by elevated plus maze (EPM) test in Control and Nestin-*Slitrk2* mice. **(g)** Representative heat maps of the time spent in the open arms of the EPM. Red represents increased time spent, and blue represents minimal time spent during the test. **(h)** Number of open arm entries and time spent in open arms. Data are presented as means  $\pm$  SEMs ('n' denotes the number of mice; Control, n = 20; Nestin-*Slitrk2*, n = 14; <sup>\*\*</sup> $p < 0.01$ ; two-tailed Mann Whitney U test).

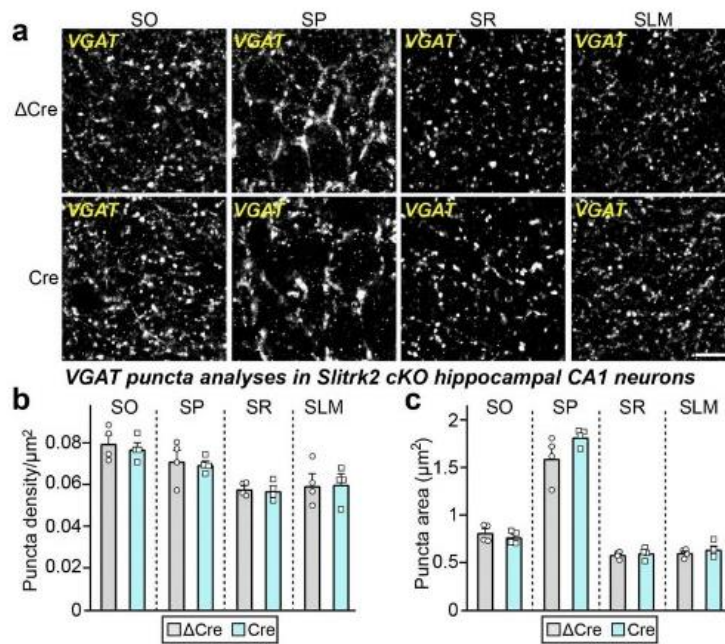
**i** Analysis of anxiety/exploration-related behavior by light-dark test (LDT) in Control and Nestin-*Slitrk2* mice. Number of entries into the light chamber and time spent in the light chamber are shown. Data are presented as means  $\pm$  SEMs ('n' denotes the number of mice; Control, n = 20; Nestin-*Slitrk2*, n = 14).

**j, k** Analysis of spatial working memory in Control and Nestin-*Slitrk2* mice using a Y-maze test. Representative heat map showing movements of the indicated mice during the Y-maze test (**j**). Red represents increased time spent, and blue represents minimal time spent during the test. (**k**) Total number of arm entries and spontaneous alternation performance ratio (% of spontaneous alternations). Data are presented as means  $\pm$  SEMs ('n' denotes the number of mice; Control, n = 20; Nestin-*Slitrk2*, n = 14).

**l** Schematic depiction of novel object-recognition memory test.

**m** Analysis of novel object recognition (NOR) memory in Control and Nestin-*Slitrk2* mice. Exploration time and novel object preference for 24-h testing phases. Mice were allowed to explore two identical objects (denoted O1 and O2), and after a 24-h delay were exposed to two different objects—one familiar object from the training phase (O1) and one novel object (O3). Data are presented as means  $\pm$  SEMs ('n' denotes the number of mice; Control, n = 12; Nestin-*Slitrk2*, n = 11; \* $p < 0.05$ ; two-tailed paired t test).

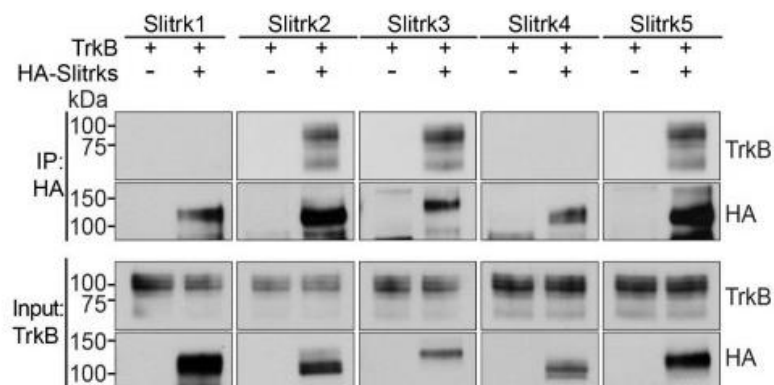
**n** Test of grip strength in Nestin-*Slitrk2* and control mice. Data are presented as means  $\pm$  SEMs ('n' denotes the number of mice; Control, n = 12; Nestin-*Slitrk2*, n = 11). See **Source Data** for raw data values and **Supplementary Table 4** for statistical details.



**Supplementary Figure 19. No alterations in inhibitory synapse maintenance in hippocampal CA1-specific *Slitr2*-cKO mice.**

**a** Representative images showing VGAT puncta in hippocampal CA1 layers (SO, SP, SR and SLM) of *Slitr2*<sup>fl/fl</sup> mice injected with AAVs expressing Cre or  $\Delta$ Cre. Scale bar, 20  $\mu\text{m}$  (applies to all images).

**b, c** Quantification of VGAT puncta density (**b**) or size (**c**). Data are shown as means  $\pm$  SEMs ( $n = 4$  mice each after averaging data from 3 sections/mouse). See **Source Data** for raw data values and **Supplementary Table 4** for statistical details.



**Supplementary Figure 20. TrkB interacts with Slitrk2, Slitrk3 and Slitrk5 in heterologous cells, but not with Slitrk1 or Slitrk4.**

HEK293T cells were singly transfected with the indicated HA-tagged Slitrk expression plasmids or co-transfected with a TrkB expression plasmid, after which coimmunoprecipitation of Slitrks with TrkB was assayed. Input, 5%. The experiments were independently repeated three times.

## Supplementary Tables

**Table S1. *SLITRK2* variants identified in the current study**

Variants are based on transcript NM\_032539.4. PPH2, polyphen2; CADD, Combined Annotation Dependent Depletion; M, male; F, female; ML/DT, membrane localization (Fig. 3) and dendritic targeting (Fig. 4); I-PTP $\delta$ , interaction with PTP $\delta$  (Fig. S4); SF: synapse formation (Fig. S5); Res-Syn, impaired rescue of synaptic maintenance and activity (Fig. 6); NA, non-available; ACMG/AMP criteria, criteria for variant interpretation according to the American College of Medical Genetics and Genomics/Association for Molecular Pathology; Criteria supporting pathogenicity: PM2, absent from controls (in gnomAD); PS2, *de novo* in a patient with the disease and no family history; PS3, well-established *in vitro* or *in vivo* functional studies supportive of a damaging effect on the gene or gene product; PP1, cosegregation with disease in multiple affected family members in a gene definitively known to cause the disease; PP3, multiple lines of computational evidence support a deleterious effect on the gene or gene product; Criteria supporting a benign effect: BS2, observed in a healthy adult individual with full penetrance expected at an early age; BP4, multiple lines of computational evidence suggest no impact on the gene or gene product. a, absent from an unaffected maternal uncle; b, amino acid change not reported in gnomAD, whereas p.R426H was observed in a hemizygous state in one male; c, this individual carried a pathogenic variant in *CUL4B* which seemed to be sufficient to explain the majority of his clinical phenotypes.

**Abbreviations:** dam, damaging; Delet., deleterious; Extracell., extracellular; N.A., not applicable; Poss., possibly; Prob., probably; and Toler., tolerable.

**Table S1. SLITRK2 variants identified in the current study**

Variant nomenclature (NM_032539.4)	Individual with NDD			Frequency		In silico predictions			Functional effects					Variant interpretation (ACMG/AMP)	
	Ind.	Sex	Inheritance	M	F	PPH2	SIFT	CADD	Domain	Affects ML/DT	Affects I-PTP6	Affects SF	Affects Res-Syn	Criteria	Class
<b>Potential disease-causing variants in individuals with NDD:</b>															
c.1381G>T, p.E461*	P1	M	de novo	-	-	N.A.	N.A.	35	N.A.	yes	yes	yes	yes	PM2, PS2, PS3, PP3	4_lik_patho
c.221T>C, p.L74S	P7	M	maternal	-	-	Prob. dam.	Delet.	26	Extracell.	yes	yes	yes	yes	PM2, PS3, PP1	4_lik_patho
c.934A>G, p.T312A	P3	M	maternal	-	1	Prob. dam.	Delet.	24.8	Extracell.	no	no	no	yes	PM2, PS3, PP3	4_lik_patho
c.1121C>G, p.P374R	P8	M	de novo	-	-	Prob. dam.	Delet.	26	Extracell.	yes	yes	yes	yes	PM2, PS3, PP1, PP3	4_lik_patho
c.1276C>T, p.R426C	P4	F	de novo	-	-	Prob. dam.	Delet.	28.2	Extracell.	yes	yes	yes	yes	PM2, PS2, PS3, PP3	4_lik_patho
c.628G>A, p.E210K	P1	M	maternal	-	-	Prob. dam.	Delet.	27.2	Extracell.	no	no	no	no	PM2, PP3	3_VUS
c.1531G>A, p.V511M	P5	M	maternal	-	2	Prob. dam.	Delet.	25.2	Extracell.	no	no	no	no	PP3	3_VUS
<b>Variants reported in Decipher:</b>															
c.26G>T, p.S9I	Dc	M	maternal	-	-	Poss. dam.	Delet.	21.7	Signal	no	no	no	N.A.	PM2, PP3	3_VUS
c.44G>A, p.G15E	Dc	M	maternal	-	-	Poss. dam.	Delet.	22.3	Signal	no	no	no	N.A.	PM2, PP3	3_VUS
<b>Other variants identified in individuals with NDD:</b>															
c.601G>A, p.V201I	P1	M	maternal	19	33	Poss. dam.	Delet.	22.8	Extracell.	no	no	no	N.A.	BS2, PP3	2_lik_benign
c.931C>G, p.P311A	P1	M	NA	1	-	Prob. dam.	Delet.	24.6	Extracell.	no	no	no	no	PP3	3_VUS
c.1451G>A, p.R484Q	P1	M	maternal	1	2	Poss. dam.	Delet.	22.7	Extracell.	no	no	no	no	PP3	3_VUS
c.1665G>C, p.E555D	P1	M	NA	-	-	Benign	Delet.	15.6	Extracell.	no	no	no	no	PM2	2_lik_benign
c.2374C>T, p.R792C	P9	M	maternal	2	3	Prob. dam.	Delet.	29.9	Intracell.	no	no	no	no	PP3	3_VUS
<b>Other missense variants from GnomAD included in the study:</b>															
c.968G>A, S323N	-	-	-	6	6	Benign	Toler.	16.4	Extracell.	no	no	no	no	BS2, BP4	2_lik_benign
c.1765G>A, V589I	-	-	-	139	128	Benign	Toler.	0.102	Extracell.	no	no	no	no	BS2, BP4	2_lik_benign

**Table S2. Stability analysis of LRR1 and LRR2 domain of human SLITRK2 WT and mutant**

The free energy of folding (stability) was calculated using FoldX 5.0.

<b>Model</b>	<b>Energy (kcal.mol<sup>-1</sup>)</b>
LRR1 (C33-D270) Wild-type	68.49
LRR1 (C33-D270) L74S	75.80
LRR1 (C33-D270) V201I	67.55
LRR1 (C33-D270) E210K	69.77
LRR2 (P341-P579) Wild-type	51.77
LRR2 (P341-P579) P374R	87.92
LRR2 (P341-P579) R426C	52.50
LRR2 (P341-P579) R484Q	52.26
LRR2 (P341-P579) V511M	58.16
LRR2 (P341-P579) E555D	52.24

**Table S3. List of all primary and secondary antibodies used at their corresponding dilutions**

Antibody	Species	Vendor	Catalog #	Dilution	RRID
SLITRK2	Rabbit	Dr. Jaewon Ko's laboratory	JK177	WB-1:1000	AB_2892626
GM130	Mouse	BD Transduction Laboratories	610822 (clone 35/GM130)	ICC-1:100	AB_398141
TfR	Rabbit	Abcam	ab214039 (clone EPR20584)	ICC-1:100	AB_2904534
VGLUT1	Guinea pig	Millipore	AB5905	ICC-1:200	AB_2301751
GAD67	Mouse	Millipore	MAB5406 (clone 1G10.2)	IHC-1:500 ICC-1:100 WB-1:1000	AB_2278725
Synaptophysin	Mouse	Sigma-Aldrich	S5768 (clone SVP-38)	ICC-1:1000	AB_477523
PSD-95	Mouse	NeuroMab	75-028 (clone K28/43)	WB-1:1000	AB_2877189
GABA <sub>A</sub> γ2	Rabbit	Synaptic Systems	224 003	ICC-1:500	AB_2263066
VGAT	Guinea pig	Synaptic Systems	131 004	IHC-1:500, ICC-1:200	AB_887873
HA	Mouse	BioLegend	901501 (clone 16B12)	ICC-1:500, WB-1:1000	AB_2565006
MAP2	Rabbit	Abcam	ab32454	ICC-1:500	AB_776174
MAP2	Mouse	Sigma-Aldrich	M1406 (clone AP-20)	ICC-1:300	AB_477171
TrkB	Rabbit	Cell Signaling	4603 (clone 80E3)	WB-1:1000	AB_2155125
phospho-TrkB	Rabbit	Thermo Fisher	PA5-36695	WB-1:1000	AB_2553666
BiP	Rabbit	Cell Signaling	3177 (clone C50B12)	WB-1:500	AB_2119845
GFP	Goat	Rockland	600-101-215	ICC-1:500	AB_218182
GluN1	Mouse	Millipore	MAB363 (clone 54.1)	WB-1:1000	AB_94946
NR2A	rabbit	Millipore	07-632	WB-1:1000	AB_310837
β-actin	Mouse	Santa Cruz	sc-47778 (clone C4)	WB-1:1000	AB_2714189
NeuN	Mouse	Millipore;	MAB377 (clone A60)	WB-1:1000	AB_2298772
VGLUT1	Rabbit	Dr. Jaewon Ko's laboratory	JK111	IHC-1:500, ICC-1:300	AB_2810945
PSD-95	Rabbit	Dr. Jaewon Ko's laboratory	JK016	IHC-1:500	AB_2722693
Shank	Rabbit	Dr. Eunjoon Kim's laboratory	1172	ICC-1:200	AB_2810261
GluA1	Rabbit	Dr. Eunjoon Kim's laboratory	1193	WB-1:1000	AB_2722772
GluA2	Rabbit	Dr. Eunjoon Kim's laboratory	1195	WB-1:1000	AB_2722773
Cy3-AffiniPure Donkey Anti-Rabbit IgG antibodies	Rabbit	Jackson ImmunoResearch	711-165-152	IHC-1:500 ICC-1:500	AB_2722772
Cy3-AffiniPure Donkey Anti-Mouse IgG antibodies	Mouse	Jackson ImmunoResearch	715-165-150	IHC-1:500 ICC-1:500	AB_2340813
Cy3-Donkey Anti-Human IgG antibodies	Human	Jackson ImmunoResearch	709-165-149	ICC-1:500	AB_2340535
Cy3-Donkey Anti-Guinea Pig IgG antibodies	Guinea Pig	Jackson ImmunoResearch	706-035-148	ICC-1:500	AB_2340447
FITC-AffiniPure Donkey Anti-Mouse	Mouse	Jackson ImmunoResearch	715-095-150	ICC-1:150	AB_2340792
IgG antibodies					30
FITC-AffiniPure Donkey Anti-Goat IgG antibodies	Goat	Jackson ImmunoResearch	705-095-147	ICC-1:150	AB_2340401
FITC-AffiniPure Donkey Anti-Rabbit IgG antibodies	Rabbit	Jackson ImmunoResearch	711-095-152	ICC-1:150	AB_2315776
Goat Anti-Guinea Pig IgG antibodies	Guinea Pig	Thermo Fisher	A-21450	ICC-1:100	AB_141882
HRP-goat anti-human IgG antibody	Human	Thermo Fisher	62-8420	WB-1:10000	AB_2533962

**Abbreviations:** ICC, immunocytochemistry; IHC, immunohistochemistry; WB, western blotting

## Supplementary Notes

### Further details of participants and molecular analyses

#### Individuals P1 and P2

##### *Clinical data*

Individual P1, a 31-year-old Caucasian male, was born at term from healthy non consanguineous parents with mild macrosomia (weight: 4,500 g, height: 51 cm, OFC: 36 cm). He had early hypotonia but walked independently at the age of 13 months. He had a speech delay and his evolution was marked by a moderate ID with learning difficulties for which he attended a special school. He also developed important behavioral troubles including anxiety and aggressiveness that are drug resistant and alter his familial and social relationships. EEG and brain MRI were normal. On clinical examination his weight was 68 kg, and height 179 cm, his OFC was 60.5 cm. He was able to make short sentences but was not able to read (**Table 1**). We noted mild facial features that included triangular face, downslanting palpebral fissures, downturned corners of the mouth, large ears and horizontal eyebrows. Patient P1's brother, individual P2, was 33-year-old at time of examination. He also had early hypotonia but walked independently at the age of 13 months, had a speech delay with mild ID. He developed major anxiety, was able to work but was very slow in daily activities. He had no seizures but his EEG was normal as well as his brain MRI. On clinical examination his weight was 68 kg, and height 180 cm, his OFC was 59.5 cm. He was not clearly dysmorphic but we note horizontal eyebrows and mildly large and floppy ears (**Table 1**).

##### *Genetic analysis*

Exome sequencing (ES) was done in part of diagnostic investigations (Strasbourg University Hospital, France). Briefly, DNA was captured using Medexome kit (Roche) and paired-end sequenced 2 \* 100 bp on HiSeq4000 sequencer (Illumina, Genomeast platform). Parental DNA were captured and sequenced in pools of six individuals (Trio-pool exome sequencing). Sequencing data were analyzed using STARK bioinformatic pipeline using GATK and



CANOES to detect respectively SNV and CNV variants. Variants were annotated and ranked using Varank<sup>1</sup> and AnnotSV<sup>2</sup>. A hemizygous nonsense variant was identified in *SLITRK2* (NM\_032539.4): c.1381G>T, p.E461\*, which had a CADD score of 35 and was absent in gnomAD database in hemizygous state (Table 2). Sanger Sequencing was performed using BigDye Terminator kit v3.1 cycle sequencing kits and run on an ABI Prism 3730XL DNA Analyzer (Perkin Elmer Applied Biosystems) and confirmed the presence of the variant in his affected brother as well as its absence in the blood of patients' mother (Fig. 1a, Fig. S1a, and Table S1).

### Individual P3

#### Clinical data

Individual 3 was a 21-year-old Caucasian male who was born at term from healthy non consanguineous parents. The birth parameters were not available. He had surgery at age two months for a craniosynostosis on the sagittal suture. He had other skeletal abnormalities including a severe kyphoscoliosis requiring surgery and short stature (height at 133 cm and weight at 31 kg). He had severe feeding difficulties and currently uses g-tube exclusively. His milestone was on the normal range until he was six, then a neurologic regression was observed. On examination, he presented with severe ID and minimal speech, spasticity on the lower limb requiring botox injections and dystonia. He developed multifocal seizures at age 10, that responded to levetiracetam. He started to have an unsteady gate at 6 years and progressed to walker then wheel chair. He did not have any behavioral issue. He experienced recurrent urinary tract infections. His brain MRI showed severe, symmetrical cerebral and cerebellar volume loss with prominence of the cortical sulci and associated ex vacuo dilation of the ventricles, atrophy of the corpus callosum and brainstem and bilateral hippocampal atrophy with increased FLAIR signal (Table 1).

#### Genetic analysis

ES (Columbia University, New York, United States) identified a hemizygous missense variant in *SLITRK2* (NM\_032539.4): c.934A>G, p.T312A, which was predicted to be pathogenic by SIFT, PPH2 and MutationTaster softwares with a CADD score of 24.8, and was absent in gnomAD database in hemizygous state. The variant was

33

inherited from the mother (Fig. 1a, Fig. S1a, and Table S1).

#### **Individual P4**

##### ***Clinical data***

Individual P4 was a 13 year-old Caucasian female without any familial history. She was born at term with disharmonious IUGR (weight: 2,460 g, height: 46 cm, OFC: 34 cm). She had failure to thrive with microcephaly with severe developmental delay and severe ID. On examination, height was 129 cm (-2.5 SDS), weight was 27.3 kg (-2.5 SDS) and OFC 51 cm (3<sup>rd</sup> percentile). She had important gastroesophageal reflux and scoliosis. Speech was absent and she could walk with support with an unsteady gait and lower limbs spasticity. She developed generalized seizures at age 11 months partially resistant to drugs. Neuropsychiatric troubles were also noted and included ASD, aggressiveness with self-injury and mutism. Her brain MRI showed thin corpus callosum, white matter diffuse reduction and leukomalacia. Hypertrichosis was noted on the upper limbs and the sacral region. Facial facial included coarse face aspect with broad eyebrows, deep set eyes, bulbous nasal tip, full and thick lips and prognathism (Table 1).

##### ***Genetic analysis***

ES was performed through Italian Telethon Undiagnosed Disease Program in which the patient had been included (Naples, Italy) and identified the heterozygous missense variant in *SLITRK2* (NM\_032539.4): c.1276C>T, p.R426C, which was predicted to be pathogenic by SIFT, PPH2 and MutationTaster softwares with a CADD score of 28.2 and was absent in gnomAD database in hemizygous state. The mother did not carry the variant, which arose *de novo* (Fig. 1a, Fig. S1a, and Table S1).

#### **Individual P8**

##### ***Clinical data***

Individual P8 was a 11 years-old male born from an uneventful pregnancy. His maternal uncle had isolated epilepsy. His birth weight was 3,250 g (term 40 WG), and height was 52 cm. He had a mild developmental delay with slight delay in verbal abilities and mild ID. He developed focal seizures at age 8, which were sensible to valproate acid. EEG showed left centrottemporal epileptic discharges. He had no neurological regression nor dystonia but had significative anxiety, became easily frustrated, had executives and interaction difficulties in the context of ADHD. He attended mainstream school with extra help. At age 11, his height was 142.5 cm and his weight was 30 kg. Brain MRI showed unspecific minor white substance punctate changes on the right side around trigonum (**Table 1**).

### **Genetic analysis**

Trio ES (Dianalund, Denmark) was performed as followed: after enrichment of exonic DNA fragments with a SureSelect Human All Exon Kit (Agilent, 50 Mb V5), sequencing was performed on a HiSeq2500 system (Illumina). The genomic regions targeted by the respective enrichment design had an average coverage of > 100 reads and > 97% were covered by  $\geq$  10 reads. An hemizygous missense variant was identified in *SLITRK2* (NM\_032539.4): c.1121C>G, p.P374R, predicted to be pathogenic by SIFT, PPH2 and MutationTaster softwares with a CADD score of 26 and was absent in gnomAD database in hemizygous state. Sanger sequencing was performed to confirm the variant, which was inherited from the asymptomatic mother but arose *de novo* in the latter since the maternal grandparents did not carry it (**Fig. 1a**, **Fig. S1a**, and **Table S1**).

### **Individual P5**

#### **Clinical data**

Individual P5 was a 12 years-old male without any family history. He was born prematurely at term 25 WG with a weight of 770 g and a height of 33.5 cm. He had a persistent ductus arteriosus in this context. In the first months he had significant feeding difficulties with GOR and failure to thrive leading to a percutaneous endoscopic gastrostomy. He still had short stature and low weight (5<sup>th</sup> centile at age 7). He was followed for a strabismus and a lumbar scoliosis. Motor and language development were slightly delayed and he had a moderate to severe ID. He had

spastic diplegia and dystonia with an unsteady gait requiring a walker and a wheelchair for support getting around. He was anxious and had signs of ADHD, which was not formerly assessed. He was not clearly dysmorphic. Brain MRI performed at the age of 11 months showed bilateral periventricular leukomalacia, enlarged lateral ventricles and a paucity of the white matter (**Table 1**).

### **Genetic analysis**

Trio ES was performed at the Yale Center for Genome Analysis (YCGA). Genomic DNA was captured using the VCRome capture kit followed by Illumina DNA sequencing. WES data was processed using two independent pipelines at the Yale School of Medicine and Phoenix Children's Hospital. At each site sequence reads were mapped to the reference genome (GRCh37) with BWA-MEM and further processed using GATK Best Practice workflows, which include duplication marking, indel realignment, and base quality recalibration. Single nucleotide variants and small indels were called with GATK HaplotypeCaller and annotated using ANNOVAR, dbSNP (v138), 1000 Genomes (August, 2015), NHLBI Exome Variant Server (EVS), and the Exome Aggregation Consortium v3 (ExAC). Rare deleterious missense variants and LOF variants (stop-gain, stop-loss, frameshift insertions/deletions, canonical splice site, and start-loss) were selected. MetaSVM and Combined Annotation Dependent Deletion (CADD v1.3) algorithms were used to predict deleteriousness of missense variants (MetaSVM-deleterious or CADD  $\geq$  20). A hemizygous missense variant was identified in *SLITRK2* (NM\_032539.4): c.1531G>A p.V511M, predicted to be pathogenic by SIFT, PPH2 and MutationTaster softwares with a CADD score of 25.2 and was absent in gnomAD database in hemizygous state. The variant was inherited from the asymptomatic mother; the maternal grand-parents were not tested (**Fig. 1a**, **Fig. S1a**, and **Table S1**).

### **Individual P10**

#### **Clinical data**

Individual P10 is an 11 year-old boy with no familial history of neurodevelopmental disabilities. He was born at full term via C-section to a 29 year-old G1 mother. Apgar score was 8 at both 1 and 5 minutes. He had a pneumothorax

diagnosed shortly after birth due to respiratory symptoms, but was discharged home at 5 days of age. He had feeding difficulties in infancy due to gastroesophageal reflux, and was fussy and “gassy”. He had abdominal pain as he got older and was found to have mildly elevated liver enzymes that later normalized. Otherwise he has been healthy and well grown. He achieved all gross motor milestones on time, but he had delays in fine motor skills development and has required therapy services. Language development was normal, but he received therapy for issues with pragmatic speech. He does not have intellectual disabilities, but he has significant neuropsychological problems, including ADHD, anxiety, obsessive compulsive behaviors, tantrums, vocal tics and features of autism (poor socialization, sensitivity to loud sounds, repetitive behaviors). His anxiety has improved since he has begun taking low dose fluoxetine. At 11 years, his height is 144.8 cm and weight is 42 kg. OFC at 8 years was 53.8 cm. He is non-dysmorphic and physical examination is normal. Brain MRI has not been performed (**Table 1**).

#### **Genetic analysis**

ES (Washington University School of Medicine, Saint-Louis, United States) identified a hemizygous missense variant in *SLITRK2* (NM\_032539.4): c.628G>A, p.E210K, which was predicted to be pathogenic by SIFT, PPH2 and MutationTaster softwares with a CADD score of 27.2 and was absent in gnomAD database in hemizygous state. The variant was inherited from the asymptomatic mother (**Fig. 1a**, **Fig. S1a**, and **Table S1**).

#### **Individual P7**

##### **Clinical data**

Individual P7 was a 12 years-old boy with no familial history. He was born prematurely at 34 WG with a weight at 2100g. Early hypotonia was noticed followed by a developmental delay, including speech delay. At presented with joint laxity, pedes planovalgi and mild thoracal scoliosis with increased lumber lordosis. At age 9 the height was 146 cm, the weight was 43.2 kg and the OFC was 55.7 cm. He had no epilepsy nor dystonia but had an unsteady gait. He developed a severe ID (IQ 20-36 at age 12 years). He had mild facial dysmorphic features including mild facial asymmetry, slightly broad forehead, full nasal tip, long and smooth philtrum, thin lips and prominent ears 37

(photographs not available). He was diagnosed with an ASD, important anxiety and hyperactivity. A brain MRI was not done (**Table 1**).

### **Genetic analysis**

ES (Radboud university medical center, Nijmegen, The Netherlands) identified the hemizygous missense variant in *SLITRK2* (NM\_032539.4) c.221T>C, p.L74S, which was predicted to be pathogenic by SIFT, PPH2 and MutationTaster softwares with a CADD score of 26,4 and was absent in gnomAD database in hemizygous state. The variant was inherited from the asymptomatic mother, as well as a hemizygous pathogenic variant in the *ARX* gene (c.1109C>T, p.A370V)<sup>3</sup> (See also **Fig. 1a** and **Table S1**). The maternal grandparents were not available but both variants were absent in the healthy maternal uncle.

### **Supplementary References**

1. Geoffroy, V., et al. VaRank: a simple and powerful tool for ranking genetic variants. *PeerJ* 3, e796 (2015).
2. Geoffroy, V., et al. AnnotSV: an integrated tool for structural variations annotation. *Bioinformatics* 34, 3572-3574 (2018).
3. Thai, M.H.N., et al. Constraint and conservation of paired-type homeodomains predicts the clinical outcome of missense variants of uncertain significance. *Hum Mutat*, 41, 1407–1424. (2020)

# Annexe 3 : Élargissement du spectre phénotypique associé aux variants pathogènes du gène *ARCN1*

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



## HHS Public Access

Author manuscript

*Genet Med.* Author manuscript; available in PMC 2023 February 13.

Published in final edited form as:

*Genet Med.* 2022 June ; 24(6): 1227–1237. doi:10.1016/j.gim.2022.02.005.

### Expanding the phenotypic spectrum of *ARCN1*-related syndrome

Alyssa L. Ritter<sup>1</sup>, Jessica Gold<sup>1</sup>, Hiroshi Hayashi<sup>1</sup>, Amanda M. Ackermann<sup>2</sup>, Stephanie Hanke<sup>2</sup>, Cara Skraban<sup>1</sup>, Sanmati Cuddapah<sup>1</sup>, Elizabeth Bhoj<sup>1</sup>, Dong Li<sup>3</sup>, Yukiko Kuroda<sup>1</sup>, Jessica Wen<sup>4</sup>, Ryojun Takeda<sup>5</sup>, Audrey Bibb<sup>6</sup>, Salima El Chehadeh<sup>7,8</sup>, Amélie Piton<sup>9,10</sup>, Jeanine Ohl<sup>11</sup>, Mary K. Kukulich<sup>12</sup>, Keisuke Nagasaki<sup>13</sup>, Kohji Kato<sup>14</sup>, Tomoo Ogi<sup>14</sup>, Tricia Bhatti<sup>15</sup>, Pierre Russo<sup>15</sup>, Bryan Krock<sup>16</sup>, Jill R. Murrell<sup>16</sup>, Jennifer A. Sullivan<sup>17</sup>, Vandana Shashi<sup>17</sup>, Nicholas Stong<sup>18</sup>, Hakon Hakonarson<sup>3</sup>, Kentaro Sawano<sup>13</sup>, Erin Torti<sup>19</sup>, Rebecca Willaert<sup>19</sup>, Yue Si<sup>19</sup>, William Ross Wilcox<sup>6</sup>, Katrine Verena Wirgenes<sup>20,21</sup>, Kristian Thomassen<sup>22</sup>, Katherine Carlotti<sup>23</sup>, Angelika Erwin<sup>23</sup>, Joanna Lazier<sup>24</sup>, Thorsten Marquardt<sup>25</sup>, Miao He<sup>26</sup>, Andrew C. Edmondson<sup>1</sup>, Kosuke Izumi<sup>1,\*</sup>

<sup>1</sup>Division of Human Genetics, Department of Pediatrics, The Children's Hospital of Philadelphia, Philadelphia, PA

<sup>2</sup>Division of Endocrinology and Diabetes, Department of Pediatrics, The Children's Hospital of Philadelphia, Philadelphia, PA

<sup>3</sup>Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, PA

<sup>4</sup>Division of Gastroenterology, Hepatology and Nutrition, Department of Pediatrics, The Children's Hospital of Philadelphia, Philadelphia, PA

<sup>5</sup>Division of Genetics, Nagano Children's Hospital, Nagano, Japan

\*Correspondence and requests for materials should be addressed to Kosuke Izumi, Division of Human Genetics, Department of Pediatrics, The Children's Hospital of Philadelphia, 3401 Civic Center Blvd, Philadelphia, PA 19104. izumik1@chop.edu.

#### Author Information

Conceptualization: A.L.R., J.G., A.C.E., K.I.; Data Curation: A.L.R., K.I.; Formal Analysis: A.L.R., K.I.; Project Administration: A.L.R.; Resources: A.L.R., J.G., Hi.H., A.M.A., S.H., C.S., S.C., E.B., D.L., Y.K., A.B., S.E.C., A.P., J.O., M.K.K., K.N., K.K., T.O., T.B., P.R., B.K., J.R.M., J.A.S., V.S., N.S., Ha.H., K.S., R.T., W.R.W., K.V.W., K.T., K.C., A.E., J.L., M.H., A.C.E., J.W., E.T., R.W., Y.S., T.M., K.I.; Supervision: K.I.; Visualization: J.G., A.C.E., A.L.R., K.I.; Writing-original draft: A.L.R., J.G., C.S., J.W., K.I.; Writing-review and editing: J.G., J.W., A.C.E., A.L.R., K.I.

#### Ethics Declaration

This study was reviewed and approved by the Children's Hospital of Philadelphia Institutional Review Board (IRB) (#16-013231) for patients identified through the institution. Each patient was enrolled in appropriate local IRB studies per their institutional requirements. Informed consent was obtained from all participants as required by the IRB and local institutional requirements. Clinical data has been de-identified. Written consent for inclusion in this publication was obtained for all participants, including for publication of participant photographs. This study adhered to the principles set out in the Declaration of Helsinki. Data from a previously published study was collected in compliance with appropriate principles of research ethics and IRB approval (Children's Hospital of Philadelphia IRB #14-011223).

#### Conflict of Interest

The following authors declare no conflicts of interest: A.L.R., J.G., Hi.H., A.M.A., S.H., C.S., S.C., E.B., D.L., Y.K., A.B., S.E.C., A.P., J.O., M.K.K., K.N., K.K., T.O., T.B., P.R., B.K., J.R.M., J.A.S., V.S., N.S., Ha.H., K.S., R.T., W.R.W., K.V.W., K.T., K.C., A.E., J.L., M.H., A.C.E., J.W., E.T., J.L., T.M., M.H., A.C.E., and K.I.

J.W. receives research support from Gilead Sciences, Inc; AbbVie; and Alexion and has consulting relationship with Gilead Sciences, Inc.

E.T., R.W., and Y.S. are employees of GeneDx, Inc.

#### Additional Information

The online version of this article (<https://doi.org/10.1016/j.gim.2022.02.005>) contains supplementary material, which is available to authorized users.

Author Manuscript

<sup>6</sup>Department of Human Genetics, Emory University School of Medicine, Emory University, Atlanta, GA

<sup>7</sup>Service de Génétique Médicale, Institut de Génétique Médicale d'Alsace (IGMA), Hôpitaux Universitaires de Strasbourg, Strasbourg, France

<sup>8</sup>Laboratoire de Génétique Médicale, UMR\_S1112, Institut de Génétique Médicale d'Alsace (IGMA), Université de Strasbourg et INSERM, Strasbourg, France

<sup>9</sup>Department of Translational Medicine and Neurogenetics, Institut Génétique Biologie Moléculaire Cellulaire, IGBMC - CNRS UMR 7104 - Inserm U 1258, Illkirch, France

<sup>10</sup>Laboratoire de Diagnostic Génétique, Hôpitaux Universitaires de Strasbourg, Strasbourg, France

<sup>11</sup>Service d'assistance Médicale à la Procréation, Centre médico-chirurgical et obstétrical (CMCO), Schiltigheim, France

<sup>12</sup>Department of Genetics, Cook Children's Medical Center, Cook Children's Health Care System, Fort Worth, TX

<sup>13</sup>Department of Pediatrics, Niigata University Medical & Dental Hospital, Niigata, Japan

<sup>14</sup>Department of Genetics, Research Institute of Environmental Medicine, Nagoya University, Nagoya, Japan

<sup>15</sup>Division of Anatomic Pathology, The Children's Hospital of Philadelphia, Philadelphia, PA

<sup>16</sup>Division of Genomic Diagnostics, Department of Pathology and Laboratory Medicine, The Children's Hospital of Philadelphia, Philadelphia, PA

<sup>17</sup>Division of Medical Genetics, Department of Pediatrics, Duke University Medical Center, Duke University School of Medicine, Durham, NC

<sup>18</sup>Institute for Genomic Medicine, Columbia University, New York, NY

<sup>19</sup>GeneDx, Gaithersburg, MD

<sup>20</sup>Department of Medical Genetics, Oslo University Hospital, Oslo, Norway

<sup>21</sup>Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway

<sup>22</sup>Department of Radiology and Nuclear Medicine, Oslo University Hospital, Oslo, Norway

<sup>23</sup>Genomic Medicine Institute, Cleveland Clinic Foundation, Cleveland, OH

<sup>24</sup>Department of Medical Genetics, Children's Hospital of Eastern Ontario, Ottawa, Ontario, Canada

<sup>25</sup>Department of Pediatrics, University Hospital of Muenster, Muenster, Germany

<sup>26</sup>Metabolic and Advanced Diagnostics, The Children's Hospital of Philadelphia, Philadelphia, PA

Author Manuscript

Author Manuscript

Author Manuscript

## Abstract

**Purpose:** This study aimed to describe the phenotypic and molecular characteristics of *ARCNI*-related syndrome.



**Methods:** Patients with *ARCNI* variants were identified, and clinician researchers were connected using GeneMatcher and physician referrals. Clinical histories were collected from each patient.

**Results:** In total, we identified 14 cases of *ARCNI*-related syndrome, (9 pediatrics, and 5 fetal cases from 3 families). The clinical features these newly identified cases were compared to 6 previously reported cases for a total of 20 cases. Intrauterine growth restriction, micrognathia, and short stature were present in all patients. Other common features included prematurity (11/15, 73.3%), developmental delay (10/14, 71.4%), genitourinary malformations in males (6/8, 75%), and microcephaly (12/15, 80%). Novel features of *ARCNI*-related syndrome included transient liver dysfunction and specific glycosylation abnormalities during illness, giant cell hepatitis, hepatoblastoma, cataracts, and lethal skeletal manifestations. Developmental delay was seen in 73% of patients, but only 3 patients had intellectual disability, which is less common than previously reported.

**Conclusion:** *ARCNI*-related syndrome presents with a wide clinical spectrum ranging from a severe embryonic lethal syndrome to a mild syndrome with intrauterine growth restriction, micrognathia, and short stature without intellectual disability. Patients with *ARCNI*-related syndrome should be monitored for liver dysfunction during illness, cataracts, and hepatoblastoma. Additional research to further define the phenotypic spectrum and possible genotype–phenotype correlations are required.

#### Keywords

ARCNI; COPI; Micrognathia

---

# Annexe 4 : Syndrome d'Ehlers-Danlos parodontal : description de 13 nouveaux cas et élargissement du phénotype

Received: 12 December 2020 | Revised: 18 April 2021  
DOI: 10.1111/cge.13972

SHORT REPORT

CLINICAL GENETICS WILEY

## Periodontal (formerly type VIII) Ehlers–Danlos syndrome: Description of 13 novel cases and expansion of the clinical phenotype

Salima El Chehadeh<sup>1,2,3</sup> | Anne Legrand<sup>4,5</sup> | Corinne Stoetzel<sup>2</sup> |  
Véronique Geoffroy<sup>2</sup> | Clarisse Billon<sup>4,5</sup> | Salma Adham<sup>5</sup> |  
Xavier Jeunemaître<sup>4,5</sup> | Roland Jaussaud<sup>6</sup> | Jean Muller<sup>2,7</sup> | Elise Schaefer<sup>1</sup> |  
Karelle Benistan<sup>8,9</sup> | Sébastien Gaertner<sup>10,11</sup> | Agnès Bloch-Zupan<sup>3,12,13</sup> |  
Aymeric Courval<sup>13</sup> | Marie-Cécile Manière<sup>3,12,13</sup> | Catherine Petit<sup>12,13,14</sup> |  
Anne-Claire Bursztejn<sup>15,16</sup> | Laurence Bal<sup>17</sup> | Anthony Reyre<sup>18</sup> |  
Agathe Chammas<sup>19</sup> | Tiffany Busa<sup>20</sup> | Hélène Dollfus<sup>1,2,21</sup> | Dan Lipsker<sup>15,22</sup>

<sup>1</sup>Service de Génétique Médicale, Institut de Génétique Médicale d'Alsace (IGMA), Hôpitaux Universitaires de Strasbourg, Hôpital de Hautepierre, Strasbourg, France

<sup>2</sup>Laboratoire de Génétique Médicale, UMR5\_1112, Institut de Génétique Médicale d'Alsace (IGMA), Université de Strasbourg et INSERM, Strasbourg, France

<sup>3</sup>Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), INSERM U1258, CNRS-UMR7104, Université de Strasbourg, Illkirch-Graffenstaden, France

<sup>4</sup>INSERM, U970, Paris Centre de Recherche Cardiovasculaire, Université de Paris, Paris, France

<sup>5</sup>Centre de Référence des Maladies Vasculaires Rares, AP-HP, Hôpital Européen Georges Pompidou, Paris, France

<sup>6</sup>Département de Médecine Interne Immunologie Clinique, CHRU de Nancy et Université de Lorraine, Nancy, France

<sup>7</sup>Laboratoire de Diagnostic Génétique, Hôpitaux Universitaires de Strasbourg, Nouvel Hôpital Civil, Strasbourg, France

<sup>8</sup>Centre de Référence des Syndromes d'Ehlers-Danlos non Vasculaires, Hôpital Raymond Poincaré, Garches, Assistance Publique Hôpitaux de Paris, Garches, France

<sup>9</sup>UMR U1179 INSERM, Université Versailles Saint-Quentin, Montigny-le-Bretonneux, France

<sup>10</sup>Service des Maladies Vasculaires - Hypertension Artérielle, Hôpitaux Universitaires de Strasbourg, Nouvel Hôpital Civil, Strasbourg, France

<sup>11</sup>Regenerative Nanomedicine, INSERM, UMR 1260, FMTS, Strasbourg, France

<sup>12</sup>Faculté de Chirurgie Dentaire, Université de Strasbourg, Strasbourg, France

<sup>13</sup>Pôle de Médecine et Chirurgie Bucco-Dentaires, Hôpital Civil, Centre de Référence des Maladies Rares Orales et Dentaires, O-Rares, Filière Santé Maladies Rares TETE COU, European Reference Network ERN CRANIO, Hôpitaux Universitaires de Strasbourg, Strasbourg, France

<sup>14</sup>INSERM, UMR 1260 'Osteoarticular and Dental Regenerative Nanomedicine', Faculté de Médecine, Fédération de Médecine Translationnelle de Strasbourg (FMTS), Strasbourg, France

<sup>15</sup>Clinique Dermatologique, Hôpitaux Universitaires de Strasbourg, Strasbourg, France

<sup>16</sup>Service de Dermatologie, Hôpitaux de Brabois, CHU de Nancy, Vandœuvre-lès-Nancy, France

<sup>17</sup>Timone Aortic Center, Department of Vascular Surgery, Timone Hospital, APHM, Marseille, France

<sup>18</sup>Service de Neuroradiologie Interventionnelle, CHU de Marseille, Hôpital La Timone, Marseille, France

<sup>19</sup>Service de radiologie, Hôpitaux Universitaires de Strasbourg, Hôpital de Hautepierre, Strasbourg, France

<sup>20</sup>Département de Génétique Médicale, CHU de Marseille, Hôpital La Timone, Marseille, France

<sup>21</sup>Filière SENSGENE, Centre de Référence Pour les Affections Rares en Génétique Ophthalmologique, CARGO, Hôpitaux Universitaires de Strasbourg, Strasbourg, France

<sup>22</sup>Faculté de Médecine, Université de Strasbourg, Strasbourg, France

### Correspondence

Salima El Chehadeh, Service de Génétique Médicale, Institut de Génétique Médicale d'Alsace (IGMA), Hôpitaux Universitaires de Strasbourg, Hôpital de Hautepierre, 1 Avenue

### Abstract

Periodontal Ehlers–Danlos syndrome (pEDS) is a rare condition caused by pathogenic variants in the *C1R* and *C1S* genes, encoding subunits C1r and C1s of the first

Molière, Strasbourg 67098, France.  
Email: salima.elchehadeh@chru-strasbourg.fr

#### Funding information

Groupement Interrégional de Recherche  
Clinique et d'Innovation Est; Société Française  
de Dermatologie

component of the classical complement pathway. It is characterized by early-onset periodontitis with premature tooth loss, pretibial hyperpigmentation and skin fragility. Rare arterial complications have been reported, but venous insufficiency is rarely described. Here we report 13 novel patients carrying heterozygous pathogenic variants in *C1R* and *C1S* including three novel *C1S* variants (c.962G > C, c.961 T > G and c.961 T > A). In addition to the pEDS phenotype, three patients and one relative displayed widespread venous insufficiency leading to persistent varicose leg ulcers. One patient suffered an intracranial aneurysm with familial vascular complications including thoracic and abdominal aortic aneurysm and dissection and intracranial aneurysm rupture. This work confirms that vascular complications can occur, although they are not frequent, which leads us to propose to carry out a first complete non-invasive vascular evaluation at the time of the diagnosis in pEDS patients. However, larger case series are needed to improve our understanding of the link between complement pathway activation and connective tissue alterations observed in these patients, and to better assess the frequency, type and consequences of the vascular complications.

#### KEYWORDS

aneurysm, aortic dissection, Ehlers–Danlos syndrome periodontal type, Ehlers–Danlos syndrome type VIII, persistent ulcers, tooth loss, vascular complications, venous insufficiency

## 1 | INTRODUCTION

Periodontal Ehlers–Danlos syndrome (pEDS) (formerly type VIII EDS) is a distinct subtype of Ehlers–Danlos syndrome, which is a genetically and phenotypically heterogeneous group of rare disorders affecting connective tissue.<sup>1</sup> Periodontal EDS is caused by autosomal dominant pathogenic variants in the *C1R* (type 1, *MIM* 613785) and *C1S* (type 2, *MIM* 120580) genes, which encode the C1r and C1s subunits of the first component of the classical complement pathway, which has a key role in the innate immune response.<sup>2</sup> Periodontal EDS is mainly characterized by early-onset and severe periodontitis,<sup>3,4</sup> pretibial hyperpigmentation, and global skin fragility including abnormal scars and easy bruising.<sup>5</sup> The periodontitis can begin very early in life, with a mean age at diagnosis of 12 years (range 2–29 years)<sup>6–8</sup> and often progresses rapidly. It is associated with peculiar signs consisting in a lack of attached gingiva and gingival recession leading to inflammatory destruction of dental attachments and premature loss of teeth.<sup>3,4,8,9</sup> Rare severe vascular complications such as potentially lethal arterial rupture have been reported,<sup>7,10,11</sup> but venous insufficiency is rarely described.<sup>12</sup> Herein, we report 13 novel pEDS cases, with a focus on the vascular features.

## 2 | CASE DESCRIPTION

Thirteen pEDS affected patients, aged 3–74 years, were recruited through a French collaborative study. All patients' families provided

written informed consent and all procedures performed in the studies were done in accordance with the ethical standards of the institutional research committee and with the Declaration of Helsinki. The diagnostic confirmation was based on genetic diagnosis either by Sanger sequencing of the *C1R* and *C1S* genes (RefSeq NM\_001733.5 and NM\_001734.4) or by exome sequencing. The vascular assessment included venous Doppler ultrasound (F1P1, F2P1, F3P1, F7P2), arterial Doppler ultrasound or angio-TDM (F1P1, F2P1, F3P1, F6P1), angio-MRI (F2P1) and cerebral arteriography (F4P1). Further details about patients' recruitment and methodology are provided in Supporting Information S1.

We were able to collect clinical and molecular data for 13 molecularly confirmed pEDS patients, including 3 sporadic and 10 familial cases. The mean age at examination was 36 years (range: 3–74 years). The clinical and molecular data are given in Table 1 and detailed clinical description can be found in the Supporting information S1. All patients presented the main clinical signs of pEDS that included pretibial hyperpigmentation (Figure 1(A–D)), early-onset periodontitis and tooth loss (Figure 1(K–M)), and global skin fragility including abnormal scars (Figure 1(A)) and easy bruising for very minor trauma (Figure 1(C)). Periodontitis occurred at a mean age of 18 years (range: 12–23 years) and complete tooth loss was observed at 24, 29, 30 and 35 years in four patients (F1P1, F2P1, F7P2, and F7P3). In addition to the pEDS phenotype, three patients and one relative had proven venous insufficiency.

Patient F1P1, a 41-year-old Caucasian male, had a pretibial hematoma that evolved into a chronic wound on the anterior part of the shin

TABLE 1 Phenotype of the pEDS cases

Phenotype of the pEDS cases		F1	F2	F3	F4	F4	F4	F5	F6	F7	F7	F7	F7	Total
Families	P1	P2	P1	P2	P1	P2	P3	P1	P1	P1	P1	P2	P3	
Patients	M	F	M	F	F	M	F	F	M	M	M	F	M	
Gender	41 y	29 y	61 y	23 y	52 y	16 y	21 y	28 y	22 y	31 y	74 y	61 y		
Age	Oral features													
Early-onset periodontitis	+	+	+	+	+	+	+	+	+	+	+	+	+	12/12
Age at diagnostic of periodontitis	18 y	12 y	<20 y	<20 y	<15 y	<14 y	<20 y	23 y	20 y	NA	17 y	18 y		
Age of complete tooth loss	24 y	29 y	NA	N/A	N/A	N/A	N/A	N/A	N/A	N/A	30 y	35 y		
Gingival recession	+	+	NA	+	NA	+	NA	+	+	NA	+	+		8/8
Absence of attached gingiva	+	+	NA	NA	NA	NA	NA	+	NA	NA	+	+	NA	4/4
Skin and nails														
Easy bruising	+	+	+	+	+	+	+	+	+	+	+	+	+	13/13
Prebital pigmentation	+	+	+	+	+	+	+	+	+	+	+	+	+	12/13
Skin fragility (Mild) elastic skin	+	+	+	+	+	+	+	+	+	+	+	+	+	13/13
Abnormal scarring	+	+	+	+	+	+	+	+	+	+	+	+	+	5/11
Thin skin, visible veins	+	+	+	+	+	+	+	+	+	+	+	+	+	13/13
Fragile nails	+	+	+	+	+	+	+	+	+	+	+	+	+	9/12
Acrogeria	+	+	+	+	+	+	+	+	+	+	+	+	+	8/11
Skeletal features	+	+	+	+	+	+	+	+	+	+	+	+	+	5/11
(Mild) joint hypermobility	+	+	+	+	+	+	+	+	+	+	+	+	+	5/13
Joint pain	+	+	+	+	+	+	+	+	+	+	+	+	+	8/13
Cyphoscoliosis	+	+	+	+	+	+	+	+	+	+	+	+	+	3/11
Vascular abnormalities														
Venous insufficiency	+	+	+	+	+	+	+	+	+	+	+	+	+	3/4
Arterial aneurysms	+	+	+	+	+	+	+	+	+	+	+	+	+	1/5
Recurrent infections	+	+	+	+	+	+	+	+	+	+	+	+	+	4/13
Other clinical features														
Nail hypoplasia of all the toes, chronic asthma	+	+	+	+	+	+	+	+	+	+	+	+	+	
Pectus excavatum	+	+	+	+	+	+	+	+	+	+	+	+	+	
Intermittent headache, anorexia nervosa	+	+	+	+	+	+	+	+	+	+	+	+	+	
Chronic asthma	+	+	+	+	+	+	+	+	+	+	+	+	+	
Recurrent inguinal hernia	+	+	+	+	+	+	+	+	+	+	+	+	+	
Stretch marks	+	+	+	+	+	+	+	+	+	+	+	+	+	
Uterine rupture	+	+	+	+	+	+	+	+	+	+	+	+	+	
Chronic asthma	+	+	+	+	+	+	+	+	+	+	+	+	+	
Mild learning difficulties	+	+	+	+	+	+	+	+	+	+	+	+	+	
Chronic asthma	+	+	+	+	+	+	+	+	+	+	+	+	+	
Intermittent headache, anorexia nervosa	+	+	+	+	+	+	+	+	+	+	+	+	+	
Digestive bleeding	+	+	+	+	+	+	+	+	+	+	+	+	+	

(Continues)

TABLE 1 (Continued)

Phenotype of the pEDS cases											
Families	F1	F2	F3	F4	F4	F4	F5	F6	F7	F7	Total
Gene involved	C1R	C1R	C1R	C1S	C1S	C1S	C1R	C1S	C1S	C1S	C1S
Variants	c.926G > T, p. (Cys309Phe)	c.976G > T, p. (Cys309Phe)	c.926G > T, p. (Cys309Phe)	c.926G > T, p. (Cys309Phe)	c.962G > C, p. (Cys321Ser)	c.962G > C, p. (Cys321Ser)	c.962G > C, p. (Cys321Ser)	c.961T > G, p. (Cys321Gly)	c.961T > A, p. (Cys321Ser)	c.961T > A, p. (Cys321Ser)	c.961T > A, p. (Cys321Ser)

Note: Some recurrent facial features were noted including an elongated face with a prominent tip and narrow root of the nose; prominent chin, marked nasolabial folds; high hair line and thin upper lip. Abbreviations: + feature is present; - feature is absent; NA, not available; N/A, not applicable; WM, white matter.

that never healed despite skin grafting (Figure 1(D)). The Doppler Ultrasound revealed deep venous insufficiency secondary to significant left popliteal reflux at 2.5 s, associated with a varicose tributary vein originating from a lower paratibial perforating vein feeding the leg ulcer, confirming the diagnosis of varicose ulcer (Figure 1(H,I)). The vascular assessment was completed by arterial Doppler ultrasound that was normal.

Patient F3P1, a 61-year-old Caucasian male, also suffered from superficial venous insufficiency, complicated by two chronic varicose ulcers. He underwent endovenous treatment of insufficiency of the left great saphenous vein. Angio-TDM showed no aneurysmal or dissecting arterial lesions. Patient F7P2, a 74-year-old female, presented varicose veins from the age of 30, complicated by an ulcer on the left shin which took more than a year to heal. The Doppler ultrasound of the left leg showed a chronic superficial venous insufficiency by ostio-truncal incontinence of the great saphenous vein with varicose tributary veins and a voluminous paratibial incontinent perforating vein. On the right leg, the small saphenous vein was discreetly varicose. Her affected mother and maternal aunt also had venous insufficiency with early varicose veins and a persistent leg ulceration requiring skin graft in her aunt. The later also had early tooth loss, tibial pigmentation and atrophic scars. With regards to the arterial vascular phenotype, the maternal cousin of F7P2's mother died at the age of 50 secondary to a thoracic aortic dissection and had a son with a dilated thoracic aorta. Unfortunately, none of them had molecular analysis.

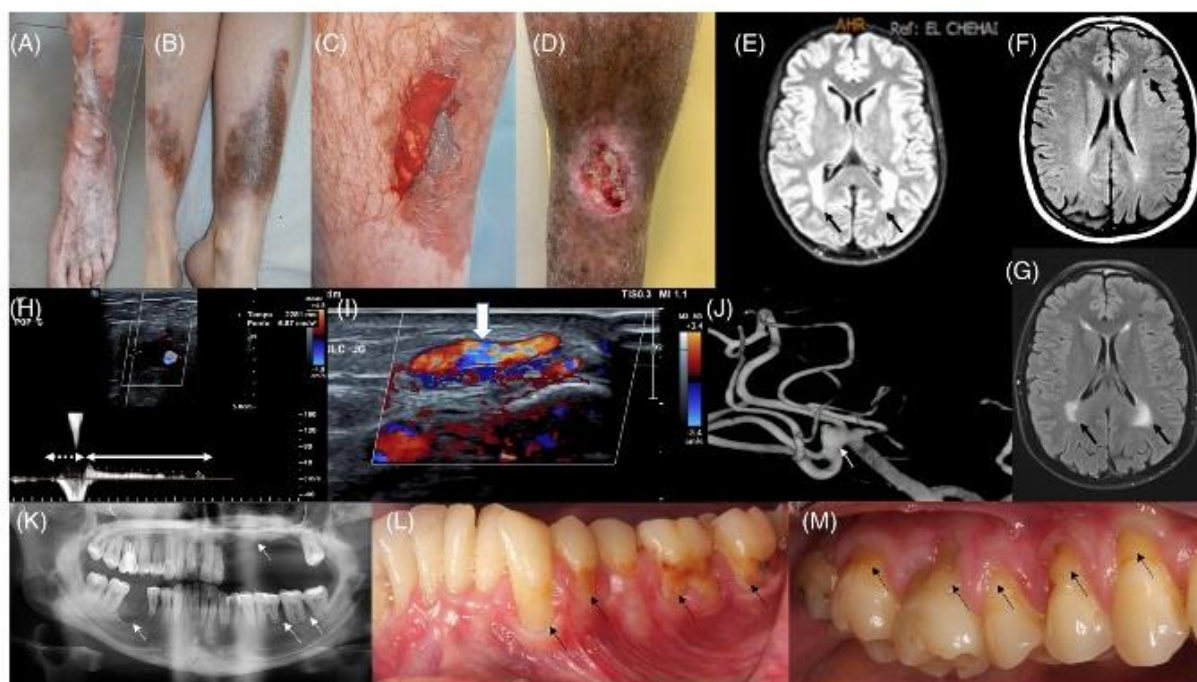
Patient F4P1, a 42-year-old Caucasian female, had a family history of pEDS with vascular complication in several relatives. Her affected maternal uncle had surgery for thoracic aortic dissection, one of her affected maternal aunts had abdominal aortic dissection and a maternal affected cousin presented with rupture of an intracranial aneurysm in her twenties. All of them had premature tooth loss and tibial pigmentation. Patient F4P1 had a brain arteriography that revealed a small aneurysm of the middle cerebral artery division that did not require any treatment (Figure 1(J)).

We were able to obtain brain MRI results for only two patients. For F1P1, who had no neurological history, brain MRI showed a 4 mm wide lacunar infarct of the left frontal area and bilateral periventricular white matter hyperintensities (Figure 1(F,G)). Brain MRI of patient F2P1 showed extensive and confluent symmetrical periventricular and deep cerebral white matter hyperintensities sparing the basal ganglia and the brainstem (Figure 1(E)).

Four patients had recurrent infections (mainly ENT and urinary tract infections) (Table 1).

### 3 | DISCUSSION

We report 13 patients with pEDS who had the typical previously described pEDS phenotype. Periodontal EDS affected patients have a pretibial pigmentation very suggestive of venous insufficiency (*ochre dermatitis*), though Doppler ultrasound is strictly normal in most of them, as in patient F2P1. Other patients however, like patients F1P1, F3P1, F7P2 and her maternal aunt, were found to have proven



**FIGURE 1** Dermatological, dental and vascular findings in pEDS patients. (A,B) Hyperpigmentation of the pre-tibial skin of patients F7P1 (A) and F4P1 (B). Note the atrophic scars in F7P1. (C) Thin and fragile aspect of pretibial skin of patient F6P1, which peels off after light contact revealing subcutaneous bleeding. This bleeding also appears to be present under the skin that has not come off, possibly explaining in part the pretibial skin pigmentation. (D) Tibial pigmentation and persistent varicose leg ulceration in patient F1P1. (E) Brain MRI of patient F2P1 (axial FLAIR) showing extensive and confluent symmetrical periventricular and deep cerebral white matter hyperintensities (arrows) sparing the basal ganglia and the brainstem. (F-G) Brain MRI of patient F1P1 (axial FLAIR) showing a left frontal lacunar lesion (arrow) (F) and periventricular white matter hyperintensities (G) (arrows). (H-I) Venous Doppler ultrasound of the lower limbs of F1P1 showing a significant left popliteal reflux greater than 1 s (full arrow) indicating deep venous insufficiency (dotted arrow: venous flushing) (H) and a varicose tributary vein originating from a lower paratibial perforating vein and feeding the left leg ulcer (arrow) (I). (J) Cerebral arteriography of patient F4P1 showing a small aneurysm on a dysplasia of the right sylvian division of about 3.2 mm in its longest axis  $\times$  1.6 mm. (K) Orthopantomogram of patient F2P1 showing the premature and progressive tooth loss with six teeth missing at the age of 26 years. Note the short roots on many teeth (11, 14, 15, 16, 21, 22, 33, 35, 36, 37, 45, 47) and periodontal bone loss in upper and lower jaws (arrows). (L-M) Intraoral photographs of patient F2P1 showing the severe periodontitis leading to soft/hard tissues destruction, the reduced amount of keratinized tissue, root surface exposure (arrows) and premature tooth loss [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

venous insufficiency with persistent chronic leg ulceration. Ronceray and collaborators also described a 32-year-old male patient with pEDS and leg ulcer in whom Doppler ultrasound ruled out an arterial disease and demonstrated varicose veins on the left leg emerging from incontinent neosaphenous and perforating veins.<sup>12</sup> Mechanisms of pretibial pigmentation in pEDS patients seem therefore to be complex and the origin of the chronic ulcers are probably multifactorial. We hypothesize that in these patients, skin and vascular (in particular small vessels) fragility predisposes to the occurrence of easy shin hematomas in the context of minimal trauma, leading to leg ulceration and that the presence of venous insufficiency leads to chronicity and non-healing of the wound, despite skin grafts. Venous abnormalities are poorly described and often not sought in EDS, especially in pEDS. In particular, varicose veins have been mainly described in patients with classical EDS (6%), spondylodysplastic EDS (25%)<sup>11</sup> and vascular EDS (28%)<sup>13</sup> but very rarely in pEDS patients.<sup>12</sup> However, in a series of 98 affected pEDS patients, Kapferer et al. reports vesperal lower

extremity swelling and stabbing pain in three patients (14: III-2, 15: III-5, and 16: II-2), one of whom displayed varicose veins, and leg ulceration (15: II-3),<sup>7</sup> elements highly suggestive of venous incontinence although not formally sought and confirmed in these patients. The observation of symptomatic venous insufficiency with chronic trophic lesions in three patients and one relative in our series (considering that only four patients had a venous assessment), associated with several cases in the literature, suggests that pEDS may predispose to early onset venous insufficiency, although it does not seem to be frequent. Venous assessments in more pEDS patients however are needed to confirm these observations using Doppler ultrasound, a non-invasive exam.

As pretibial findings like haemosiderin depositions and posttraumatic ulcers are common with pEDS, clinical distinction from venous insufficiency can be difficult, even though the latter can be accompanied by vesperal swollen lower extremities and pain. For these reasons, we propose that venous assessment should be done in these patients to

diagnose venous insufficiency early, and begin appropriate care to prevent or limit the aggravation and chronicity of trophic lesions.

A systematic review of vascular abnormalities in non-vascular EDS found 6% of vascular complications in pEDS patients, including cerebral aneurysms leading to hemorrhages and fatal aortic dissection.<sup>11</sup> In 98 pEDS patients, an incidence of arterial aneurysms of 16% was found. In particular, this study reported a pEDS patient who was first thought to have vascular EDS because of early death from arterial rupture in four maternal relatives (aged 23–46 years).<sup>7</sup> In our series, family F4 is a good representation of arterial fragility in pEDS since F4P1 had a cerebral aneurysm and three relatives suffered from thoracic/abdominal aortic aneurysm and dissection and rupture of cerebral aneurysm (Table 1, Figure 1(J)). Furthermore, the maternal little cousin of patients F7P2 and F7P3 died at the age of 50 secondary to a thoracic aortic dissection and had a son with a thoracic aortic dilation. Unfortunately, the latter never benefited from molecular analysis thus we do not know whether their vascular complications were related to pEDS in isolation, or if another connective tissue condition may have been present in this family in addition, such as Marfan syndrome. These life-threatening vascular complications have however important consequences, suggesting the interest of carrying out more systematic and complete vascular assessment in these patients.

The link between vascular complications and alteration of the complement pathway cannot to date be explained but it is probable that pEDS patients display fragility of both the venous and arterial walls. This hypothesis is supported by the fact that several reported pEDS patients had profuse bleeding after surgery,<sup>7,11</sup> and by the occurrence of tibial haematomas after minimal trauma in these patients, requiring surgical evacuation on three occasions in patient F1P1. Similarly, the tibial skin of patient F6P1 peeled off after very light contact revealing subcutaneous bleeding which also appeared to be present under the healthy skin (Figure 1(C)). This phenomenon, which occurs very frequently, possibly explains in part the pretibial skin pigmentation and supports the hypothesis of a fragility of the small vessels.

It has been recently shown that brain extensive white matter alterations suggestive of an underlying small vessel disease progressive with age, are a feature of periodontal EDS as well as lacunar cerebral lesions. In these series, neurological examination was unremarkable in all individuals but one, who had mild cognitive decline, ataxia and experienced seizures. Two patients had frequent headache and one patient suffered from depression as patient F2P1.<sup>14–16</sup> In patient F1P1, who had no neurological history, brain MRI showed a lacunar infarct of the left frontal area (Figure 1(F)) and bilateral periventricular white matter hyperintensities that have spread with age (Figure 1(G)). Brain MRI of patient F2P1 showed extensive and confluent symmetrical periventricular and deep cerebral white matter hyperintensities sparing the basal ganglia and the brainstem (Figure 1(E)). Brain CT angiography showed no definite vascular anomaly. These white matter hyperintensities most likely correspond to the “leukoencephalopathy” previously described in pEDS.<sup>14–16</sup> These MRI results could have led to the suspicion of a genetic vascular encephalopathy but the normal aspect of the basal ganglia and the brain stem, associated with the absence of

neurological symptomatology was not in favor of that diagnoses. Similarly, the topography of the bilateral and symmetrical lesions and their extensive evolution from the outset were not suggestive of an inflammatory demyelinating leukopathy. In addition, none of our patients had vascular risk factor including diabetes type II, hypercholesterolemia and hypertension. Unfortunately, brain MRI was not performed in the other patients. Further MRI and neurological descriptions of this rare condition are therefore needed to better understand these findings and their potential clinical consequences.

Complete tooth loss occurred at the ages of 24, 29, 30, and 35 years in F1P1, F2P1, F7P2, and F7P3, respectively, consistent with reported cases (mean age of complete tooth loss: 14–48 years)<sup>7</sup> (Table 1). Early tooth loss secondary to severe periodontitis can begin very early in childhood and should lead to early and specialized care.<sup>7,14</sup> Patient F7P1 had regular dental follow-up and has lost only five teeth at age 31 years suggesting that an early diagnosis of pEDS may delay the tooth loss through specialized anticipatory dental care.

The complement system is a major component of innate immunity that is activated through recognition of molecular patterns associated with microorganisms, abnormal host cells and modified molecules.<sup>17</sup> The classical complement pathway, which is one of the three complement activation pathways, is initiated when the C1 complex, formed by C1q and C1r2s2, binds to an activator. The main models state that C1 binding-activation leads to a conformational rearrangement in C1q collagen stems that causes structural reorganization of the C1r2s2.<sup>18,19</sup> The five variations identified in our patients alter cysteines in the Sushi CCP1 domain of C1r and C1s (except for C1R c.905A > G, p.[Tyr302Cys] localized in CUB2 domain<sup>3</sup>). We were not able to perform functional studies regarding the three C1S novel variants we identified but there are strong arguments in favor of their pathogenicity. These novel C1S variants are classified as likely pathogenic according to the ACMG criteria (Table S2). To strengthen these criteria: (a) They are localized in the Sushi 1 domain of C1S, like the other missense variant described in C1S gene by Kapferer-Seebacher et al.<sup>7</sup> (b) They are totally absent from gnomAD. (c) They are predicted to be pathogenic by SIFT, Align-GVGD, MutationTaster, and PolyPhen-2, and affect a strongly conserved amino acid up to zebrafish. Furthermore, the Cysteine in position 321 is affected in our study in three independent families including six patients with a pEDS phenotype (F4, F6 and F7), suggesting an important functional role of this residue.

In conclusion, the association of pretibial hyperpigmentation with possible persistent varicose ulcer and early-onset tooth loss should lead to the suspicion of pEDS and subsequent screening of the C1R and C1S genes. Early diagnosis of pEDS allows starting appropriate dental care to try to delay the tooth loss. Rare vascular complications exist including venous insufficiency and arterial aneurysm and dissection that can be fatal. However, more cases with enhanced vascular characterization are required in order to complete the phenotypic description of this very rare condition and determine if a systematic vascular monitoring, including both arterial and venous assessment, would be recommended. This would also improve the understanding of pEDS pathogenesis, especially the link between the complement pathway activation and the multi-system connective tissue damage.

## ACKNOWLEDGMENTS

The authors thank the patients and their parents for their participation in this study, the Laboratory of medical genetics U1112 and the IGBMC (Institut de Génétique et de Biologie Moléculaire et Cellulaire), the BICS and BISTRO bioinformatics platforms in Strasbourg, for their technical support. The authors also thank the GIRCI EST (Groupement Interrégional de Recherche Clinique et d'Innovation Est) and the Société Française de Dermatologie who financially supported the Dermaseq Study. The authors thank Dr Louise F. Porter for her English review of the manuscript.

## CONFLICT OF INTEREST

The authors declare no conflicts of interest.

## PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/cge.13972>.

## DATA AVAILABILITY STATEMENT

The datasets used and/or analyzed during the current study are available from the corresponding author on request. This manuscript contains shared data.

## ORCID

Salima El Chehadeh  <https://orcid.org/0000-0003-1613-6570>

Clarisse Billon  <https://orcid.org/0000-0002-7104-8496>

Salma Adham  <https://orcid.org/0000-0003-2955-0922>

Jean Muller  <https://orcid.org/0000-0002-7682-559X>

Agathe Chammas  <https://orcid.org/0000-0002-1045-6103>

## REFERENCES

- Malfait F, Francomano C, Byers P, et al. The 2017 international classification of the Ehlers-Danlos syndromes. *Am J Med Genet Part C Semin Med Genet*. 2017;175(1):8-26. <https://doi.org/10.1002/ajmg.c.31552>.
- Gröbner R, Kapferer-Seebacher I, Amberger A, et al. C1R mutations trigger constitutive complement 1 activation in periodontal Ehlers-Danlos syndrome. *Front Immunol*. 2019;10(2537):1-14. <https://doi.org/10.3389/fimmu.2019.02537>.
- Jepsen S, Caton JG, Albandar JM, et al. Periodontal manifestations of systemic diseases and developmental and acquired conditions: consensus report of workgroup 3 of the 2017 world workshop on the classification of periodontal and Peri-implant diseases and conditions. *J Clin Periodontol*. 2018;45(Suppl 20):S219-S229. <https://doi.org/10.1111/jcpe.12951>.
- Albandar JM, Susin C, Hughes FJ. Manifestations of systemic diseases and conditions that affect the periodontal attachment apparatus: case definitions and diagnostic considerations. *J Clin Periodontol*. 2018;45(Suppl 20):S171-S189. <https://doi.org/10.1111/jcpe.12947>.
- Malfait F. Vascular aspects of the Ehlers-Danlos syndromes. *Matrix Biol*. 2018;71-72:380-395. <https://doi.org/10.1016/j.matbio.2018.04.013>.
- Stewart RE, Hollister DW, Rimoin DL. A new variant of Ehlers-Danlos syndrome: an autosomal dominant disorder of fragile skin, abnormal scarring, and generalized periodontitis. *Birth Defects Orig Artic Ser*. 1977;13(3B):85-93.
- Kapferer-Seebacher I, Heiss-Kisielewsky I, Pepin M, et al. Periodontal Ehlers-Danlos syndrome is caused by mutations in C1R and C1S, which encode subcomponents C1r and C1s of complement. *Am J Hum Genet*. 2016;99(5):1005-1014. <https://doi.org/10.1016/j.ajhg.2016.08.019>.
- Kapferer-Seebacher I, Lundberg P, Malfait F, Zschocke J. Periodontal manifestations of Ehlers-Danlos syndromes: a systematic review. *J Clin Periodontol*. 2017;44(11):1088-1100. <https://doi.org/10.1111/jcpe.12807>.
- Rinner A, Zschocke J, Schossig A, Gröbner R, Strobl H, Kapferer-Seebacher I. High risk of peri-implant disease in periodontal Ehlers-Danlos syndrome. A case series. *Clin Oral Implants Res*. 2018;29(11):1101-1106. <https://doi.org/10.1111/clr.13373>.
- Cikla U, Sadighi A, Bauer A, Başkaya MK. Fatal ruptured blood blister-like aneurysm of middle cerebral artery associated with Ehlers-Danlos syndrome type VIII (periodontitis type). *J Neurol Surg Rep*. 2014;75(2):e210-e213. <https://doi.org/10.1055/s-0034-1387185>.
- D'Hondt S, Van Damme T, Malfait F. Vascular phenotypes in nonvascular subtypes of the Ehlers-Danlos syndrome: a systematic review. *Genet Med*. 2018;20(6):562-573. <https://doi.org/10.1038/gim.2017.138>.
- Ronceray S, Miquel J, Lucas A, et al. Ehlers-Danlos syndrome type VIII: a rare cause of leg ulcers in young patients. *Case Rep Dermatol Med*. 2013;2013:469505. <https://doi.org/10.1155/2013/469505>.
- Frank M, Albuissou J, Ranque B, et al. The type of variants at the COL3A1 gene associates with the phenotype and severity of vascular Ehlers-Danlos syndrome. *Eur J Hum Genet*. 2015;23(12):1657-1664. <https://doi.org/10.1038/ejhg.2015.32>.
- Kapferer-Seebacher I, Waisfisz Q, Boesch S, et al. Periodontal Ehlers-Danlos syndrome is associated with leukoencephalopathy. *Neurogenetics*. 2019;20(1):1-8. <https://doi.org/10.1007/s10048-018-0560-x>.
- Seo GH, Kim Y-M, Ghang B, Kim G-H, Lee BH. Ehlers-Danlos syndrome VIII with novel C1R variant accompanying white matter changes. *J Gene Med*. 2019;16(1):43-47. <https://doi.org/10.5734/jgm.2019.16.1.43>.
- Spranger M, Spranger K, Kirchhof BS. Ehlers-Danlos syndrome type VIII and Leukodystrophy. *Am J Med Genet*. 1996;66:2.
- Bajic G, Degn SE, Thiel S, Andersen GR. Complement activation, regulation, and molecular basis for complement-related diseases. *EMBO J*. 2015;34(22):2735-2757. <https://doi.org/10.15252/embj.201591881>.
- Budayova-Spano M, Lacroix M, Thielens NM, Arlaud GJ, Fontecilla-Camps JC, Gaboriaud C. The crystal structure of the zymogen catalytic domain of complement protease C1r reveals that a disruptive mechanical stress is required to trigger activation of the C1 complex. *EMBO J*. 2002;21(3):231-239. <https://doi.org/10.1093/emboj/21.3.231>.
- Kardos J, Harmat V, Palló A, et al. Revisiting the mechanism of the autoactivation of the complement protease C1r in the C1 complex: structure of the active catalytic region of C1r. *Mol Immunol*. 2008;45(6):1752-1760. <https://doi.org/10.1016/j.molimm.2007.09.031>.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** El Chehadeh S, Legrand A, Stoetzel C, et al. Periodontal (formerly type VIII) Ehlers-Danlos syndrome: Description of 13 novel cases and expansion of the clinical phenotype. *Clinical Genetics*. 2021;1-7. <https://doi.org/10.1111/cge.13972>



# Annexe 5 : Description phénotypique du syndrome de Frank-ter Haar

European Journal of Medical Genetics 63 (2020) 103857



Contents lists available at ScienceDirect

European Journal of Medical Genetics

journal homepage: [www.elsevier.com/locate/ejmg](http://www.elsevier.com/locate/ejmg)



## A severe case of Frank-ter Haar syndrome and literature review: Further delineation of the phenotypical spectrum



Benjamin Durand<sup>a</sup>, Corinne Stoetzel<sup>b</sup>, Elise Schaefer<sup>a</sup>, Nadège Calmels<sup>c</sup>, Sophie Scheidecker<sup>c</sup>, Nadine Kempf<sup>c</sup>, Charlie De Melo<sup>d</sup>, Anne-Sophie Guilbert<sup>d</sup>, Dana Timbolschi<sup>e</sup>, Leonardo Donato<sup>e</sup>, Dominique Astruc<sup>d,e</sup>, Arnaud Sauer<sup>f</sup>, Maria Cristina Antal<sup>g</sup>, Hélène Dollfus<sup>a,b</sup>, Salima El Chehadeh<sup>a,\*</sup>

<sup>a</sup> Service de génétique médicale, Institut de Génétique Médicale d'Alsace (IGMA), Hôpitaux Universitaires de Strasbourg, Hôpital de Hautepierre, Strasbourg, France

<sup>b</sup> Laboratoire de Génétique Médicale, INSERM U1112, Institut de Génétique Médicale d'Alsace (IGMA), Université de Strasbourg, Strasbourg, France

<sup>c</sup> Laboratoire de Diagnostic Génétique, Institut de Génétique Médicale d'Alsace (IGMA), Hôpitaux Universitaires de Strasbourg, Nouvel Hôpital Civil, Strasbourg, France

<sup>d</sup> Service de réanimation pédiatrique spécialisée, Pôle médico-chirurgical de pédiatrie, Hôpitaux Universitaires de Strasbourg, Hôpital de Hautepierre, Strasbourg, France

<sup>e</sup> Service de pédiatrie II, Médecine et réanimation du nouveau-né, Pôle médico-chirurgical de pédiatrie, Hôpitaux Universitaires de Strasbourg, Hôpital de Hautepierre, Strasbourg, France

<sup>f</sup> Strasbourg, France

<sup>g</sup> Service d'ophtalmologie, Hôpitaux Universitaires de Strasbourg, Nouvel Hôpital Civil, Strasbourg, France

<sup>h</sup> Service de Pathologie, UF6349 Fœtopathologie, Hôpitaux Universitaires de Strasbourg, Hôpital de Hautepierre, Strasbourg, France

### ARTICLE INFO

**Keywords:**  
Frank-ter Haar syndrome  
*SH3PXD2B*  
Congenital glaucoma  
Megalocornea

### ABSTRACT

Frank-ter Haar syndrome (FTHS) is a rare autosomal recessive syndrome resulting from mutations in the *SH3PXD2B* gene involved in the formation of podosomes and invadopodia which have a role in extracellular matrix remodelling and cell migration. FTHS is characterized by facial dysmorphism, megalocornea, inconstant glaucoma, variable developmental delay, skeletal and cardiac anomalies. To date, 40 patients have been reported in the literature with a clinical diagnosis of FTHS, only 20 patients having identified mutations. We present a review of these 20 reported patients and describe a patient born to non-consanguineous parents, with intrauterine growth retardation, hypotonia, congenital glaucoma, caudal appendix, scoliosis, camptodactyly, ventricular septal defect, thin corpus callosum and craniofacial features suggestive of FTHS. Clinical evolution resulted in buphthalmos worsening, coarsening of the facial features and respiratory failure leading to death at 4,5 months. Diagnosis was confirmed by the identification of a previously known homozygous mutation c.969delG, p.(Arg324Glyfs\*19) in *SH3PXD2B*. This is the first description of very severe phenotype with lethal respiratory impairment in FTHS. Since very few patients are described in the literature, and 2 out of the 3 patients carrying the c.969delG mutation had a favourable clinical course, more cases are needed to better characterize the phenotype and understand the natural history of this syndrome. Furthermore, we hypothesize that the alteration of podosomes function could lead to a reduction of the extracellular matrix degradation and accumulation of the latter in the extracellular space, which might explain the coarsening of the facial features and the severe refractory glaucoma.

### 1. Introduction

Mutations in *SH3PXD2B* cause Frank-ter Haar syndrome (FTHS) (MIM# 249420) a very rare disease characterized by craniofacial anomalies, including hypertelorism, brachycephaly, wide anterior fontanelle, prominent forehead, prominent eyes, associated with skeletal, cardiovascular and ocular anomalies including megalocornea with or without glaucoma, and variable developmental delay (Maas

et al., 2004). Frank et al. described for the first time in 1973 a Bedouin girl born to consanguineous parents with megalocornea, dysmorphic facial features, multiple skeletal anomalies, and developmental delay (Frank et al., 1973). Later, ter Haar et al. described three Dutch siblings with bone changes and craniofacial features and suggested an autosomal recessive form of Melnick-Needles Syndrome (MNS) (ter Haar et al., 1982). In 2010, Iqbal et al. identified, through homozygosity mapping studies, *SH3PXD2B* as the most plausible candidate gene

\* Corresponding author. Service de génétique médicale, Hôpital de Hautepierre, Hôpitaux Universitaires de Strasbourg, 1 Avenue Molière, 67098, Strasbourg Cedex, France.

E-mail address: [salima.elchehadeh@chru-strasbourg.fr](mailto:salima.elchehadeh@chru-strasbourg.fr) (S. El Chehadeh).

<https://doi.org/10.1016/j.ejmg.2020.103857>

Received 2 October 2019; Received in revised form 7 December 2019; Accepted 17 January 2020

Available online 21 January 2020

1769-7212/ © 2020 Elsevier Masson SAS. All rights reserved.

involved in FTFS. Screening of 17 patients from 13 families revealed 5 different homozygous mutations in *SH3PXD2B* (4 intragenic mutations and one deletion of the entire gene) in 10 patients and no *SH3PXD2B* mutation was detected in the 7 remaining patients (Iqbal et al., 2010). In 2014, Wilson et al. identified homozygous mutations in *SH3PXD2B* in three patients originally diagnosed with Borrone Dermato-Cardio-Skeletal syndrome (BDCSS), a multisystemic disorder affecting the skin, bone, joints and heart, first classified as a distinct clinical entity of FTFS (Borrone et al., 1993; Wilson et al., 2014). FTFS and BDCSS are since then considered as the same entity (Wilson et al., 2014).

The *SH3PXD2B* gene (MIM<sup>®</sup> 613293, GenBank NM\_001017995.2) contains 13 exons and is localized on chromosome 5q35.1. This gene encodes an adaptor protein involved in the formation of actin-rich membrane protrusions called podosomes and invadopodia, which play a role in cell migration, and extracellular matrix (ECM) degradation (Buschman et al., 2009). It has been shown in mice models that *SH3PXD2B* protein plays a role in postnatal growth and development of the craniofacial structures, the ocular iridocorneal angle, and white adipose tissue (Mao et al., 2009). Moreover, Iqbal et al. demonstrated that *SH3PXD2B* mutant mice created via a gene-trap strategy presented craniofacial developmental defects, skeletal dysplasia, glaucoma and cardiac abnormalities similar to those observed in patients with FTFS (Iqbal et al., 2010).

To date, among the 40 patients reported in the literature with a clinical diagnosis of FTFS, only 20 have been shown to carry mutations in *SH3PXD2B*. Here we report the case of a female patient who presents particularly severe ophthalmic features, hypotonia with feeding difficulties and unusual lethal respiratory failure. By combining our results with a review of all mutated patients reported to date in the literature, we aim to further describe the clinical heterogeneity of this rare disease. We also hypothesize that the alteration of the podosomes function could lead to a reduction of the ECM degradation and accumulation of the latter in the extracellular space, which might explain the coarsening of the facial features and the severe refractory glaucoma.

## 2. Clinical report

The patient, a female, is the first-born child of healthy Armenian parents who were not consanguineous but came from the same small ethnic group. The familial history noticed that the patient's mother previously experienced a miscarriage at 16 weeks of gestation (WG): the autopsy showed that this male fetus had tetralogy of Fallot, normal karyotype (46, XY) and normal 22q11.2 FISH analysis. The patient had also a younger sister who died at the age of 27 days because of non-ketotic hyperglycinemia due to a homozygous mutation in the *GLDC* gene.

Concerning the prenatal period, ultrasounds examination revealed unilateral kidney agenesis and unexplained severe intrauterine growth retardation. She was delivered by caesarean section, at 37 WG, due to fetal heart rate anomalies. Her birth weight was 2490 g (−1,6 SD), her length was 43.5 cm (−2,8 SD), her occipitofrontal head circumference (OFC) was 32.5 cm (−1,5 SD) and Apgar score was 9, 9 and 10 after 1, 5 and 10 min, respectively.

Neonatal clinical examination revealed global hypotonia with dysmorphic features including brachycephaly, a wide anterior fontanelle, prominent forehead, hypertelorism, prominent subocular folds, bilateral buphthalmos with megalocornea, downward slanting palpebral fissures, midface hypoplasia, flat nasal bridge, anteverted nostrils, thin upper lip, long philtrum, downturned corners of the mouth, micro-retrognathia, ears abnormalities (protruding, large, low-set and dysplastic), and excess of nuchal skin (Fig. 1a, Fig. 1b, Fig. 1c and 2a). She also had skeletal anomalies including a caudal appendix (Fig. 1d), dorsolombar kyphoscoliosis (Fig. 2b), malposition of some toes with bilateral hallux valgus (Figs. 1e and 2c) and bilateral camptodactyly of the second, the third and the fourth finger, with short hands (Figs. 1f and 2d). She also presented a systolic murmur which led to the

diagnosis of a wide ventricular septal defect (6 mm), a patent foramen ovale and a persistent ductus arteriosus, with normal valves. Abdominal ultrasound found a solitary pelvic right kidney without parenchymal anomalies. Hip ultrasound revealed a left hip dysplasia. Hearing assessment was normal. Brain MRI showed a slight thinning of the corpus callosum (with normal morphology and topography) and a mega cisterna magna (Fig. 2e). Ophthalmological assessment confirmed the megalocornea and revealed severe bilateral congenital glaucoma. The patient had been referred to a neonatal care unit because of hypotonia and feeding difficulties in addition to the congenital malformations. The first bilateral trabeculectomy was performed at the age of two days. Subsequent controls showed persistent buphthalmos with a corneal edema despite surgical and local treatment.

Clinical evolution was marked by the coarsening of the facial features, with accentuation of the prominent forehead, arched and thick eyebrows, trichomegaly, broad mouth and full cheeks (Fig. 3a and Fig. 3b). Camptodactyly was treated with physiotherapy. At last examination, at the age of 2 months, her weight was 3140 g (−3,6 SD), her height was 53 cm (−2 SD) and her OFC was 35 cm (−2,5 SD). She had early and severe axial hypotonia and poor eye contact. Due to the persistence of ocular hypertonia (intraocular pressure (IOP): 25–30 mmHg) with appearance of a central corneal opacity on the right eye, a second unilateral trabeculectomy with peripheral iridectomy was performed at the age of one month and 11 days. The IOP in the left eye was between 15 and 20 mmHg and did not require a new surgery at that time. The patient experienced a third bilateral trabeculectomy at 3 months and 21 days. Despite these surgeries, buphthalmos and corneal edema worsening was observed (Fig. 3b) due to a persistent high intraocular pressure leading to a severe visual prognosis. Feeding difficulties, including swallowing difficulties, and important gastro-oesophageal reflux, appeared gradually and required nasogastric tube feeding, followed by a Nissen fundoplication and gastrostomy at the age of 3,5 months. At the age of two months, she presented sleep apnea and multiple episodes of desaturation requiring oxygen therapy. She was transferred at the age of 2,5 months to intensive care unit due to acute respiratory failure with severe hypercapnia that progressively led to a non-invasive ventilation with oxygen-dependence. Viral infection was suspected but not confirmed. Bronchoscopy with bronchoalveolar lavage showed a major tracheobronchial congestion and an inflammatory aspect compatible with lung disease caused by chronic aspiration of saliva. Chest tomodensitometry did not find underlying parenchymal, upper airways or vascular anomalies. The cardiac examination did not reveal any anomaly. Severe axial hypotonia with worsening of the scoliosis might have participated to the chronic respiratory failure, which led to death at the age of 4 months and 20 days. Autopsy was not performed. Frank-ter Haar syndrome was suspected on the clinical phenotype.

## 3. Genetic analysis and diagnosis

After obtaining written consent of the family, peripheral blood was collected from the patient and her parents. Genomic DNA was extracted and pangenomic analysis was first performed in the patient sample using the array Comparative Genomic Hybridization method (Human genome CGH Microarray, 180K Agilent<sup>®</sup>), according to the manufacturer's instructions. The results revealed no pathogenic imbalance (arr (1-22,X)x2). Then, standard polymerase chain reaction (PCR) amplification and Sanger sequencing of the entire *SH3PXD2B* gene were performed using standard methods. Obtained sequences were aligned with the reference sequence of the *SH3PXD2B* gene and Sanger sequencing results revealed a homozygous variant in the exon 10 of *SH3PXD2B* (NM\_001017995.2): c.969delG; p.(Arg324Glyfs\*19) in the patient. Both parents carry this variant in a heterozygous state (Fig. 4). This mutation is a one base pair deletion leading to a frameshift with the occurrence of a premature termination STOP codon 19 amino acids downstream. The variant has already been submitted to ClinVar



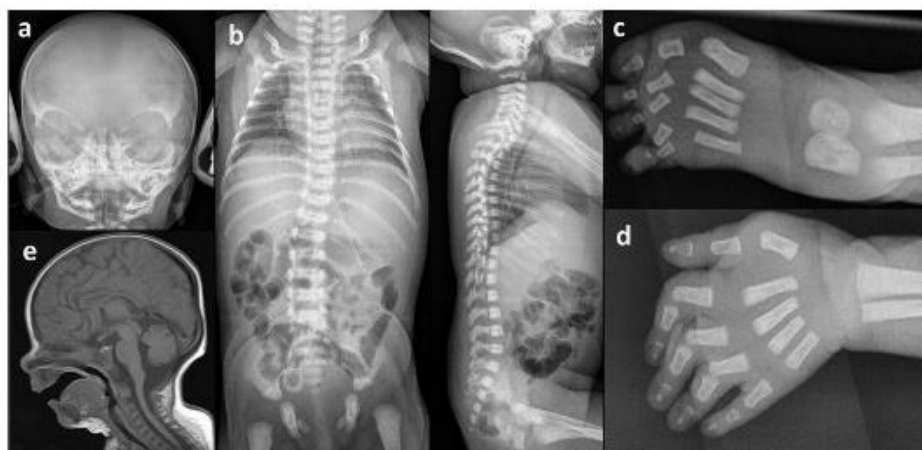
**Fig. 1. Morphological features of the patient at birth.** (a–b) Frontal and lateral view of patient's face at birth showing the dysmorphic features including hypotonic face, brachycephaly, prominent forehead, hypertelorism, with prominent subocular folds, buphthalmos with megalocornea, downward slanting palpebral fissures, midface hypoplasia, flat nasal bridge, anteverted nostrils, thin upper lip, long philtrum, downturned corners of the mouth, microretrognathia, ears abnormalities which are protruding, large, low-set and dysplastic. (c) Excess of nuchal skin. (d) Caudal appendix. (e) Malposition of some toes (high implantation of the 4th toe with hallux valgus). (f) Camptodactyly of the second, the third and the fourth finger, and short hands.

Database by OMIM (<http://www.ncbi.nlm.nih.gov/clinvar/>, accession number: SCV00020355). We also submitted our patient's variant to ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>, accession number: SCV000999009). Iqbal et al. first reported this mutation in two families (a Turkish family initially described by Maas et al. in 2004 (Maas et al., 2004) and an Arabic family) (Iqbal et al., 2010). These results confirmed the clinical diagnosis of FTGS.

#### 4. Discussion

In 2010, Iqbal et al. identified by homozygosity mapping studies the

*SH3PXD2B* gene on chromosome 5q35.1 which encodes a 911 amino-acid protein characterized by four Src homology 3 (SH3) domains and a Phox homology (PX) domain, called TKS4 protein (tyrosine kinase substrate with four SH3 domains) (Iqbal et al., 2010). TKS4 is an adaptor protein required for functional invadopodia and podosomes formation (Buschman et al., 2009; Courtneidge, 2012). TKS4 interacts with matrix metalloproteinase ADAM15 (Mao et al., 2009) and recruits to podosomes the MT1-MMP (membrane-type 1 matrix metalloproteinase, also known as MMP14) allowing extracellular matrix degradation (Buschman et al., 2009). Podosomes and invadopodia are actin-rich membrane protrusions involved in cell adhesion, migration and



**Fig. 2. Imaging exams.** (a) Frontal skull radiography showing wide anterior fontanelle, hypertelorism and micrognathia. (b–d) X-ray showing moderate dorso-lumbar kyphoscoliosis and left hip dysplasia (b), toes deformities with hallux valgus (c), flexion deformity of several fingers (d), (e) Brain MRI showing a thin corpus callosum and a mega cisterna magna.



Fig. 3. Evolution of the facial features with time. (a) Left to right: frontal view showing the coarsening of the facial features with time, from birth to 3.5 months of age. (b) Worsening of the buphthalmos (left) and trichomegaly (right).

extracellular matrix (ECM) remodelling. These structures are necessary for the adhesion and migration of a variety of cell types, including macrophages, lymphocytes, dendritic cells, osteoclasts and endothelial cells (Gimona et al., 2008).

To date, a total of 40 patients from 27 families with a clinical diagnosis of Frank-ter Haar syndrome have been reported in the literature including 20 patients from 13 families with identified mutations in *SH3PXD2B* (19 patients with homozygous mutations and 1 patient with compound heterozygous mutation) (Bendon et al., 2012; Borrone et al., 1993; Chang et al., 2017; Frank et al., 1973; Hamel et al., 1995; Iqbal et al., 2010; Maas et al., 2004; Mégarbané et al., 1997; ter Haar et al., 1982; Wilson et al., 2014; Zemojtel et al., 2014; Zrhidri et al., 2017). The 20 remaining patients did not carry any *SH3PXD2B* mutation or were not tested. The mutations reported include 9 intragenic variations (1 missense variant, 2 splice variants (of which 1 splice variant

predicted to lead to a frameshift) and 6 variants leading to premature termination STOP codon (Chang et al., 2017; Iqbal et al., 2010; Wilson et al., 2014; Zemojtel et al., 2014; Zrhidri et al., 2017), two homozygous deletion of exon 13 (Bendon et al., 2012; Wilson et al., 2014) and one entire gene deletion, that also includes the *UBTD2* gene (Iqbal et al., 2010) (Fig. 5). The molecular and clinical data of the 20 reported patients and the present case are summarized in Table 1.

Frank-ter Haar syndrome seems to have genetic heterogeneity since Iqbal et al. found 5 different homozygous mutations in 7 unrelated families but failed to identify *SH3PXD2B* mutation in 6 other families. The authors suggested that it is possible that mutations in *cis*-regulatory elements of *SH3PXD2B* exist in these families, but have not been highlighted (Iqbal et al., 2010). Furthermore, the same authors also showed other homozygosity regions in patients without *SH3PXD2B* mutation in favour of the involvement of multiple other loci. They

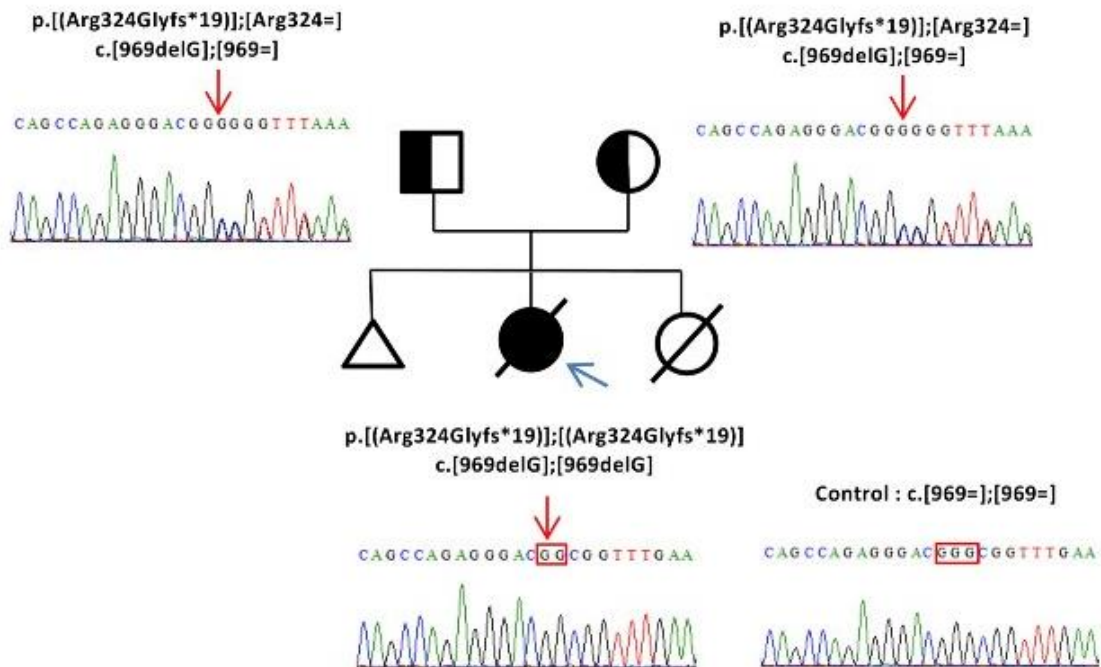
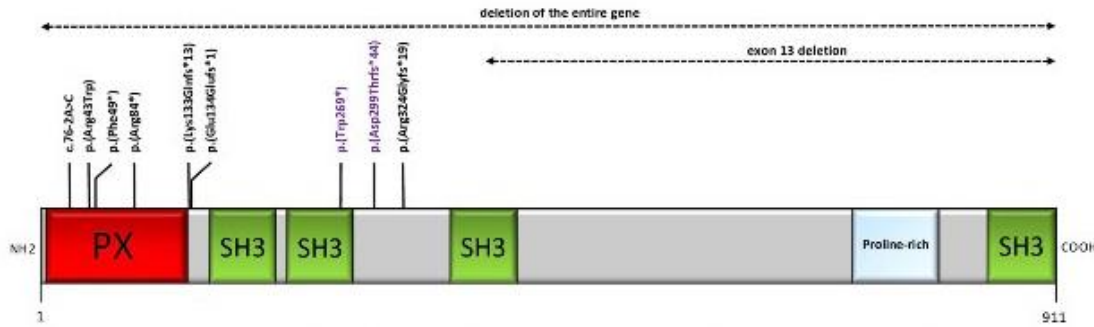


Fig. 4. Pedigree and Sanger sequencing results of the *SH3PXD2B* gene. The full circle represents the affected female patient.



**Fig. 5.** Schematic plan of the SH3PXD2B protein (TKS4) with location of all mutations causing FTSH identified to date in the literature. The compound heterozygous mutations appear in purple.

PX indicates the Phox homology domain and SH3 indicates the Src homology 3 domain. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article).

hypothesized that individuals with FTSH without *SH3PXD2B* mutation probably have mutations in genes that interact with TKS4 and/or are involved in podosomes formation or function (Iqbal et al., 2010). This assumption is consistent with the fact that FTSH was first considered as an autosomal recessive form of Melnick-Needles syndrome (MNS) (ter Haar et al., 1982), a rare disease caused by gain-of-function mutations in *FLNA* gene (Xq28) that shares some clinical features with the FTSH including facial gestalt (prominent forehead, proptosis, full cheeks, microretrognathia, abnormal ears) and skeletal anomalies (scoliosis joint subluxations and bowing of long bones) (Robertson et al., 2003). In addition and interestingly, the filamin A protein is present in the podosome belt and its cleavage by calpain is necessary for osteoclast motility during bone development (Marzia et al., 2006). Filamin A is also required for podosome rosette formation, proteolysis of the extracellular matrix mediated by podosomes in macrophages, and three-dimensional mesenchymal cells migration (Guiet et al., 2012). Additionally, Cejudo-Martin and Courtneidge reported that the syndromes due to mutations in genes encoding podosomal proteins are characterized by serious developmental defects involving the craniofacial area, skeleton and heart (Cejudo-Martin and Courtneidge, 2011). Therefore, the clinical similarities of FTSH and MNS could be partially explained by a common pathophysiological basis. Furthermore, recently, De Vos et al. identified a homozygous missense mutation in *MMP14* (matrix metalloprotease 14), leading to the diagnosis of Winchester syndrome, in two brothers originally described with BDCSS and no mutation in *SH3PXD2B* (de Vos et al., 2018). Winchester Syndrome (WS, MIM# 277950) is a rare autosomal recessive skeletal dysplasia characterized by progressive joint destruction and osteolysis. This syndrome is part of the “vanishing bone” syndromes, a group of skeletal disorders characterized by excessive bone resorption which also includes Multicentric osteolysis nodulosis and arthropathy syndrome (MONA, MIM# 259600) caused by homozygous mutation in *MMP2* (matrix metalloprotease 2). All the protein products of *SH3PXD2B*, *MMP14* and *MMP2* cooperate and interact in collagen remodelling. Given these observations, the clinical overlap and common physiopathology, De Vos et al. suggest to gather these syndromes (FTSH, WS and MONA) in a new entity: the “defective collagen-remodelling spectrum (DECORS)” (de Vos et al., 2019). Therefore, these observations, and the similarity between these three syndromes, make us believe that patients with apparently clinical diagnosis of FTSH might have defects in *MMP14* or *MMP2* genes.

We report the particularly severe phenotype of a female patient including severe congenital glaucoma (three trabeculectomy failures), buphthalmos worsening, major feeding troubles, and an early death due to respiratory failure without any infection at 4,5 months. At our knowledge the female patient we report is the one that had the shorter survival. To date, among the 20 patients with *SH3PXD2B* identified

mutations, 6 patients (30%) died early (before 25 years), because of cardiorespiratory failure (7 years old and 8 years old) (Hamel et al., 1995; Wilson et al., 2014), heart failure (12 months old and 24 years old) (Borrone et al., 1993; ter Haar et al., 1982) or not reported causes (5,5 months and not reported) (Iqbal et al., 2010; Zemojtel et al., 2014) (Table 1). Therefore, FTSH seems to be clinically heterogeneous, although most of the reported patients were young at the time of the description and we have no information about their evolution. It cannot be excluded that they died during childhood or later. Only two adult patients have been reported to date, one of which died at the age of 24 years, and the other was described at the age of 36 years. In addition, our patient had an early and severe axial hypotonia with a delayed motor development and poor eye contact whereas in the literature, most of the patients seem to have a more favourable course: 8/17 patients (47%) have a mild motor delay and among the 12 patients in whom neuropsychological assessment was performed, 4 have moderate learning disabilities or a mildly delayed cognitive development (Table 1). Interestingly, 3 patients in the literature also have unspecific brain abnormalities (Table 1). However, it was difficult to conclude that brain structural anomalies could be part of the phenotype since only 4 patients, including our patient, had brain imaging and since the reported brain abnormalities were heterogeneous, unspecific and different between the patients. In addition, brain abnormalities do not seem to be correlated with the severity of the developmental delay, but too few patients have been assessed regarding this part of the phenotype. Regarding the ophthalmological features, we noted that 5/17 reported patients (29%) had congenital glaucoma, and 2 of them required surgery (once for one patient and twice for the other) (Mégarbané et al., 1997) (Table 1). Unfortunately, the data about the ophthalmological evolution was lacking for these 5 cases. A postnatal growth retardation has been reported in 3/5 patients (60%) and is also very frequent in *Sh3pxd2b* null mice (Iqbal et al., 2010). However, we cannot exclude that the very severe phenotype of our patient may be partly due to a second rare disease. Indeed, the patient's younger sister died at the age of 27 days because of nonketotic hyperglycinemia, a rare metabolic disorder due to a homozygous mutation in the *GLDC* gene, found in each parent at a heterozygous state. Our patient has been tested and she carried this mutation in a heterozygous state. This second recessive disease in the family suggests the existence of a consanguinity which was not known to the parents.

The genotype-phenotype correlation is difficult to establish because of the too low number of identified mutated patients. In addition, among the three other patients sharing the same homozygous c.969delG mutation, two had a favourable clinical course conversely to our patient (Iqbal et al., 2010; Maas et al., 2004). These patients and our patient all share similar craniofacial features, variable skeletal (2/4 kyphosis, 2/4 caudal appendix, 2/4 camptodactyly, 1/4 bowing of long

**Table 1**  
Phenotype of the confirmed cases of Frank-ter Haar syndrome with known SH3PXD2B mutations.

		Iqbal et al. (2010)						
Family	13	12	2	1	6	9	7	
Patient	1	1	2	1	1	2	1	
Gender	M	M	M	F	F	M	M	
Age	ND	5 y 6 m	1 y	> 1 y	6 y 4 m	3 y 6 m	NR	
Consanguinity	Yes	Yes	Yes	No	Yes	Yes	Yes	
Premature death (age; cause)	-	7 y; CR failure	1 y; heart failure	-	-	-	unknown cause and age	
<b>Facial features</b>								
Prominent forehead	+	+	+	+	+	+	+	
Brachycephaly	+	+	+	+	+	+	+	
Wide anterior fontanel	+	+	+	+	+	+	+	
Proptosis	+	+	+	+	+	+	+	
Hypertelorism	+	+	+	+	+	+	+	
Full cheeks	+	+	+	+	+	+	+	
Anteverted nostrils	+	+	+	+	+	+	+	
Long philtrum	ND	ND	+	+	+	+	+	
Thin upper lip	ND	ND	+	+	+	+	+	
Broad mouth	+	+	+	+	+	+	+	
Dental anomalies	ND	ND	ND	ND	ND	ND	ND	
Gingival hyperplasia	ND	ND	ND	ND	ND	ND	ND	
Micrognathia	+	+	+	+	+	+	+	
Low-set ears	ND	ND	+	+	+	+	+	
Large protruding ears	+	+	+	+	+	+	+	
Coarse features	ND	ND	+	+	+	+	+	
<b>Skeletal anomalies</b>								
Kyphoscoliosis	+	+	+	+	+	+	+	
Pectus excavatum	ND	ND	ND	ND	+	+	+	
Caudal appendix	+	+	+	+	+	+	+	
Bowing of long bones	+	+	+	+	+	+	+	
Short hands	+	+	+	+	+	+	+	
Campodactyly	+	+	+	+	+	+	+	
Toes anomalies	ND	ND	+	+	+	+	+	
Clubfeet	+	+	+	+	+	+	+	
Other	-	-	-	-	-	-	-	
Pes valgus, dislocated knees, hips, elbows								
OP, joints hyperlaxity, fetal pads								
Wormian bones								
<b>Growth parameters (postnatal period)</b>								
Height	ND	ND	ND	ND	ND	ND	ND	
Weight	ND	-1.3SD	ND	ND	-3SD	ND	-2SD	
OPC	ND	-1.3SD	0.5D	ND	-1.3SD	ND	-2SD	
Ocular anomalies	-	-1.3SD	-	ND	-1.5D	ND	-1.5SD	
Congenital glaucoma	-	+	ND	+	-	-	+	

Table 1 (continued)

	Ighod et al. (2010)										
Megalocornea	+	+	+	+	+	+	+	+	+	+	+
Bilateral buphthalmos	ND	ND	+	+	+	+	+	+	+	+	+
Other	RD	Cloudy cornea									
<b>Cardiac anomalies</b>											
Valves anomalies	-	MVP, AVP	MVP	ND	ND	MVR (0)	-	-	-	ND	ND
VSD	+	+	ND	ND	ND	+	-	-	+	-	+
Other	ASD, PDA	Cardiomegaly	DORV	DORV	DORV						
<b>Developmental delay</b>											
Motor retardation	+	ND	(G)	+	+	+	+	+	+	+	+
Intellectual disability	ND	ND	-	ND	ND	+	ND	ND	ND	ND	+
Learning difficulties	ND	ND	-	ND	ND	+	ND	ND	ND	ND	NA
Brain imaging	ND	ND	ND	ND	ND	-	ND	ND	ND	Small pituitary gland, atrophy of chiasma/optic tracts	NP
<b>Metabolic analysis</b>											
Metabolic analysis	ND	ND	nor.	nor.	nor.	nor.	ND	ND	ND	nor.	NP
<b>Additional clinical features</b>											
Additional clinical features	ND	2 café au lait spots									Bilat. inguinal hernia, loose skin, gall stones, hirsutism
<b>SH3BPXD2B mutations</b>											
SH3BPXD2B mutations	Homozygous c.76-2A > C	Homozygous c.127C > T, p.(Arg43Trp)	Homozygous c.147insT, p.(Phe49*)	Homozygous c.969delG, p.(Arg324Glyfs*19)							Homozygous 5q35.1 deletion
<b>Family</b>											
Family	Zemajtel et al. (2014)	Chang et al. (2017)	Zhidars et al. (2017)	Bendon et al. (2012)	Wilson et al. (2014)						Our patient
Patient	4	1	1	1	3	3	3	3	1	1	14
Gender	F	F	F	F	M	M	M	M	M	M	F
Age	5.5 m	3 y	2 y	2 y	13 y 10 m	7 y 7 m	7 y 1 m	36 y	23 y	8 y	4.5 m
Consanguinity	Yes	No	No	Yes	Yes	Yes	Yes	No	No	Yes	No
Premature death (age; cause)	5.5 m; ND	-	-	-	-	-	-	-	24 y; heart failure	8 y; CR failure	4.5 m; respiratory failure
<b>Facial features</b>											
Facial features	+	+	+	+	+	+	+	+	+	+	+
Prominent forehead	+	ND	+	+	+	+	+	+	+	+	+
Brachycephaly	+	+	+	+	+	+	+	+	+	+	+
Wide anterior fontanel	+	+	+	+	+	+	+	+	+	+	+

Table 1 (continued)

	Zemolčić et al. (2014)	Chang et al. (2017)	Zhidari et al. (2017)	Bendon et al. (2012)	Wilson et al. (2014)	Our patient	Total
<b>Proptosis</b>	ND	+	+	+	ND	+	18/18
<b>Hypertelorism</b>	+	+	+	+	+	+	21/21
<b>Full cheeks</b>	+	+	+	+	+	+	21/21
<b>Anteverted nostrils</b>	ND	+	-	+	+	+	14/20
<b>Long philtrum</b>	ND	+	ND	-	+	+	13/17
<b>Thin upper lip</b>	ND	ND	ND	-	+	+	8/15
<b>Broad mouth</b>	ND	ND	+	+	+	+	18/18
<b>Dental anomalies</b>	ND	+	ND	+	ND	ND	7/9
<b>Gingival hyperplasia</b>	ND	+	ND	+	+	ND	11/11
<b>Micrognathia</b>	+	ND	+	+	+	+	16/19
<b>Low-set ears</b>	ND	+	ND	+	+	+	13/16
<b>Large protruding ears</b>	ND	ND	+	+	ND	+	13/15
<b>Coarse features</b>	+	+	ND	-	+	+	14/17
<b>Skeletal anomalies</b>							
<b>Kyphoscoliosis</b>	ND	+	+	-	+	+	13/18
<b>Pectus excavatum</b>	ND	+	ND	ND	ND	+	5/8
<b>Caudal appendix</b>	+	+	+	-	-	+	14/21
<b>Bowing of long bones</b>	ND	ND	+	ND	-	+	11/15
<b>Short hands</b>	ND	+	+	+	+	+	20/20
<b>Camptodactyly</b>	+	+	+	+	+	+	14/21
<b>Toes anomalies</b>	ND	ND	ND	+	ND	+	8/8
<b>Clubfeet</b>	+	-	+	+	ND	+	11/19
<b>Other</b>		Osteolysis, GV, metatarsus adductus, thick joints	Osteolysis, GV, metatarsus adductus, thick joints	CS (d), size discrepancy in feet	Osteolysis, GR	Osteolysis, GV, radial head disloc.	Left hip dysplasia
<b>Growth parameters (postnatal period)</b>							
<b>Height</b>	ND	ND	ND	ND	ND	ND	-2SD
<b>Weight</b>	ND	ND	ND	ND	ND	ND	-3.6SD
<b>OFC</b>	< -2SD	ND	2SD	> 2SD	ND	ND	-2.5SD
<b>Ocular anomalies</b>							
<b>Congenital glaucoma</b>	+	-	-	-	ND	+	6/18
<b>Megalocornea</b>	ND	+	+	-	ND	ND	14/18
<b>Bilateral buphthalmos</b>	ND	+	+	-	ND	+	8/14
<b>Other</b>	Sclerocornea, RD	Unilateral iridoretinal coloboma (g)	Myelinated nerve fiber layer (g)	ERG: reduced cortico-retinal conduction			
<b>Cardiac anomalies</b>							
<b>Valves anomalies</b>	-	MVP	-	MVP, MVR, AVR	MVP	MVP	10/17
<b>VSD</b>	+	ASD, TVP, dilated aortic root, PAH	Hypoplastic left ventricle, dextroposition of aorta, DORV	TVR	-	-	8/18
<b>Other</b>	SVAS, PLSVC		Small shunt left-right				Patent foramen ovale
<b>Developmental delay</b>							



Table 1 (continued)

	Zemojzel et al. (2014)	Chang et al. (2017)	Zhidis et al. (2017)	Bendon et al. (2012)	Wilson et al. (2014)	Our patient	Total
<b>Motor retardation</b>	ND	+	-	-	-	-	9/18
<b>Intellectual disability</b>	NA	-	-	-	-	+	1/12
<b>Learning difficulties</b>	NA	ND	ND	+	-	NA	4/8
<b>Brain imaging</b>	Delay CNS myelination	-	ND	-	Hypoplasia of inferior vermis	ND	ND
<b>Metabolic analysis</b>	ND	ND	ND	ND	nor.	nor.	NP
<b>Additional clinical features</b>	Cholestasis	Umbilical hernia	Idiopathic intracranial hypertension	Bilat. legs lymphedema, subcutaneous nodules	Recurrent urinary tract infections subcutaneous nodules	Acne, inguinal hernia	Acne
<b>SH3PXD2B mutations</b>	Homozygous c.250C > T, p. (Arg84*)	Homozygous c.396dup, p. (Lys133Glnfs*13)	C heterozygous c.806G > A, p. (Trp269*); c.892delC p. (Asp299Thrfs*44)	Homozygous exon 13 deletion	Homozygous c.401 + 1G > A, p. (Glu134Gluafs*1)	Homozygous c.969delG, p. (Arg324Glyfs*19)	Homozygous c.969delG, p. (Arg324Glyfs*19)

Abbreviations: + feature is present; - feature is absent.

ASD, atrial septal defect; AVP, aortic valve prolapse; AVR, aortic valve regurgitation; bilat., bilateral; C, heterozygous, compound heterozygous; CNS, central nervous system; CR, cardiorespiratory; CS, craniosynostosis; d, days; disloc., dislocation; DORV, double-outlet right ventricle; ERG, electroretinogram; GR, genu recurvatum; GV, genu valgum; m, months; MCP, metacarpophalangeal; MVR, mitral valve regurgitation; MVP, mitral valve prolapse; NA, not applicable; ND, not determined; nor., normal; NP, not performed; OFC, occipitofrontal circumference; OP, osteopenia; PAH, pulmonary arterial hypertension; PDA, patent ductus arteriosus; PLSVC, persistent left superior vena cava; RD, retinal detachment; SD, standard deviation; SVAS, supravalvular aortic stenosis; TVP, tricuspid valve prolapse; TVR, tricuspid valve regurgitation; VSD, ventricular septal defect; y, years.

a: teeth hypoplasia; b: class III malocclusion; c: short toes; d: sagittal suture synostosis; e: surgery for bilateral glaucoma at 2 days and 3 months; f: surgery for glaucoma at 10 days of age; g: tortuous vessels, amblyopia, anisometropia; h: artificial mitral valve implanted at 23 months; i: muscle weakness (normal electromyography); j: generalized hypotonia.

Families from Iqbal et al., 2010; Family 2 was described by ter Haar et al., 1982 and Hamel et al., 1995; family 6 was described by Maas et al., 2004 and family 7 was described by Mègarbané et al., 1997. Family 3 from Wilson et al. (2014), was described by Borrone et al. (1993).

<sup>a</sup> Elevated intraocular pressure (IOP) in patient 1 was attributed to systemic steroid response, and elevated IOP could not be ruled out in patient 2 since it could have been masked by the chronic systemic acetazolamide therapy for an idiopathic intracranial hypertension.

<sup>b</sup> Microdeletion including SH3PXD2B and UBTD2 genes.

bones) and cardiac anomalies (1/4 mitral valve anomaly, 3/4 ventricular septal defect). Concerning ocular anomalies, only 1/4 of the patients have congenital glaucoma, but they all have a megalocornea (Table 1). However, it is interesting to note that the three siblings with homozygous deletion of exon 13 described by Bendon et al. have no ocular anomalies and they have a distinct phenotype including (in addition to the classically FTHS associated signs): lymphoedema, class III malocclusion, and sagittal synostosis (Bendon et al., 2012). Conversely, these features were not present in the two brothers described by Borrone et al., who also carry a deletion of exon 13 (Borrone et al., 1993; Wilson et al., 2014). Nevertheless, Bendon et al. conclude that this particular phenotype could involve the two other regions of shared homozygosity identified on chromosomes 8 and 12 in this family (Bendon et al., 2012).

Frank-ter Haar syndrome includes evolving clinical features which previously led to the comparison to a metabolic storage disease (Maas et al., 2004), in particular coarsening of the facial features with time, tendency to hirsutism, skeletal and cardiac valve anomalies, gingival hyperplasia, inguinal hernia and motor retardation with hypotonia. For this reason, metabolic investigations were performed on several patients of the literature but were all normal (Maas et al., 2004; Mégarbané et al., 1997; ter Haar et al., 1982). For instance, Wilson et al. suspected lysosomal storage disease in their patient with *SH3PXD2B* mutation and gingival hyperplasia. The histological analysis of the gingiva showed that the fibroblasts had accumulated cytoplasmic membrane-bound structures containing collagen fibers but *in vitro* analysis did not reveal any abnormal collagen synthesis or secretion, and ultrastructural skin study, biochemical and enzymatic analysis excluded the diagnosis of metabolic disorder (Wilson et al., 2014). Therefore, there is no current evidence of metabolic abnormalities in FTHS. However, based on the Tks4 protein function (involved in matrix metalloproteinase recruitment and ECM degradation), we can hypothesize that podosome alteration due to *SH3PXD2B* mutations may cause impairment in degradation of the ECM and accumulation of its components into the extracellular space. This could explain the normality of the collagen synthesis and secretion analyses, and the observed accumulation of collagen fibers (Wilson et al., 2014). In the same way, this physiopathological hypothesis of accumulation of ECM components could partly explain the origin and the progressive worsening of the glaucoma in FTHS patients, in particular in our patient who experienced 3 trabeculotomy failures with striking aggravation of the buphthalmos. This is corroborated by experimental data from mouse models, since mutant mice bearing homozygous *Sh3pxd2b* *nee* mutation exhibit anterior segment dysgenesis that causes early-onset glaucoma with progression over time. Given these observations, the authors concluded that podosomes through their role in ECM remodelling, are implicated in the normal development of the iridocorneal angle and that mutation in *SH3PXD2B* contribute to the development of glaucoma (Mao et al., 2011). Interestingly, podosomes have also been observed in cells of the trabecular meshwork in the eye (Aga et al., 2008), and it has been shown that the ECM has a well-established role in regulating the function of the mature trabecular meshwork (Fuchshofer and Tamm, 2009). Moreover, other genes associated with the ECM remodelling are involved in anterior segment dysgenesis, like ADAMs and members of the TGF $\beta$  superfamily (Mao et al., 2011). Recently, De Vos et al. suggests that FTHS, WS and MONA share common phenotypical features and belong to a same single clinical spectrum ("defective collagen-remodelling spectrum (DECORS)") and emphasize the central role of impaired collagen remodelling in the three disorders (de Vos et al., 2019).

In conclusion, we report a very severe neurodevelopmental phenotype of FTHS with lethal respiratory impairment, confirming the clinical heterogeneity of this rare condition. We were not able to determine a genotype-phenotype correlation since very few patients are described in the literature and since 2 out of the 3 other patients carrying the same mutation had a favourable clinical course. It could be

hypothesized that the evolving features including the cardiac valve anomalies, the gingival hyperplasia, the coarsening of the facial traits and the aggravation of the glaucoma in some patients, may be due to an accumulation of ECM components in the extracellular space secondary to the alteration of the podosomes function. However, more cases are needed to better characterize the phenotype and understand the natural history of the disease.

#### CRedit authorship contribution statement

**Benjamin Durand:** Conceptualization, Methodology, Investigation, Writing - original draft, Writing - review & editing. **Corinne Stoetzel:** Investigation, Validation. **Elise Schaefer:** Resources. **Nadège Calmels:** Investigation. **Sophie Scheidecker:** Investigation. **Nadine Kempf:** Investigation. **Charlie De Melo:** Resources. **Anne-Sophie Guilbert:** Resources. **Dana Timbolschi:** Resources. **Leonardo Donato:** Resources. **Dominique Astruc:** Resources. **Arnaud Sauer:** Resources. **Maria Cristina Antal:** Resources. **Hélène Dollfus:** Supervision, Project administration. **Salima El Chehadeh:** Supervision, Conceptualization, Methodology, Writing - review & editing.

#### Declaration of competing interest

All authors declare that they have no conflict of interest.

#### Acknowledgments

The authors thank the parents of the patient for their participation in this study.

The authors thank the Laboratory of medical genetics U1112 for its technical support.

#### References

- Aga, M., Bradley, J.M., Keller, K.E., Kelley, M.J., Acott, T.S., 2008. Specialized podosome- or invadopodia-like structures (PILS) for focal trabecular meshwork extracellular matrix turnover. *Invest. Ophthalmol. Vis. Sci.* 49, 5353–5365. <https://doi.org/10.1167/iov.07-1666>.
- Bendon, C.L., Fenwick, A.L., Hurst, J.A., Nürnberg, G., Nürnberg, P., Wall, S.A., Wilkie, A.O.M., Johnson, D., 2012. Frank-ter Haar syndrome associated with sagittal craniosynostosis and raised intracranial pressure. *BMC Med. Genet.* 13, 104. <https://doi.org/10.1186/1471-2350-13-104>.
- Borrone, C., Di Rocco, M., Crovato, F., Camera, G., Gambini, C., 1993. New multisystemic disorder involving heart valves, skin, bones, and joints in two brothers. *Am. J. Med. Genet.* 46, 228–234. <https://doi.org/10.1002/ajmg.1320460225>.
- Buschman, M.D., Bromann, P.A., Cejudo-Martin, P., Wen, F., Past, I., Courtneidge, S.A., Brugge, J., 2009. The novel adaptor protein Tks4 (*SH3PXD2B*) is required for functional podosome formation. *Mol. Biol. Cell* 20, 1302–1311. <https://doi.org/10.1091/mbc.08-09-0949>.
- Cejudo-Martin, P., Courtneidge, S.A., 2011. Podosomal proteins as causes of human syndromes: a role in craniofacial development? *Genes. N. Y. N* 49, 209–221. <https://doi.org/10.1002/dvg.20732>.
- Chang, T.C., Bauer, M., Puerta, H.S., Greenberg, M.B., Cavuoto, K.M., 2017. Ophthalmic findings in Frank-ter Haar syndrome: report of a sibling pair. *J. AAPOS Off. Publ. Am. Assoc. Pediatr. Ophthalmol. Strabismus* 21, 514–516. <https://doi.org/10.1016/j.jaapos.2017.07.216>.
- Courtneidge, S.A., 2012. Cell migration and invasion in human disease: the Tks adaptor proteins. *Biochem. Soc. Trans.* 40, 129–132. <https://doi.org/10.1042/BST20110685>.
- de Vos, L.J.H.M., Tao, E.Y., Ong, S.L.M., Goggi, J.L., Scerri, T., Wilson, G.R., Low, C.G.M., Wong, A.S.W., Grussu, D., Stegmann, A.P.A., van Geel, M., Janssen, R., Amor, D.J., Bahlo, M., Dunn, N.R., Carney, T.J., Lockhart, P.J., Coull, B.J., van Steensel, M.A.M., 2018. Functional analysis of a hypomorphic allele shows that MMP14 catalytic activity is the prime determinant of the Winchester syndrome phenotype. *Hum. Mol. Genet.* 27, 2775–2788. <https://doi.org/10.1093/hmg/ddy168>.
- de Vos, L.J.H.M., Wong, A.S.W., Welting, T.J.M., Coull, B.J., van Steensel, M.A.M., 2019. Multicentric osteolytic syndromes represent a phenotypic spectrum defined by defective collagen remodeling. *Am. J. Med. Genet. A* 179, 1652–1664. <https://doi.org/10.1002/ajmg.a.61264>.
- Frank, Y., Zipirkowski, M., Romano, A., Stein, R., Katznelson, M.B., Cohen, B., Goodman, R.M., 1973. Megalocornea associated with multiple skeletal anomalies: a new genetic syndrome? *J. Genet. Hum.* 21, 67–72.
- Fuchshofer, R., Tamm, E.R., 2009. Modulation of extracellular matrix turnover in the trabecular meshwork. *Exp. Eye Res.* 88, 683–688. <https://doi.org/10.1016/j.exer.2009.01.005>.
- Gimona, M., Buccione, R., Courtneidge, S.A., Linder, S., 2008. Assembly and biological

- role of podosomes and invadopodia. *Curr. Opin. Cell Biol.* 20, 235–241. <https://doi.org/10.1016/j.ccb.2008.01.005>.
- Griet, R., Vérolet, C., Lamsoul, I., Cougoule, C., Poincloux, R., Labrousse, A., Calderwood, D.A., Glogauer, M., Lutz, P.G., Maridonneau-Parini, I., 2012. Macrophage mesenchymal migration requires podosome stabilization by filamin A. *J. Biol. Chem.* 287, 13051–13062. <https://doi.org/10.1074/jbc.M111.307124>.
- Hamel, B.C., Draaisma, J.M., Pinckers, A.J., Boetes, C., Hoppe, R.L., Ropers, H.H., Brunner, H.G., 1995. Autosomal recessive Melnick-Needles syndrome or ter Haar syndrome? Report of a patient and reappraisal of an earlier report. *Am. J. Med. Genet.* 56, 312–316. <https://doi.org/10.1002/ajmg.1320560320>.
- Iqbal, Z., Cejudo-Martin, P., de Brouwer, A., van der Zwaag, B., Ruiz-Lozano, P., Scimia, M.C., Lindsey, J.D., Weinreb, R., Albrecht, B., Megarbane, A., Alanay, Y., Ben-Neriah, Z., Amenduni, M., Artuso, R., Veltman, J.A., van Beusekom, E., Oudakker, A., Millán, J.L., Hennekam, R., Hamel, B., Courtneidge, S.A., van Bokhoven, H., 2010. Disruption of the podosome adaptor protein TKS4 (SH3PXD2B) causes the skeletal dysplasia, eye, and cardiac abnormalities of frank-ter haar syndrome. *Am. J. Hum. Genet.* 86, 254–261. <https://doi.org/10.1016/j.ajhg.2010.01.009>.
- Maas, S.M., Kayserili, H., Lam, J., Apak, M.Y., Hennekam, R.C.M., 2004. Further delineation of frank-ter haar syndrome. *Am. J. Med. Genet. A* 131, 127–133. <https://doi.org/10.1002/ajmg.a.30244>.
- Mao, M., Hedberg-Buenz, A., Koehn, D., John, S.W.M., Anderson, M.G., 2011. Anterior segment dysgenesis and early-onset glaucoma in *nee* mice with mutation of *Sh3pxd2b*. *Invest. Ophthalmol. Vis. Sci.* 52, 2679–2688. <https://doi.org/10.1167/iovs.10.5993>.
- Mao, M., Thedens, D.R., Chang, B., Harris, B.S., Zheng, Q.Y., Johnson, K.R., Donahue, L.R., Anderson, M.G., 2009. The podosomal-adaptor protein SH3PXD2B is essential for normal postnatal development. *Mamm. Genome Off. J. Int. Mamm. Genome Soc.* 20, 462–475. <https://doi.org/10.1007/s00335-009-9210-9>.
- Marzia, M., Chiusaroli, R., Neff, L., Kim, N.-Y., Chishti, A.H., Baron, R., Horne, W.C., 2006. Calpain is required for normal osteoclast function and is down-regulated by calcitonin. *J. Biol. Chem.* 281, 9745–9754. <https://doi.org/10.1074/jbc.M513516200>.
- Mégarbané, A., Tomey, K., Wakim, G., 1997. Congenital glaucoma, limb deformities, skeletal dysplasia, and facial anomalies: report of another family. *Am. J. Med. Genet.* 73, 67–71.
- Robertson, S.P., Twigg, S.R.F., Sutherland-Smith, A.J., Biancalana, V., Gorlin, R.J., Horn, D., Kenwright, S.J., Kim, C.A., Morava, E., Newbury-Ecob, R., Orstavik, K.H., Quarrell, O.W.J., Schwartz, C.E., Shears, D.J., Suri, M., Kendrick-Jones, J., Wilkie, A.O.M., OPD-spectrum Disorders Clinical Collaborative Group, 2003. Localized mutations in the gene encoding the cytoskeletal protein filamin A cause diverse malformations in humans. *Nat. Genet.* 33, 487–491. <https://doi.org/10.1038/ng1119>.
- ter Haar, B., Hamel, B., Hendriks, J., de Jager, J., 1982. Melnick-Needles syndrome: indication for an autosomal recessive form. *Am. J. Med. Genet.* 13, 469–477. <https://doi.org/10.1002/ajmg.1320130418>.
- Wilson, G.R., Sunley, J., Smith, K.R., Pope, K., Bromhead, C.J., Fitzpatrick, E., Di Rocco, M., van Steensel, M., Coman, D.J., Leventer, R.J., Delatycki, M.B., Amor, D.J., Bahlo, M., Lockhart, P.J., 2014. Mutations in SH3PXD2B cause Borroni dermatocardi-skeletal syndrome. *Eur. J. Hum. Genet. EJHG* 22, 741–747. <https://doi.org/10.1038/ejhg.2013.229>.
- Zemajtel, T., Kohler, S., Mackenroth, L., Jager, M., Hecht, J., Krawitz, P., Graul-Neumann, L., Doelken, S., Ehmke, N., Spielmann, M., Oien, N.C., Schweiger, M.R., Kruger, U., Frommer, G., Fischer, B., Kornak, U., Flottmann, R., Ardeshirdavani, A., Moreau, Y., Lewis, S.E., Haendel, M., Smedley, D., Horn, D., Mundlos, S., Robinson, P.N., 2014. Effective diagnosis of genetic disease by computational phenotype analysis of the disease-associated genome. *Sci. Transl. Med.* 6. <https://doi.org/10.1126/scitranslmed.3009262>. 252ra123-252ra123.
- Zrhidri, A., Jaouad, I.C., Lyahyai, J., Raymond, L., Egéa, G., Taoudi, M., El Mouatassim, S., Sefiani, A., 2017. Identification of two novel SH3PXD2B gene mutations in Frank-Ter Haar syndrome by exome sequencing: case report and review of the literature. *Gene* 628, 190–193. <https://doi.org/10.1016/j.gene.2017.07.011>.

## Annexe 6 : La perte de fonction de PLAAT3 provoque un syndrome lipodystrophique et neurologique par altération de la signalisation PPAR $\gamma$

nature genetics

Article

<https://doi.org/10.1038/s41588-023-01535-3>


# Loss of phospholipase PLAAT3 causes a mixed lipodystrophic and neurological syndrome due to impaired PPAR $\gamma$ signaling

Received: 22 December 2022

A list of authors and their affiliations appears at the end of the paper

Accepted: 16 September 2023

Published online: 02 November 2023

 Check for updates

Phospholipase A/acyltransferase 3 (PLAAT3) is a phospholipid-modifying enzyme predominantly expressed in neural and white adipose tissue (WAT). It is a potential drug target for metabolic syndrome, as *Plaat3* deficiency in mice protects against diet-induced obesity. We identified seven patients from four unrelated consanguineous families, with homozygous loss-of-function variants in *PLAAT3*, who presented with a lipodystrophy syndrome with loss of fat varying from partial to generalized and associated with metabolic complications, as well as variable neurological features including demyelinating neuropathy and intellectual disability. Multi-omics analysis of mouse *Plaat3*<sup>-/-</sup> and patient-derived WAT showed enrichment of arachidonic acid-containing membrane phospholipids and a strong decrease in the signaling of peroxisome proliferator-activated receptor gamma (PPAR $\gamma$ ), the master regulator of adipocyte differentiation. Accordingly, CRISPR-Cas9-mediated *PLAAT3* inactivation in human adipose stem cells induced insulin resistance, altered adipocyte differentiation with decreased lipid droplet formation and reduced the expression of adipogenic and mature adipocyte markers, including PPAR $\gamma$ . These findings establish PLAAT3 deficiency as a hereditary lipodystrophy syndrome with neurological manifestations, caused by a PPAR $\gamma$ -dependent defect in WAT differentiation and function.

Human phospholipase A and acyltransferase 3 (PLAAT3), previously known as phospholipase A2, group XVI (PLA2G16), is part of a family of acyltransferases and phospholipases, which catalyze phospholipase A (PLA) and acyltransferase activities<sup>1</sup>. In mice, *Plaat3* is highly expressed in white adipose tissue (WAT) where it exhibits PLA<sub>1</sub> and PLA<sub>2</sub> activities, which hydrolyze fatty acids linked to the sn-1 or sn-2 positions of membrane phospholipids<sup>2-4</sup>. *Plaat3*-deficient mice are resistant to diet-induced obesity, and PLAAT3 was therefore suggested as a potential therapeutic target for metabolic syndrome<sup>5</sup>.

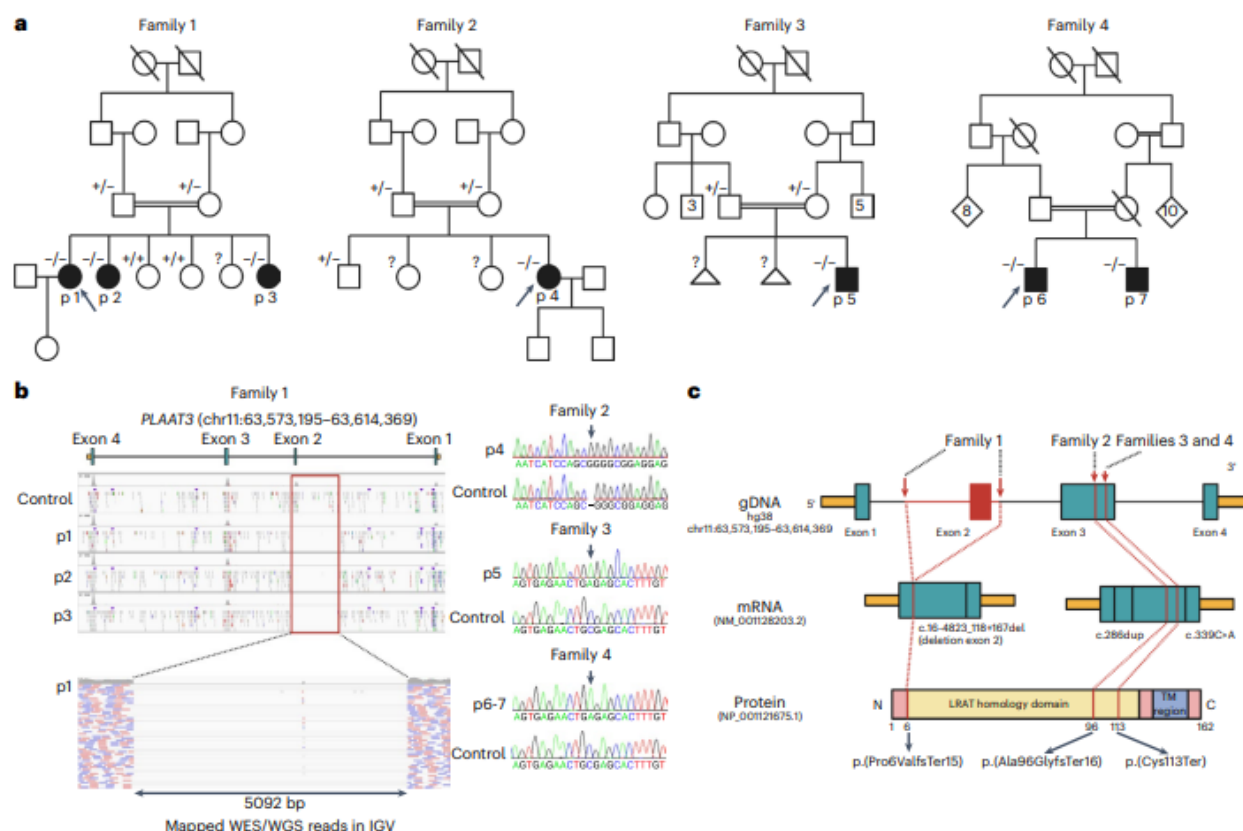
Lipodystrophy syndrome is characterized by generalized or partial lipodystrophy with reduced WAT due to a defect in adipose tissue storage of triglycerides (TGs)<sup>6</sup>. This results in lipid infiltration of nonadipose tissues leading to insulin resistance, increased liver

glucose production, hypertriglyceridemia and hepatosteatosis. Approximately 30 genes have been implicated in hereditary lipodystrophy syndrome<sup>7,8</sup>. Although these disorders often remain genetically unexplained, there is growing interest in identifying their molecular bases to improve genetic counseling, personalize treatment<sup>9,10</sup> and uncover signaling pathways involved in diabetes, hepatosteatosis or metabolic syndrome<sup>9,10</sup>. Patients with lipodystrophy syndrome show striking metabolic similarities with obese participants who display an incapacity to store surplus energy in overwhelmed WAT resulting in ectopic lipid accumulation.

In this study, we combined homozygosity mapping, whole-exome sequencing and whole-genome sequencing (WES/WGS) to identify homozygous null variants in *PLAAT3* as a cause of lipodystrophy

 e-mail: [bart.dermaut@ugent.be](mailto:bart.dermaut@ugent.be)

Nature Genetics



**Fig. 1 | Family pedigree structures, genetic findings and schematic display of the *PLAAT3* null variants in families 1, 2, 3 and 4. a**, Pedigrees of families 1, 2, 3 and 4. Affected family members are shown with filled circles (females) or squares (males). Diamond-shaped symbols are used when sex is unspecified. Arrows indicate the index patients. Double horizontal lines indicate consanguinity and diagonal lines are used for deceased individuals. Genotypes, if known, are depicted in the pedigrees ( $-/-$  for homozygotes,  $+/-$  for heterozygotes). Arrows point to the index patient in the pedigree. **b**, Left, integrative genome viewer (IGV) screenshots showing WES (top-left) and WGS (bottom-left) reads mapped to the reference *PLAAT3* sequence. No reads are mapped to exon 2 of the *PLAAT3* gene in patients 1, 2 and 3, indicative of a homozygous deletion. Right, Sanger

sequencing of the homozygous single-nucleotide insertion in patient 4 of family 2 and the nonsense variant in patients 5, 6 and 7 in families 3 and 4. The variants are indicated with an arrow. **c**, Schematic overview of the *PLAAT3* homozygous null variants identified through WES/WGS in families 1, 2, 3 and 4: the deletion of exon 2 (indicated in red) of the *PLAAT3* gene in family 1, the single base insertion in exon 3 in family 2 and the nonsense variant in exon 3 in families 3 and 4 (top). The consequences of the molecular defects are also shown at the mRNA level (middle) and at the protein level (bottom). The enzymatically active lecithin retinol acyltransferase (LRAT) and a transmembrane region (TM) of *PLAAT3* are depicted. All three variants result in a premature termination codon and are predicted to cause NMD.

syndrome with neurological features. We further assessed the metabolic consequences of the loss of *PLAAT3* using a multi-omics approach in mouse and human WAT. The impact of the loss of *PLAAT3* activity on human adipocyte differentiation and function was also evaluated using a CRISPR–Cas9-mediated genome-editing approach.

## Results

### Identification of homozygous *PLAAT3* null variants in seven patients

In a consanguineous family with three children (patients 1–3; family 1; Fig. 1a) presenting with a lipodystrophy syndrome combining metabolic and neurological manifestations, we performed single-nucleotide polymorphism (SNP) array-based homozygosity mapping in the three affected and two unaffected siblings. The largest homozygous region (~43 Mb), exclusively shared between the affected siblings, was found on chromosome 11p11.2–q14.2 and contained the *BSCL2* gene, implicated in congenital generalized lipodystrophy type 2 (Supplementary Table 1). Normal Sanger sequencing of *BSCL2* coding exons, followed by normal cDNA analysis of peripheral leukocytes, excluded *BSCL2* as

the causal gene. WES was performed in the three affected sisters and two unaffected siblings, but we found no homozygous pathogenic single-nucleotide variants in the region of interest (Supplementary Table 2). Copy number variant analysis of WES data using ExomeDepth<sup>11</sup> revealed a homozygous 103 bp deletion corresponding to exon 2 of the *PLAAT3* gene (NM\_00128203.2) exclusively in the three affected sisters (Fig. 1b, top, and Supplementary Table 3). The deletion was located within the ~43 Mb shared homozygous region and was confirmed by low-pass WGS (copy number variant sequencing (CNV-seq))<sup>12</sup>. WGS in one unaffected and two affected sisters determined the genomic deletion breakpoints (chr11: 63597894–63602986) revealing a homozygous 5092 bp deletion (Fig. 1b, bottom). Deletion of *PLAAT3* exon 2 (c.16-4823\_118+167del) results in a frameshift leading to a premature termination codon (p.(Pro6ValfsTer15)) (Fig. 1c). RNA sequencing confirmed that the molecular defect induced nonsense-mediated decay (NMD; Supplementary Fig. 1).

Using the GeneMatcher platform<sup>13</sup>, we identified a female patient (patient 4, family 2; Fig. 1a) in whom WES revealed a homozygous single base duplication in exon 3 of *PLAAT3* leading to a frameshift and



**Fig. 2 | DEXA scan and clinical pictures of PLAAT3-deficient patients.**

**a**, DEXA scan of patient 1 showing a lipotrophic body composition with a total fat percentage of 13.8%. **b**, Lateral view of patient 1 showing muscular hypertrophy in the upper and lower limbs. **c**, Frontal view of patient 4 demonstrating masculine features with muscle hypertrophy and lipotrophy of the upper and lower limbs with submental accumulation of adipose tissue. **d**, Frontal view of

patient 5 showing generalized lipodystrophy, gynecomastia, genu valgum and muscle hypertrophy of the upper limbs. **e**, Lateral view of patient 5 showing gynecomastia. **f**, Dorsal view of patient 5 demonstrating thoracic hyperkyphosis. **g**, Foot deformities in patient 5 including pes cavus and hammer toes. **h**, Frontal view of patient 6 showing lipotrophy of the trunk and gynecomastia. **i**, Acanthosis nigricans in the axillary region of patient 6.

premature termination codon, predicted to induce NMD (c.286dupG, p.(Ala96GlyfsTer16)) (Fig. 1b,c)<sup>13</sup>. Through data sharing we found three additional patients from two unrelated families (patient 5, family 3; patients 6–7, family 4; Fig. 1a) in which WES identified a homozygous nonsense variant (c.339C>A; p.(Cys113Ter)) in exon 3 of *PLAAT3* (Fig. 1c). The variants identified in families 2, 3 and 4 were confirmed by Sanger sequencing (Fig. 1b). Only the p.(Cys113Ter) variant was present in gnomAD v3.1.2 (in one individual of African origin in the heterozygous state).

#### Loss of PLAAT3 is associated with lipodystrophy syndrome

**Family 1.** Three sisters (patients 1–3), born from consanguineous parents of Turkish origin, presented with lipotrophy of limbs and trunk and lipohypertrophy in the submental and posterior cervical region in patients 2 and 3 (Fig. 1a). The patients were lean with a body-mass index (BMI) ranging from 19.6 to 20.5 kg m<sup>-2</sup>. In patient 1, the lipotrophy was confirmed by dual-energy X-ray absorptiometry (DEXA) with a total fat mass of 13.8% (Fig. 2a). Serum leptin levels in patient 1 and patient 2 were decreased (Supplementary Table 4). Other clinical signs typical for lipodystrophy syndrome included muscle hypertrophy, insulin-resistant diabetes with hyperinsulinemia, acanthosis nigricans and increased fasting glucose, hypertriglyceridemia with low high-density lipoprotein (HDL) cholesterol, liver steatosis and polycystic ovary syndrome with hirsutism (Fig. 2b, Table 1 and Supplementary Table 4). Neurological symptoms comprised migraine (patient 1), demyelinating neuropathy (patient 1) and intellectual disability (patient 2; Table 1 and Supplementary Table 4). Patient 1 presented with unilateral carpal tunnel syndrome. Additional features included generalized musculoskeletal pain, hypertension, subclinical hypothyroidism, precocious puberty and acromegalic features (Supplementary Table 4). Short stature was seen in all three affected siblings, which was probably constitutional given the short stature in both unaffected parents (mother, 158 cm and father, 165 cm).

**Family 2.** Patient 4 (female) was born from Algerian first-degree cousins (Fig. 1a). She presented with an android habitus with generalized muscle hypertrophy and lipotrophy in the limbs with a relative accumulation of adipose tissue in the face, neck and submental region resulting in a cushingoid appearance (Fig. 2c). She had a BMI of 24.6 kg m<sup>-2</sup> with insulin-resistant diabetes, hypertriglyceridemia with low HDL cholesterol, liver steatosis, polycystic ovary syndrome and hirsutism. She complained of chronic muscle pain and was diagnosed with demyelinating neuropathy (Table 1). Arterial hypertension, transient hypothyroidism, carpal tunnel syndrome, obstructive sleep apnea syndrome and glomerulopathy without renal insufficiency were observed (Supplementary Table 4). Recently, she was diagnosed with thyroid vesicular carcinoma requiring thyroidectomy.

**Family 3.** Patient 5 (male) was born from Algerian first-degree cousins (Fig. 1a). His mother had two early miscarriages before his birth. He presented with generalized lipodystrophy, muscular hypertrophy, insulin-resistant diabetes and liver steatosis (Fig. 2d and Table 1). The patient was lean with a BMI of 20.7 kg m<sup>-2</sup> and a body fat percentage of 10.1% (Table 1 and Supplementary Table 4). Serum leptin levels were very low. He developed gynecomastia at the age of 12 years, which persisted after puberty (Fig. 2d,e). Sex hormone levels were normal, with slightly elevated free testosterone levels due to decreased sex hormone-binding globulin levels associated with insulin resistance. Luteinizing hormone (LH) was within the normal range, and follicle stimulating hormone (FSH) was slightly upregulated with low inhibin B levels, suggesting spermatogenic insufficiency (Supplementary Table 4). He presented with facial dysmorphism including low-set and posteriorly rotated ears, prominent nose and forehead, high nasal bridge and prognathism. Musculoskeletal manifestations included thoracic hyperkyphosis, genu valgum, talipes equinovarus, retraction of Achilles tendons, hammer toes and bilateral syndactyly of the second and third toes (Fig. 2d–g). The patient presented with a rather

**Table 1 | Clinical and biological characteristics of patients with biallelic loss-of-function *PLAAT3* variants**

Homozygous <i>PLAAT3</i> variant	Family 1: patient 1	Family 1: patient 2	Family 1: patient 3	Family 2: patient 4	Family 3: patient 5	Family 4: patient 6	Family 4: patient 7	Summary
	c.16-4823_118+167del p.(Pro6ValfsTer15)	c.16-4823_118+167del p.(Pro6ValfsTer15)	c.16-4823_118+167del p.(Pro6ValfsTer15)	c.286dupGp. (Ala96GlyfsTer16)	c.339C>A p.(Cys113Ter)	c.339C>A p.(Cys113Ter)	c.339C>A p.(Cys113Ter)	
Sex	Female	Female	Female	Female	Male	Male	Male	Female (4/7), male (3/7)
Age at diagnosis (years)	38	37	25	42	2	15	11	Mean: 27
Age at symptom onset (years)								
Metabolic features	33	N/A	18	14	4	N/A	N/A	
Neurological/skeletal features	19	N/A	8	11	1	4	4	-
Lipodystrophy syndrome features								
Lipoatrophy	Generalized	Partial	Partial	Partial	Generalized	Partial	Partial	Partial (5/7), generalized (2/7)
Body-mass index (kg m <sup>-2</sup> )	19.7	20.5	20.5	24.6	20.7	22.2	19.1	Mean: 21
Insulin resistance	+	+	+	+	+	+	+	7/7
Liver steatosis	+	+	+	+	+	+	+	7/7
Dyslipidemia/hypertriglyceridemia	+	+	+	+	-	+	+	6/7
Neurological features								
Psychomotor retardation/intellectual disability	-	+	-	-	+	+	-	3/7
Demyelinating peripheral neuropathy	+	N/A	N/A	+	+	+	+	5/7
Musculoskeletal features								
Kyphoscoliosis	-	-	-	-	+	+	+	3/7
Chronic muscle/joint pain	+	N/A	+	+	-	+	+	4/7
Acromegalic features	+	-	+	-	-	+	+	4/7
Additional features								
Gynecomastia	N/A	N/A	N/A	N/A	+	+	+	3/3
Dysmorphic features	Prominent nose and chin	-	Prominent nose and chin	-	Low-set ears, deep-set eyes, prognathism, beaked nose and flat forehead	Short neck and low hair implantation	Short neck and low hair implantation	5/7

N/A, not available/not applicable.

short stature, which was also seen in his unaffected parents (Supplementary Table 4). During childhood, advanced bone age with bone hypermineralization ( $z$  score  $> 3.5$  s.d.) was observed. Neurological features consisted of psychomotor delay, intellectual disability, behavioral problems, demyelinating neuropathy and spastic gait (Table 1). Magnetic resonance imaging (MRI) of the brain was normal.

**Family 4.** Patients 6 and 7 (male brothers) were born from Algerian first-degree cousins (Fig. 1a). They were lean with normal BMIs and displayed facial acromegalic features, kyphoscoliosis and lumbar hyperlordosis (Table 1). Both patients presented with walking difficulties at the age of 4 years. Neuromusculoskeletal manifestations included tiptoe walking, absence of deep tendon reflexes, retraction of Achilles tendons and muscle pain. Metabolic features included partial lipoatrophy affecting the shoulder girdle, upper extremities and trunk associated with muscular hypertrophy and pronounced subcutaneous veins, gynecomastia, insulin resistance, acanthosis nigricans, diabetes, liver steatosis, hypertriglyceridemia in the older brother and decreased

HDL cholesterol in the younger brother (Fig. 2h,i, Table 1 and Supplementary Table 4). Development of secondary sex characteristics was normal, reflected by normal gonadotropin and sex hormone levels (Supplementary Table 4). A demyelinating sensorimotor neuropathy was present in both patients. Brain MRI revealed Arnold–Chiari malformation type 1 in patient 6 (Supplementary Table 4).

#### *Plaat3*<sup>-/-</sup> mouse WAT displays a peroxisome proliferator-activated receptor gamma (PPAR $\gamma$ )-dependent differentiation defect

The mouse *Plaat3* protein is evolutionary, highly conserved and is 83% identical and 91% similar to human *PLAAT3* (<https://www.flyrnai.org/diopt>). *Plaat3*<sup>-/-</sup> mice have been reported to be resistant to diet-induced obesity and show a lipodystrophy syndrome-like phenotype<sup>5</sup>. This phenotype was suggested to be due to loss of *PLA $\gamma$*  activity resulting in a decline in arachidonic acid (AA, C20:4) release and prostaglandin E2 (PGE2) synthesis and an increase in levels of cyclic adenosine monophosphate (cAMP), leading to increased lipolysis in normally

differentiated adipocytes<sup>5</sup>. Other studies have suggested an important role for PLAAT3 in adipocyte differentiation<sup>14–16</sup>.

We used *Plaat3*<sup>-/-</sup> mouse-derived inguinal WAT samples<sup>17</sup> for morphological and multi-omics studies to further clarify the pathogenic mechanism of PLAAT3-related lipodystrophy syndrome. We collected 12 inguinal *Plaat3*<sup>-/-</sup> WAT biopsy specimens and 12 *Plaat3*<sup>+/+</sup> inguinal WAT control specimens. Light microscopic analysis of *Plaat3*<sup>-/-</sup> samples and *Plaat3*<sup>+/+</sup> controls showed that adipocytes in *Plaat3*-deficient tissue were significantly smaller in size, in line with a previous report<sup>5</sup>, while no difference in shape (circularity) was observed (Fig. 3a,b).

Next, we evaluated the effect of *Plaat3* deficiency on the lipid content in inguinal WAT using a liquid chromatography-tandem mass spectrometry (LC-MS/MS) lipidomics approach. Dimensionality reduction showed clustering of samples according to sex along the first component and according to genotype (*Plaat3*<sup>-/-</sup> or *Plaat3*<sup>+/+</sup>) along the second component, indicating that loss of *Plaat3* activity influences the abundance of certain lipid classes and species (Fig. 3c). Given the known enzymatic function of PLAAT3 in membrane phospholipid remodeling, we checked the overall abundance of the 11 phospholipid subclasses in *Plaat3*<sup>-/-</sup> and *Plaat3*<sup>+/+</sup> WAT (Supplementary Fig. 2). Increased levels were seen for phosphatidylethanolamine (PE), phosphatidylinositol (PI), phosphatidylserine (PS) and 1-alkenyl, 2-acyl phosphatidylethanolamine in both female and male *Plaat3*<sup>-/-</sup> WAT, with PI and PS showing the most striking differences (Supplementary Fig. 2). Within each phospholipid subclass, we filtered all phospholipid species with statistically significant ( $P < 0.05$ ) differences between *Plaat3*<sup>-/-</sup> and *Plaat3*<sup>+/+</sup> WAT that showed alterations in the same direction in female and male samples. This revealed a total of 14 lipid species (Supplementary Table 5) of which two, PI(18:0/20:4) and PE(18:0/18:2), represent the most common species within their lipid class and were clearly upregulated in *Plaat3*<sup>-/-</sup> WAT (Plaat3 mouse lipidomics.html). In animal tissues, PI(18:0/20:4) is the primary source of AA (C20:4) required for the biosynthesis of eicosanoids (including prostaglandins) via PLA<sub>2</sub> activity, which releases these bioactive fatty acids from position sn-2 (ref. 18). However, mediator lipidomics showed no statistically significant differences for AA or AA-derived eicosanoids, including PGE<sub>2</sub>, between *Plaat3*<sup>-/-</sup> and *Plaat3*<sup>+/+</sup> WAT (Fig. 3d; Plaat3 mouse mediator lipidomics.html).

We also performed RNA sequencing on seven *Plaat3*<sup>-/-</sup> (3 males and 4 females) and eight *Plaat3*<sup>+/+</sup> mouse (4 males and 4 females) WAT samples. A total of 99 genes were significantly differentially expressed ( $|\log_2$  fold change (FC)|  $\geq 1$ , raw  $P < 0.05$ ), of which 59 were upregulated and 40 downregulated (Fig. 3e). We used Metascape to search for enriched pathways within the differentially expressed gene (DEG) sets (top 100 downregulated and upregulated genes ranked by FC with raw  $P < 0.05$ ) and found that genes involved in synaptic transmission, glycolipid

biosynthetic processes and fat cell differentiation were downregulated, whereas genes involved in myelination and inflammatory responses were upregulated in *Plaat3*<sup>-/-</sup> WAT (Fig. 3f). We used the epigenetic landscape in silico deletion analysis (LISA) tool to identify transcriptional regulators (TRs) of gene networks within the downregulated and upregulated genes (Fig. 3g)<sup>19</sup>. Interestingly, the most important TR of the downregulated genes was PPAR $\gamma$  (or PPARG), which is responsible for adipocyte differentiation together with its coactivator retinoid X receptor alpha (RXRA or RXR $\alpha$ ) with which it forms a heterodimer to induce transcription of target genes<sup>20,21</sup>. Transcript levels of the three lipolytic enzymes (*Pnpla2*, *Lipe* and *Mgl1*) were not altered in *Plaat3*<sup>-/-</sup> WAT compared to *Plaat3*<sup>+/+</sup> (Supplementary Fig. 3).

Finally, LC-MS/MS-based shotgun proteomics analysis was performed on nine *Plaat3*<sup>-/-</sup> (4 males and 5 females) and nine *Plaat3*<sup>+/+</sup> (4 males and 5 females) inguinal WAT biopsy specimens, which reliably quantified a total of 4,033 protein groups in all samples. Differential protein intensity analysis ( $|\log_2$ (FC)|  $\geq 0.5$ , raw  $P < 0.05$ ) showed 41 significantly downregulated and 8 upregulated proteins (Fig. 3h). Again, we looked for enriched pathways within the top 100 downregulated and upregulated proteins (ranked according to FC with raw  $P < 0.05$ ). Proteins involved in fatty acid metabolic and lipid biosynthetic processes were less abundant in *Plaat3*<sup>-/-</sup> WAT, whereas proteins involved in autophagy and catabolism were more abundant (Fig. 3i). With the LISA tool, PPAR $\gamma$  and its coactivators, such as CCAAT enhancer binding protein  $\alpha$  and  $\beta$  (CEBPA and CEBPB), were again identified as the strongest TRs within the set of downregulated proteins (Fig. 3j). The most strongly upregulated protein was Plin2 ( $|\log_2$ (FC)| = 2.69, raw  $P < 0.0001$ ; Fig. 3h). PLIN2 is known to be expressed on lipid droplet membranes of pre-adipocytes and is replaced by PLIN1 after completion of the cell maturation process<sup>22</sup>. Along with PLIN3, PLIN2 also contributes to the formation of lipid droplet-lysosome contacts and its removal by chaperone-mediated autophagy facilitates lipolysis<sup>23,24</sup>. Taken together, these results show that mouse *Plaat3*<sup>-/-</sup> WAT consists of smaller adipocytes containing increased levels of the PI(18:0/20:4) membrane phospholipid and showing a PPAR $\gamma$ -dependent adipocyte differentiation defect with abnormal lipid droplet metabolism.

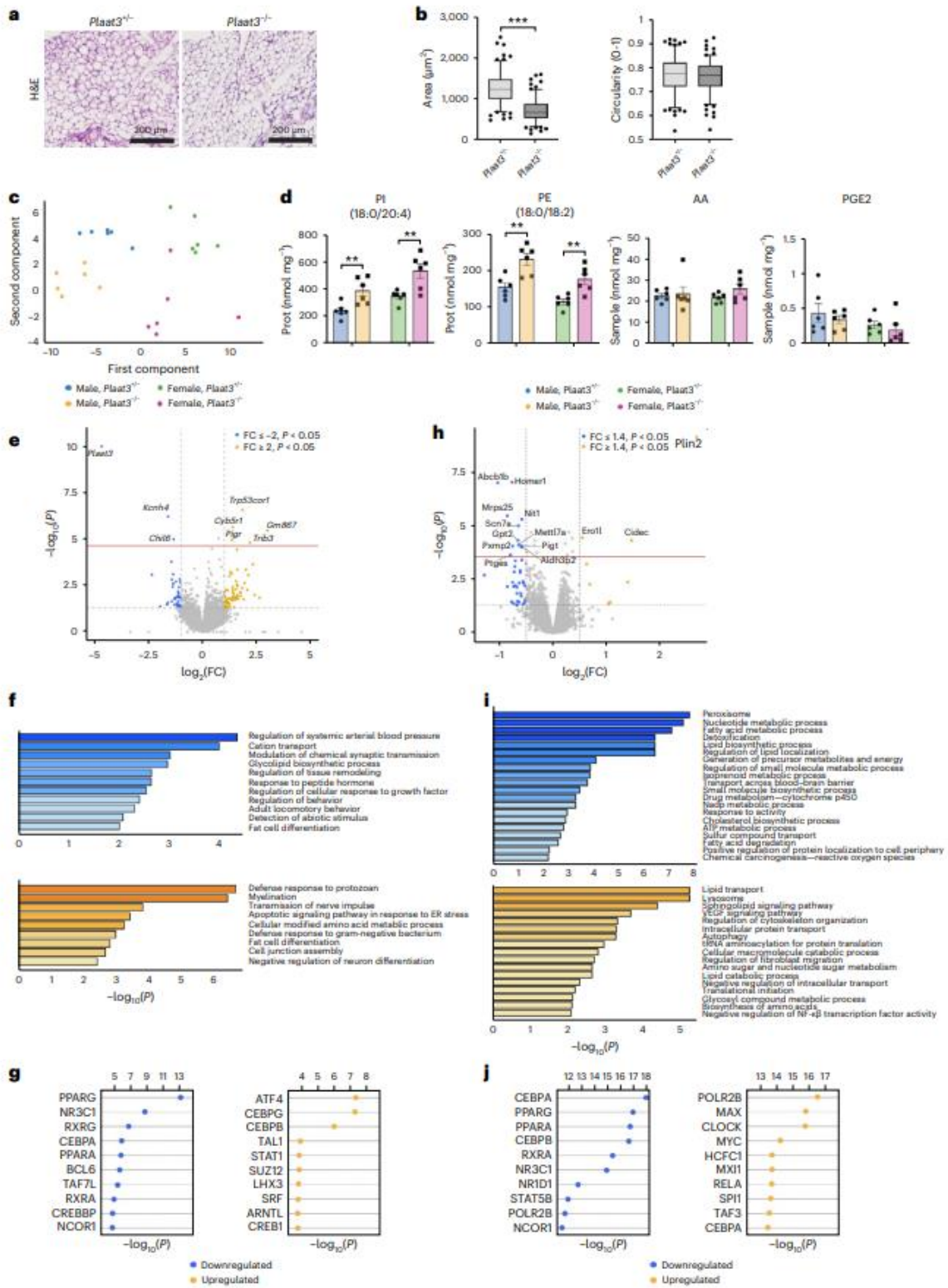
### Patient WAT shows a PPAR $\gamma$ -mediated adipogenesis defect

We collected subcutaneous WAT samples of female patients 1 and 4, as well as six healthy female controls with a normal BMI (19–23 kg m<sup>-2</sup>). Light microscopic analysis showed that patient adipocytes were larger and showed a statistically significant difference in ‘circularity’ (Fig. 4a,b). Patient WAT also displayed increased inflammation as assessed by CD68 staining (Fig. 4a, middle). Immunohistochemistry using a 70-kDa peroxisomal membrane protein (PMP70) antibody did not show remarkable differences in peroxisome abundance (Fig. 4a, right).

**Fig. 3 | Histopathology, lipidomics, proteomics and differentially expressed gene analysis in *Plaat3*<sup>-/-</sup> and *Plaat3*<sup>+/+</sup> mouse WAT. a**, Light microscopic images after staining with H&E of inguinal WAT from a *Plaat3*<sup>-/-</sup> and a *Plaat3*<sup>+/+</sup> male mouse. Scale bars represent 200  $\mu$ m. **b**, Box plot of adipocyte size measurement ( $\mu$ m<sup>2</sup>;  $P < 0.001$ , left) and morphology expressed as ‘circularity’ (0–1, right) of *Plaat3*<sup>-/-</sup> and *Plaat3*<sup>+/+</sup> mouse WAT. The center line represents the median, the box limits represent the interquartile range (IQR) and the whiskers indicate the minimum and maximum values. *P* values were calculated using two-sided independent samples *t* test. \*\*\* $P < 0.001$ . **c**, PCA of lipidomics on WAT from male ( $n = 6$ ) and female ( $n = 6$ ) *Plaat3*<sup>-/-</sup> and male ( $n = 6$ ) and female ( $n = 6$ ) *Plaat3*<sup>+/+</sup> mice. **d**, Bar charts of PI (18:0/20:4; male: raw  $P = 0.008$ , female: raw  $P = 0.007$ ), PE (18:0/18:2; male: raw  $P = 0.003$ , female: raw  $P = 0.004$ ), AA and PGE<sub>2</sub> levels depicted as the means  $\pm$  standard error of the mean (s.e.m.). *P* values were calculated using one-way analysis of variance (ANOVA), and the Benjamini–Hochberg procedure was conducted to correct for multiple testing. \*\* $P < 0.01$ . **e**, Volcano plot showing downregulated (blue) and upregulated (yellow) DEGs with  $|\text{FC}| \geq 2$  and raw  $P < 0.05$  (Exact Test in edgeR) in WAT from *Plaat3*-deficient male ( $n = 3$ ) and female ( $n = 4$ ) mice compared to male ( $n = 4$ )

and female ( $n = 4$ ) *Plaat3*<sup>+/+</sup> mice. DEGs above the red line, FDR-adjusted  $P < 0.05$  (Benjamini–Hochberg correction). **f**, KEGG and GO pathway enrichment analysis of DEGs (top 100 downregulated (blue) and upregulated (yellow) genes ranked according to FC with raw  $P < 0.05$ ) using Metascape (hypergeometric test and Benjamini–Hochberg correction). **g**, Top ten TRs of the downregulated (blue) and upregulated (yellow) genes (top 100 ranked according to FC with raw  $P < 0.05$ ) identified by LISA (one-sided Wilcoxon rank-sum test is used to assess statistical significance). **h**, Volcano plot showing downregulated (blue) and upregulated (yellow) proteins with differential protein intensities with  $|\text{FC}| \geq 1.4$  and raw  $P < 0.05$  (two-sided independent samples *t* test) in WAT from male ( $n = 4$ ) and female ( $n = 5$ ) *Plaat3*<sup>-/-</sup> and male ( $n = 4$ ) compared to female ( $n = 5$ ) *Plaat3*<sup>+/+</sup> mice. Proteins above the red line, FDR-adjusted  $P < 0.05$  (Benjamini–Hochberg correction). **i**, KEGG and GO pathway enrichment analysis of differentially abundant proteins (top 100 downregulated (blue) and upregulated (yellow) ranked according to FC with raw  $P < 0.05$ ) using Metascape. **j**, Top ten TRs of downregulated (blue) and upregulated (yellow) proteins (top 100 ranked according to FC with raw  $P < 0.05$ ) identified using LISA.





LC-MS/MS-based lipidomics analysis was performed on the two patient WAT samples and six controls. Principal component analysis (PCA) based on all the measured lipid species across all classes showed that the lipid profile in the patients clearly differed from that in the controls (Fig. 4c). We again focused on phospholipids and observed increased levels for 9 of the 11 phospholipid subclasses (Supplementary Fig. 4). Out of 26 phospholipid species with statistically significant ( $P < 0.05$ ) differences between patient and control WAT, 25 were upregulated (Supplementary Table 6). When ranked according to  $P$  value, AA-containing phospholipids lysophosphatidylcholine (LPC; 20:4) ( $\log_2(\text{FC}) = 2.05 \pm 0.38$ ,  $P < 0.001$ ) and lysophosphatidylethanolamine (LPE; 20:4) ( $\log_2(\text{FC}) = 1.24 \pm 0.28$ ,  $P = 0.004$ ) were among the top-ranked lipid species (Supplementary Table 6 and Fig. 4d). These results are in line with decreased PLA<sub>2</sub> activity in human PLAAT3-deficient WAT (Fig. 4d, right; PLAAT3 human lipidomics.html).

Finally, RNA sequencing was performed on WAT samples from patient 1 and three healthy controls to identify DEGs as a consequence of PLAAT3 deficiency. A total of 8,711 genes were identified that were significantly upregulated or downregulated in patient WAT ( $|\log_2(\text{FC})| \geq 1$ , raw  $P < 0.05$ ). Of these 8,711 genes, the large majority ( $n = 6,710$ ) were downregulated in patient WAT, whereas 2,001 genes were significantly upregulated (Fig. 4e). Pathway enrichment analysis of the 100 most downregulated genes (protein coding and noncoding) identified cytoplasmic translation as the most enriched cluster, indicative of a metabolically inactive state of adipocytes. Strikingly, in human WAT, genes involved in the PPAR signaling pathway were also shown to be downregulated (Fig. 4f). Transcript levels of well-established PPAR $\gamma$  target genes involved in adipocyte differentiation (*PLINI* and *FABP4*), fatty acid transport (*LPL* and *CD36*) and energy homeostasis (*LEP* and *ADIPOQ*) were significantly decreased in patient adipocytes (Fig. 4e). The same analysis was performed on the top 100 most strongly upregulated genes that identified olfactory transduction and cellular dopamine response as the only significantly upregulated biological pathways. The LISA tool was used to identify TRs of gene networks within the 500 most strongly downregulated and upregulated genes (Fig. 4g). In line with the results in mouse WAT, out of the top ten ranked TRs of downregulated genes in human WAT, four were strongly linked to PPAR signaling including PPAR $\gamma$  itself, its two main cooperating TRs CEBPA and CEBPB and its coactivator mediator complex subunit 1 (refs. 20,21,25). We used Gene Set Enrichment Analysis (GSEA) software to look for enriched gene sets among the DEGs. Among the downregulated genes, we observed an enrichment of MYC Proto-Oncogene (MYC) target genes and genes associated with oxidative phosphorylation and adipogenesis (Fig. 4h,i). Together, these results are in line with our *Plaat3*<sup>-/-</sup> mouse data and suggest a PPAR $\gamma$ -mediated adipogenesis defect accompanied by reduced PLA<sub>2</sub> activity in human PLAAT3-deficient WAT.

### PLAAT3 knockout (KO) in adipose stem cells (ASCs) disrupts lipid droplet formation

To better address the role of PLAAT3 in human adipogenesis, a custom-designed single-guide RNA (sgRNA)/Cas9 expression vector targeting the second exon of PLAAT3 was used (Fig. 5a,b) in human ASCs, isolated from abdominal subcutaneous adipose tissue<sup>26</sup>. Human ASCs were chosen as a model because of their ability to differentiate into mature adipocytes after stimulation in vitro (Fig. 5c). The efficiency of PLAAT3 KO was confirmed by Sanger sequencing of PLAAT3 exon 2 in genomic DNA from KO ASC cells, which revealed a high level of on-target indels with 70% insertions and 26% deletions (Fig. 5a,b). Control (CTL) cells, expressing scrambled sgRNA, differentiated into adipocytes within 20 d (D20; Fig. 5c,d) and displayed strong accumulation of lipid droplets and TG content in the cytoplasm (Fig. 5e,f). In contrast, PLAAT3 KO led to a strong and significant decrease in lipid droplet formation ( $P < 0.0001$ ; Fig. 5d,e) as well as TG content ( $P < 0.0001$ ; Fig. 5f). The perilipin level, encoded by *PLINI*, a mature adipocyte marker, was sharply decreased in PLAAT3 KO cells (Fig. 5g). Interestingly, the PLAAT3 KO cells displayed decreased expression of PPAR $\gamma$  (Fig. 5g). We next investigated the effect of PLAAT3 loss on insulin sensitivity. In wild-type (WT) and control pre-adipocytes (D0) stimulated with insulin, Western blot analysis revealed a strong increase in the phosphorylation of AKT and extracellular-regulated kinase (ERK; Fig. 5h). In contrast, the PLAAT3 KO cells at D0 were resistant to insulin, as shown by the lack or strong decrease in the phosphorylation of these intermediates upon insulin stimulation (Fig. 5h). Altogether, these findings confirm that PLAAT3 regulates adipocyte differentiation through a cellular process involving PPAR $\gamma$ .

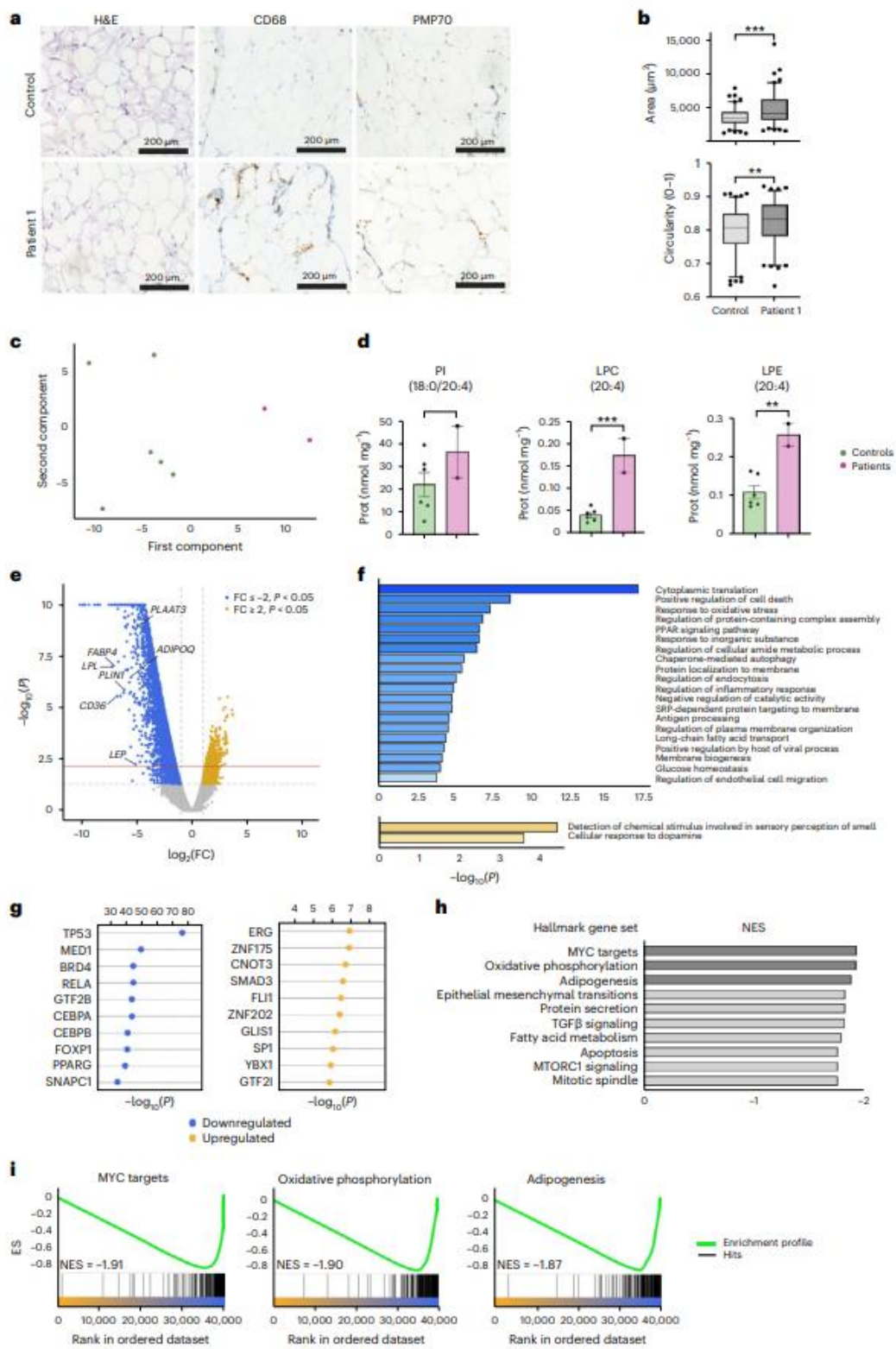
### Discussion

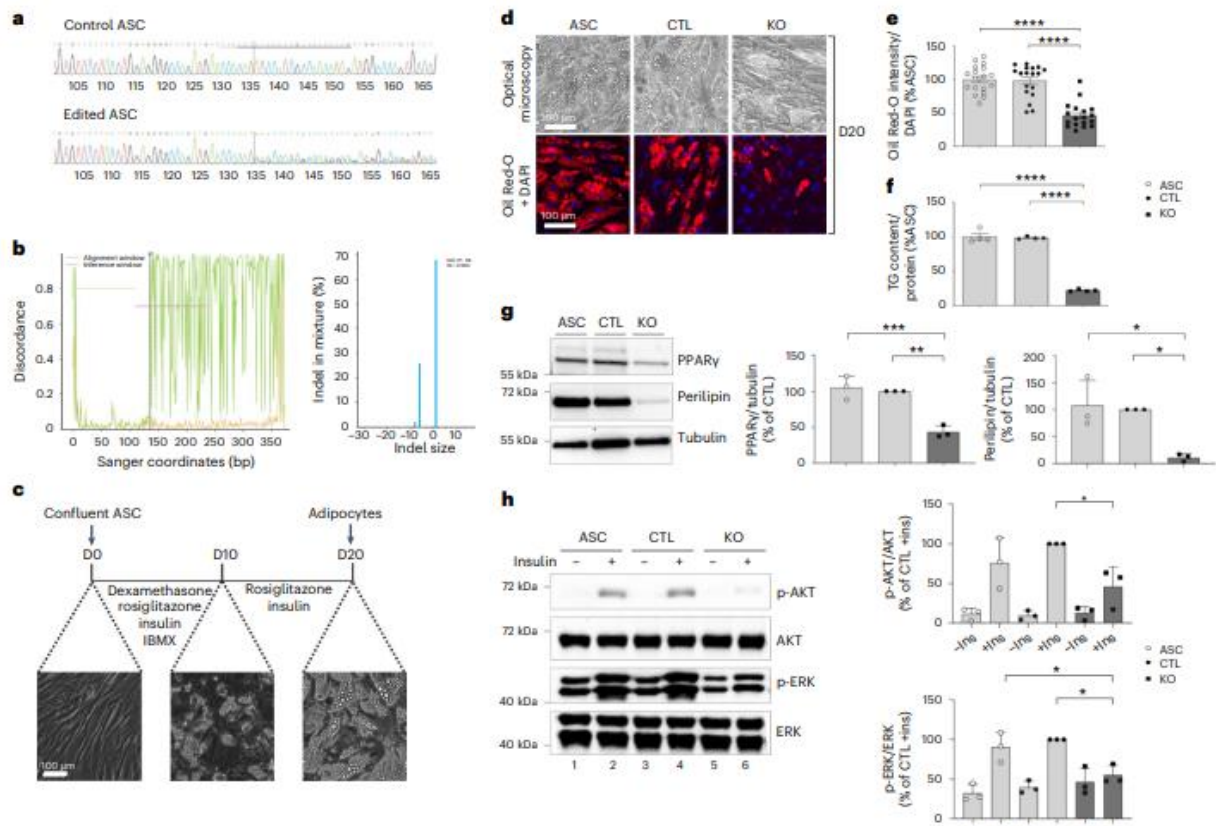
This study establishes PLAAT3 deficiency as a monogenic lipodystrophy syndrome associated with neurological manifestations. Our data suggest a key pathogenic role for PLAAT3 in PPAR $\gamma$ -mediated adipogenesis and thus identify PLAAT3 as a potential therapeutic handle to regulate PPAR $\gamma$  signaling, a promising therapeutic target to increase insulin sensitivity in type 2 diabetes<sup>27</sup>.

We provide genetic evidence for the involvement of PLAAT3 in a complex metabolic and neurological syndrome. We report seven patients from four unrelated families harboring different biallelic loss-of-function variants in PLAAT3 and displaying a clinical presentation combining lipodystrophy syndrome with neurological features, consistent with the expression profile of PLAAT3, which is mainly expressed in adipose tissue, brain and peripheral nerves (<https://www.gtexportal.org/home/gene/PLA2G16>). In all patients, various degrees of lipodystrophy were seen, associated with metabolic features secondary to the loss of healthy adipose tissue. The historical classification of lipodystrophy syndrome distinguishes the generalized forms (congenital generalized lipodystrophy) with generalized fat loss apparent at birth and the partial forms (familial partial lipodystrophy

**Fig. 4 | Histopathology, lipidomics and differentially expressed gene analysis of patient and control WAT biopsies. a–i.** Data were obtained from upper arm subcutaneous WAT of patients 1 and 4, as well as six control participants ( $n = 8$ ). **a**, Light microscopic images after H&E, CD68 and PMP70 staining of control and patient (p1) WAT. Scale bars represent 200  $\mu\text{m}$ . **b**, Box plot of adipocyte size measurement ( $\mu\text{m}^2$ ;  $P < 0.001$ , top) and morphology expressed as ‘circularity’ (0–1;  $P = 0.003$ , bottom) patient and control WAT. The center line represents the media, the box limits represent the IQR and the whiskers indicate the minimum and maximum values.  $P$  values were calculated using two-sided independent samples  $t$  test. **c**, PCA of lipidomics in two patient samples (in purple) and six control samples (in green). **d**, Bar charts of PI (18:0/20:4), LPC (20:4; raw  $P < 0.001$ , FDR-adjusted  $P = 0.017$ ) and LPE (20:4; raw  $P = 0.004$ , FDR-adjusted  $P = 0.044$ ) levels expressed as the means  $\pm$  s.e.m.  $P$  values were calculated using one-way ANOVA, and the Benjamini–Hochberg procedure was conducted to correct for multiple testing.

**e**, Volcano plot showing the downregulated (blue) and upregulated (yellow) DEGs with  $|\text{FC}| \geq 2$  and raw  $P < 0.05$  (exactTest in edgeR) in WAT from patient 1 compared to three controls ( $n = 4$ ). DEGs above the red line, FDR-adjusted  $P < 0.05$  (Benjamini–Hochberg correction). **f**, KEGG and GO pathway enrichment analysis of DEGs (top 100 downregulated (blue) and upregulated (yellow) genes ranked according to FC with raw  $P < 0.05$ ) using Metascape (hypergeometric test and Benjamini–Hochberg correction). **g**, Top ten TRs of downregulated (blue) and upregulated (yellow) genes (top 500 ranked according to FC with raw  $P < 0.05$ ) identified using LISA (one-sided Wilcoxon rank-sum test is used to assess statistical significance). **h**, Top ten enriched hallmark gene sets within the downregulated genes identified using GSEA software, ranked according to the NES. **i**, Enrichment plots with NES of the top three enriched hallmark gene sets. NES, normalized enrichment score; ES, enrichment score.





**Fig. 5 | *PLAAT3* deficiency suppresses white adipocyte differentiation of ASCs.** **a–h**, Data were obtained from human ASCs, ASCs with a CRISPR–Cas9-mediated *PLAAT3* KO and ASCs transduced with a CTL Cas9/scramble gRNA plasmid. *P* values were determined by one-way ANOVA with the Geisser–Greenhouse correction and Tukey’s multiple comparisons test. The results are expressed as means  $\pm$  s.e.m., with \**P* < 0.05, \*\**P* < 0.01 and \*\*\**P* < 0.001, \*\*\*\**P* < 0.0001. **a**, Sanger sequencing confirmed high recombination in the target region (exon 2). The expected break site is indicated by a vertical dotted line and the gRNA sequence is underlined. **b**, The CRISPR indel pattern analysis in CTL and edited ASCs using Synthego software. The discordance plot showed a high rate of sequence misalignment after the gRNA target site, validating an editing efficiency of 96%. **c**, A timeline representation of ASC differentiation with a hormonal cocktail. The scale bar represents 100  $\mu$ m. **d**, Adipocyte differentiation was assessed through Oil Red-O lipid staining, with representative images shown by optical and fluorescence microscopy. Scale bars represent 100  $\mu$ m. **e**, Quantification of Oil Red-O fluorescence normalized to DNA content (DAPI). The measurements represent three independent experiments (six replicates each; *P* < 0.0001). Light gray bars correspond to control conditions

(ASC and CTL), and dark gray bars correspond to KO condition. **f**, Intracellular triglyceride content at D20. The measurements represent four independent experiments (*P* < 0.0001). Light gray bars correspond to control conditions (ASC and CTL), and dark gray bars correspond to KO condition. **g**, Western blot analysis of adipocyte markers at D20 in ASC cells. Images are representative of three independent experiments. Quantification was performed using Fiji software and normalized to the value of CTL cells at D20. PPAR $\gamma$ , KO versus ASC (*P* = 0.0010) and KO versus CTL (*P* = 0.0015). Perilipin, KO versus ASC (*P* = 0.0113) and KO versus CTL (*P* = 0.0165). **h**, Activation of insulin signaling in ASCs before differentiation at D0. ASC, CTL and *PLAAT3* KO cells were deprived of serum for 6 h, stimulated with 20 nM insulin for 5 min or left untreated, and subjected to immunoblotting with antibodies against total and phospho-AKT, and ERK1/2. Western blot images are representative of three independent experiments and were quantified using Fiji software, with normalization to CTL cells at D20 treated with insulin. pAKT, KO versus CTL (*P* = 0.0239). pERK, KO versus ASC (*P* = 0.00471) and KO versus CTL (*P* = 0.0103). D0, day 0; D10, day 10; D20, day 20; IBMX, 3-isobutyl-1-methylxanthine.

(FPLD)) beginning later in life, frequently in adolescence. However, this dichotomy does not apply to all situations, and some patients with genetic lipodystrophy syndrome have already been reported with normal fat distribution at birth and the appearance of generalized fat loss later in life<sup>28,29</sup>. In patients with *PLAAT3* deficiency, both partial lipodystrophy associated with facio-cervical lipohypertrophy and generalized lipodystrophy are observed. Biallelic mutations in *PCYT1A* represent another example of an unclassified genetic lipodystrophy syndrome that can present as either generalized or partial fat loss<sup>30,31</sup>. All investigated *PLAAT3*-deficient patients displayed common complications associated with lipodystrophy syndrome (for example,

insulin resistance, diabetes, hypertriglyceridemia, liver steatosis and polycystic ovary syndrome). Most patients also presented with neurological features. Demyelinating sensorimotor neuropathy was present in all investigated patients. Psychomotor retardation or intellectual disability was present in half of them.

Using multi-omics studies in human and mouse WAT, we discovered a downregulated PPAR $\gamma$ -mediated gene network as a downstream consequence of *PLAAT3* inactivation. PPAR $\gamma$  is the master regulator of adipocyte differentiation and function and was identified as the causal gene for an autosomal dominant lipodystrophy syndrome called FPLD3 (ref. 32). Accordingly, our study, using CRISPR–Cas9

KO of *PLAAT3* in human ASCs, revealed an adipocyte differentiation defect characterized by a major decrease in intracellular neutral lipid levels as well as decreased expression of adipogenesis and mature adipocyte markers including PPAR $\gamma$ . Insulin signaling was also altered, even in pre-adipocytes. These functional data are consistent with the lipotrophic and insulin-resistant phenotype of *PLAAT3* patients. A similar adipocyte differentiation defect has been reported in other lipodystrophy syndromes of various genetic origins<sup>29,33–35</sup>. Our data are in line with earlier studies suggesting a central role for *PLAAT3* in adipocyte differentiation<sup>14–16</sup>. A close link between *PLAAT3* and PPAR $\gamma$  is supported by a mouse adipocyte differentiation study demonstrating that *Plaat3* mRNA levels started to increase between 6 and 12 h after PPAR $\gamma$ -mediated initiation of differentiation<sup>14</sup>. A similar observation was made for human *PLAAT3* transcripts whose expression began to rise 6 h after induction of human pre-adipocyte differentiation and continued to increase during adipogenesis, such as PPAR $\gamma$ , whose expression started a few hours earlier<sup>16</sup>. The expression of *PLAAT3* is therefore appropriately timed to mediate PPAR $\gamma$ -driven adipogenesis. The exact mechanisms by which *PLAAT3* and PPAR $\gamma$  are linked remain however partially elusive. In contrast to our results indicating a disturbed PPAR $\gamma$ -mediated gene network downstream of *PLAAT3* inactivation, mouse *Plaat3* was identified as a direct transcriptional target of PPAR $\gamma$ , placing PPAR $\gamma$  upstream of *PLAAT3* in adipogenesis<sup>14</sup>. One explanation could be that, after initial induction of PPAR $\gamma$  by TFs including C/EBP $\beta$  and C/EBP $\delta$ <sup>21</sup>, PPAR $\gamma$  quickly induces *PLAAT3* expression<sup>14</sup>, which is then needed to further sustain PPAR $\gamma$  activity during and possibly after adipocyte differentiation (present study). To date, an endogenous and physiologically relevant ligand for PPAR $\gamma$  has not been identified, but it was shown that naturally occurring bioactive compounds including polyunsaturated fatty acids such as AA efficiently modify PPAR $\gamma$  activity<sup>36,37</sup>. It is thus tempting to speculate that *PLAAT3* has a role as a PPAR $\gamma$  ligand generator by liberating AA or related bioactive lipids from membrane phospholipids in WAT. Although the increased levels of AA-containing (lyso)phospholipids in mouse and human *PLAAT3*-deficient WAT observed in the lipidomics studies would support this hypothesis, we could not detect significantly altered free AA levels or other bioactive lipid mediators in mouse *Plaat3*-deficient WAT. Further research is needed to explore this possibility.

Furthermore, it cannot be excluded that *PLAAT3* has an additional PPAR $\gamma$ -dependent role in mature adipocytes. Indeed, next to its central involvement in adipogenesis, PPAR $\gamma$  is crucial for mature fat cell function where it regulates lipid metabolism and glucose homeostasis<sup>27</sup>. In line with this idea, a previous report in mice has shown chronically upregulated lipolysis in normally differentiated *Plaat3*<sup>-/-</sup> adipocytes due to decreased PGE2 levels as a consequence of abrogated PLA activity, which is a known inhibitor of lipolysis<sup>5,38</sup>. However, our mediator lipidomics study could not confirm the reduced PGE2 levels in mouse *Plaat3*-deficient WAT, and no upregulation of lipolytic enzymes on the transcript or protein level was seen. Nevertheless, we did not formally investigate lipolysis in the current study, and it is possible that it occurs in human *PLAAT3* patients. Similar observations have been made for the *LIPE* gene involved in another form of lipodystrophy syndrome<sup>39</sup>.

Comparison of *Plaat3* null mice described in 2009 with patients carrying homozygous pathogenic *PLAAT3* variants reveals a number of similarities, which are as follows: a lipotrophic phenotype, decreased circulating levels of leptin and adiponectin, insulin resistance and liver steatosis<sup>5</sup>. Nevertheless, some differences can also be noted. Serum TGs were lower in KO mice, whereas most patients display hypertriglyceridemia. Mice display smaller adipocytes (current study and previous data), whereas adipocyte size in patient WAT seems to be slightly increased<sup>5</sup>. Such discrepancies are not unusual and, while there is no doubt about the usefulness of mouse models to study human lipodystrophy syndrome, differences between human and mouse fat distribution and lipid metabolism represent known limitations<sup>40,41</sup>.

Due to early rodent studies suggesting potential antiobesity and antiviral properties, modifying *PLAAT3* activity was thought to be an interesting therapeutic strategy. Interestingly, while it was shown that drug targets with human genetic support are twice as likely to lead to approved drugs<sup>42,43</sup>, our study suggests that blocking *PLAAT3* activity too strongly could lead to unwanted metabolic side-effects secondary to downregulation of PPAR $\gamma$ -regulated gene networks.

## Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41588-023-01535-3>.

## References

- Pang, X. Y. et al. Structure/function relationships of adipose phospholipase A2 containing a Cys–His–His catalytic triad. *J. Biol. Chem.* **287**, 35260–35274 (2012).
- Duncan, R. E., Sarkadi-Nagy, E., Jaworski, K., Ahmadian, M. & Sul, H. S. Identification and functional characterization of adipose-specific phospholipase A2 (AdPLA). *J. Biol. Chem.* **283**, 25428–25436 (2008).
- Uyama, T., Jin, X. H., Tsuboi, K., Tonai, T. & Ueda, N. Characterization of the human tumor suppressors TIG3 and HRASLS2 as phospholipid-metabolizing enzymes. *Biochim. Biophys. Acta* **1791**, 1114–1124 (2009).
- Uyama, T. et al. The tumor suppressor gene H-Rev107 functions as a novel Ca<sup>2+</sup>-independent cytosolic phospholipase A1/2 of the thiol hydrolase type. *J. Lipid Res.* **50**, 685–693 (2009).
- Jaworski, K. et al. AdPLA ablation increases lipolysis and prevents obesity induced by high-fat feeding or leptin deficiency. *Nat. Med.* **15**, 159–168 (2009).
- Hussain, I. & Garg, A. Lipodystrophy syndromes. *Endocrinol. Metab. Clin. North Am.* **45**, 783–797 (2016).
- Jéru, I. Genetics of lipodystrophy syndromes. *Presse Med.* **50**, 104074 (2021).
- Brown, R. J. et al. The diagnosis and management of lipodystrophy syndromes: a multi-society practice guideline. *J. Clin. Endocrinol. Metab.* **101**, 4500–4511 (2016).
- Letourneau, L. R. & Greeley, S. A. W. Congenital forms of diabetes: the  $\beta$ -cell and beyond. *Curr. Opin. Genet. Dev.* **50**, 25–34 (2018).
- Sollier, C. et al. Lipodystrophic syndromes: from diagnosis to treatment. *Ann. Endocrinol. (Paris)* **81**, 51–60 (2020).
- Plagnol, V. et al. A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. *Bioinformatics* **28**, 2747–2754 (2012).
- Raman, L., Dheedene, A., De Smet, M., Van Dorpe, J. & Menten, B. WisecorDX: improved copy number detection for routine shallow whole-genome sequencing. *Nucleic Acids Res.* **47**, 1605–1614 (2019).
- Sobreira, N., Schiettecatte, F., Valle, D. & Hamosh, A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Hum. Mutat.* **36**, 928–930 (2015).
- Hummasti, S., Hong, C., Bensinger, S. J. & Tontonoz, P. HRASLS3 is a PPAR $\gamma$ -selective target gene that promotes adipocyte differentiation. *J. Lipid Res.* **49**, 2535–2544 (2008).
- Wang, Z. et al. LncPLAAT3-AS regulates PLAAT3-mediated adipocyte differentiation and lipogenesis in pigs through miR-503-5p. *Genes (Basel)* **14**, 161 (2023).
- Michaud, A. et al. Expression of genes related to prostaglandin synthesis or signaling in human subcutaneous and omental adipose tissue: depot differences and modulation by adipogenesis. *Mediators Inflamm.* **2014**, 451620 (2014).
- Morishita, H. et al. Organelle degradation in the lens by PLAAT phospholipases. *Nature* **592**, 634–638 (2021).

18. Wishart, D. S. et al. HMDB 5.0: the human metabolome database for 2022. *Nucleic Acids Res.* **50**, D622–d31 (2022).
19. Qin, Q. et al. Lisa: inferring transcriptional regulators through integrative modeling of public chromatin accessibility and ChIP-seq data. *Genome Biol.* **21**, 32 (2020).
20. Cataldi, S., Costa, V., Ciccociola, A. & Aprile, M. PPAR $\gamma$  and diabetes: beyond the genome and towards personalized medicine. *Curr. Diab. Rep.* **21**, 18 (2021).
21. Lefterova, M. I., Haakonsson, A. K., Lazar, M. A. & Mandrup, S. PPAR $\gamma$  and the global map of adipogenesis and beyond. *Trends Endocrinol. Metab.* **25**, 293–302 (2014).
22. Zhang, K., Chen, X., Zhang, P. & Liu, G. Perilipin2 is an earlier marker than perilipin1 for identifying adipocyte regeneration in fat grafts. *Aesthet. Surg. J.* **41**, Np646–Np652 (2021).
23. Olzmann, J. A. & Carvalho, P. Dynamics and functions of lipid droplets. *Nat. Rev. Mol. Cell Biol.* **20**, 137–155 (2019).
24. Kaushik, S. & Cuervo, A. M. Degradation of lipid droplet-associated proteins by chaperone-mediated autophagy facilitates lipolysis. *Nat. Cell Biol.* **17**, 759–770 (2015).
25. Ge, K. et al. Alternative mechanisms by which mediator subunit MED1/TRAP220 regulates peroxisome proliferator-activated receptor  $\gamma$ -stimulated adipogenesis and target gene expression. *Mol. Cell Biol.* **28**, 1081–1091 (2008).
26. Gorwood, J. et al. SIV infection and the HIV proteins tat and nef induce senescence in adipose tissue and human adipose stem cells, resulting in adipocyte dysfunction. *Cells* **9**, 854 (2020).
27. Ahmadian, M. et al. PPAR $\gamma$  signaling and metabolism: the good, the bad and the future. *Nat. Med.* **19**, 557–566 (2013).
28. Mory, P. B. et al. Atypical generalized lipodystrophy and severe insulin resistance due to a heterozygous LMNA p.T10I mutation. *Arq. Bras. Endocrinol. Metabol.* **52**, 1252–1256 (2008).
29. Gautheron, J. et al. Loss of thymidine phosphorylase activity disrupts adipocyte differentiation and induces insulin-resistant lipodystrophic diabetes. *BMC Med.* **20**, 95 (2022).
30. Payne, F. et al. Mutations disrupting the Kennedy phosphatidylcholine pathway in humans with congenital lipodystrophy and fatty liver disease. *Proc. Natl Acad. Sci. USA* **111**, 8901–8906 (2014).
31. Mann, J. P. & Savage, D. B. What lipodystrophies teach us about the metabolic syndrome. *J. Clin. Invest.* **129**, 4009–4021 (2019).
32. Barroso, I. et al. Dominant negative mutations in human PPAR $\gamma$  associated with severe insulin resistance, diabetes mellitus and hypertension. *Nature* **402**, 880–883 (1999).
33. Capel, E. et al. MFN2-associated lipomatosis: clinical spectrum and impact on adipose tissue. *J. Clin. Lipidol.* **12**, 1420–1435 (2018).
34. Sollier, C. et al. LIPE-related lipodystrophic syndrome: clinical features and disease modeling using adipose stem cells. *Eur. J. Endocrinol.* **184**, 155–168 (2021).
35. Gautheron, J. et al. EPHX1 mutations cause a lipodystrophic diabetes syndrome due to impaired epoxide hydrolysis and increased cellular senescence. *eLife* **10**, e68445 (2021).
36. Hernandez-Quiles, M., Broekema, M. F. & Kalkhoven, E. PPAR $\gamma$  in metabolism, immunity, and cancer: unified and diverse mechanisms of action. *Front. Endocrinol. (Lausanne)* **12**, 624112 (2021).
37. Kim, Y. G., Lou, A. C. & Saghatelian, A. A metabolomics strategy for detecting protein-metabolite interactions to identify natural nuclear receptor ligands. *Mol. Biosyst.* **7**, 1046–1049 (2011).
38. Civelek, E. & Ozen, G. The biological actions of prostanoids in adipose tissue in physiological and pathophysiological conditions. *Prostaglandins Leukot. Essent. Fatty Acids* **186**, 102508 (2022).
39. Hajnal, A., Klemenz, R. & Schäfer, R. Subtraction cloning of H-rev107, a gene specifically expressed in H-ras resistant fibroblasts. *Oncogene* **9**, 479–490 (1994).
40. Rochford, J. J. Mouse models of lipodystrophy and their significance in understanding fat regulation. *Curr. Top. Dev. Biol.* **109**, 53–96 (2014).
41. Le Lay, S., Magré, J. & Prieur, X. Not enough fat: mouse models of inherited lipodystrophy. *Front. Endocrinol. (Lausanne)* **13**, 785819 (2022).
42. King, E. A., Davis, J. W. & Degner, J. F. Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval. *PLoS Genet.* **15**, e1008489 (2019).
43. Nelson, M. R. et al. The support of human genetic evidence for approved drug indications. *Nat. Genet.* **47**, 856–860 (2015).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2023

**Nika Schuermans**<sup>1,2,32</sup>, **Salima El Chehadeh**<sup>3,4,5,32</sup>, **Dimitri Hemelsoet**<sup>6,32</sup>, **Jérémie Gautheron**<sup>7,32</sup>, **Marie-Christine Vantuyghem**<sup>8,9</sup>, **Sonia Nouioua**<sup>10,11</sup>, **Meriem Tazir**<sup>11,12</sup>, **Corinne Vigouroux**<sup>7,13</sup>, **Martine Auclair**<sup>7,13</sup>, **Elke Bogaert**<sup>1,2</sup>, **Sara Dufour**<sup>2,14,15</sup>, **Fumiya Okawa**<sup>16</sup>, **Pascale Hilbert**<sup>17</sup>, **Nike Van Doninck**<sup>18</sup>, **Marie-Caroline Taquet**<sup>19</sup>, **Toon Rosseel**<sup>1</sup>, **Griet De Clercq**<sup>1,2</sup>, **Elke Debackere**<sup>1,2</sup>, **Carole Van Haverbeke**<sup>20</sup>, **Ferroudja Ramdane Cherif**<sup>10,11</sup>, **Jon Andoni Urtizberea**<sup>21</sup>, **Jean-Baptiste Chanson**<sup>22</sup>, **Benoit Funalot**<sup>23,24</sup>, **François-Jérôme Authier**<sup>24,25</sup>, **Sabine Kaya**<sup>26</sup>, **Wim Terry**<sup>27</sup>, **Steven Callens**<sup>28</sup>, **Bernard Depypere**<sup>29</sup>, **Jo Van Dorpe**<sup>20</sup>, **Program for Undiagnosed Diseases (UD-PrOZA)**<sup>\*</sup>, **Bruce Poppe**<sup>1,2</sup>, **Francis Impens**<sup>2,14,15</sup>, **Noboru Mizushima**<sup>16</sup>, **Christel Depienne**<sup>4,26</sup>, **Isabelle Jéru**<sup>7,30,33</sup> & **Bart Dermaut**<sup>1,2,33</sup> ✉

<sup>1</sup>Center for Medical Genetics, Ghent University Hospital, Ghent, Belgium. <sup>2</sup>Department of Biomolecular Medicine, Faculty of Medicine and Health Sciences, Ghent University, Ghent, Belgium. <sup>3</sup>Service de Génétique Médicale, Institut de Génétique Médicale d'Alsace (IGMA), Hôpitaux Universitaires de Strasbourg, Strasbourg, France. <sup>4</sup>Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), INSERM U1258, CNRS-UMR7104, Université de Strasbourg, Strasbourg, France. <sup>5</sup>Laboratoire de Génétique Médicale, UMRS\_1112, Institut de Génétique Médicale d'Alsace (IGMA), Université de Strasbourg et INSERM, Strasbourg, France. <sup>6</sup>Department of Neurology, Ghent University Hospital, Ghent, Belgium. <sup>7</sup>Sorbonne Université, INSERM UMRS\_938, Centre de Recherche Saint-Antoine (CRSA), Paris, France. <sup>8</sup>Endocrinology, Diabetology, Metabolism Department, National Competence

Centre for Rare Diseases of Insulin Secretion and Insulin Sensitivity (PRISIS), Lille University Hospital, Lille, France.<sup>9</sup>University of Lille, INSERM U1190, European Genomic Institute for Diabetes, Lille, France.<sup>10</sup>Department of Neurology of the EHS of Cherrhell, University Centre of Blida, Tipaza, Algeria.<sup>11</sup>NeuroSciences Research Laboratory, University of Algiers Benyoucef Benkhedda, Algiers, Algeria.<sup>12</sup>Department of Neurology, CHU Algiers (Mustapha Pacha Hospital), Algiers, Algeria.<sup>13</sup>Assistance Publique—Hôpitaux de Paris, Saint-Antoine University Hospital, National Reference Center for Rare Diseases of Insulin Secretion and Insulin Sensitivity (PRISIS), Department of Endocrinology, Diabetology and Reproductive Endocrinology, and Department of Molecular Biology and Genetics, Paris, France.<sup>14</sup>VIB-UGent Center for Medical Biotechnology, VIB, Ghent, Belgium.<sup>15</sup>VIB Proteomics Core, VIB, Ghent, Belgium.<sup>16</sup>Department of Biochemistry and Molecular Biology, Graduate School and Faculty of Medicine, The University of Tokyo, Bunkyo, Japan.<sup>17</sup>Department of Molecular and Cellular Biology, Institute of Pathology and Genetics, Charleroi, Belgium.<sup>18</sup>Department of Endocrinology and Diabetology, General Hospital VITAZ, Sint-Niklaas, Belgium.<sup>19</sup>Department of Internal Medicine and Nutrition, Hopitaux Universitaires Strasbourg, Strasbourg, France.<sup>20</sup>Department of Pathology, Ghent University Hospital, Ghent, Belgium.<sup>21</sup>Institut de Myologie, Paris, France.<sup>22</sup>Service de Neurologie et Centre de Référence Neuromusculaire Nord/Est/Ile de France, Hôpital de Hautepierre, Strasbourg, France.<sup>23</sup>Department of Medical Genetics, Hôpital Henri Mondor, Université Paris-Est-Créteil, Créteil, France.<sup>24</sup>INSERM UMR955, Team Relaix, Faculty of Medicine, Créteil, France.<sup>25</sup>Centre Expert de Pathologie Neuromusculaire/Histologie, Département de Pathologie, Hôpital Henri Mondor, Université Paris-Est-Créteil, Créteil, France.<sup>26</sup>Institut für Humangenetik, Universitätsklinikum Essen, Essen, Germany.<sup>27</sup>Department of Nephrology, Jan Yperman Hospital, Ieper, Belgium.<sup>28</sup>Department of General Internal Medicine, Ghent University Hospital, Ghent, Belgium.<sup>29</sup>Department of Plastic and Reconstructive Surgery, Ghent University Hospital, Ghent, Belgium.<sup>30</sup>Department of Medical Genetics, DMU BioGeM, Sorbonne Université, AP-HP, Pitié-Salpêtrière Hospital, Paris, France.<sup>32</sup>These authors contributed equally: Nika Schuermans, Salima El Chehadeh, Dimitri Hemelsoet, Jérémie Gautheron.<sup>33</sup>These authors jointly supervised this work: Isabelle Jéru, Bart Dermaut. \*A list of authors and their affiliations appears at the end of the paper. ✉e-mail: [bart.dermaut@ugent.be](mailto:bart.dermaut@ugent.be)

### Program for Undiagnosed Diseases (UD-ProZA)

**Steven Callens<sup>28</sup>, Bart Dermaut<sup>1,2,33</sup>, Dimitri Hemelsoet<sup>6,32</sup>, Christel Depienne<sup>4,26</sup>, Nika Schuermans<sup>1,2,32</sup>, Wim Terry<sup>27</sup>, Arnaud V. Vanlander<sup>31</sup>, Patrick Verloo<sup>31</sup> & Paul J. Coucke<sup>1,2</sup>**

<sup>31</sup>Division of Pediatric Neurology and Metabolic Diseases, Department of Pediatrics, Ghent University Hospital, Ghent, Belgium.

## Methods

This research complies with all relevant ethical regulations and was approved by the Ethics Committee of the University Hospital of Ghent (EC:2019/1430). Data collection and analysis were not performed blind to the conditions of the experiment. No animals or data points have been excluded from the analyses.

### Patient clinical characteristics

All seven patients were clinically evaluated by an endocrinologist and a neurologist. Blood samples were taken to determine the levels of fasting glucose, HbA1c, insulin, cholesterol, TGs and leptin. Plasma insulin measurements were performed using the Roche ECLIA technology, and leptin was measured through a radioimmunoassay. Lipoatrophy and regional fat accumulation were objectified using MRI or computed tomography scan imaging. In two patients, DEXA was performed to evaluate body composition. In five patients, electromyography (EMG) was performed, and in four patients, brain MRI was performed. In one patient and one control, biopsy specimens of upper arm subcutaneous WAT were obtained for histopathological examination with hematoxylin and eosin (H&E) and anti-CD68 staining after informed consent of the patient and its caregiver. Quantification of the results was performed using ImageJ software<sup>7</sup>. Statistical analysis was done in SPSS Statistics 27. Written informed consent for multi-omics analysis and publication of clinical pictures was obtained from all patients and control individuals. For multi-omics analysis, we obtained upper arm subcutaneous WAT biopsies of patients 1 and 4, as well as six healthy, lean, female controls with ages between 25 and 60 years. The sample size was restricted due to the limited availability of WAT biopsies. No power analysis was performed.

### Genomics

Genetic studies in family 1 were initiated by the UD-ProZA team at Ghent University Hospital and included SNP-array-based homozygosity mapping, Sanger sequencing and cDNA analysis of *BSCL2*, WES, low-pass shallow WGS (CNV-seq) and WGS.

**Homozygosity mapping.** Genotyping for homozygosity mapping was performed in three affected and two unaffected members of family 1 using 200K genome-wide HumanCytoSNP-12 v2 BeadChip SNP arrays (Illumina). The position of the probes was based on the National Center for Biotechnology Information (NCBI) build GRCh37. Homozygous regions shared between affected members were detected using the PLINK algorithm (v1.07, default settings)<sup>44</sup>.

***BSCL2* mutation screening.** Sequencing of *BSCL2* was performed using Sanger sequencing after PCR amplification of all exons including splice junctions. Extracted mRNA from cultured lymphocytes was reverse-transcribed to cDNA. The region covering *BSCL2* exons 2–9 was amplified by PCR from this cDNA. The PCR product was migrated on an agarose gel, and its size was compared to that obtained from a control participant.

**WES.** WES was performed on the Illumina NovaSeq 6000 Platform after enrichment of gDNA with SureSelectXT Low Input Human All Exon v7 (Agilent Technologies). The BWA-MEM 0.7.17 algorithm was used for read mapping against the human genome reference sequence (NCBI, GRCh37.p5/hg19), duplicate read removal and variant calling. Variant calling and filtering were performed using Seqplorer, an in-house developed tool for the analysis of WES data. The position of the called variants is based on NCBI build GRCh38. A minimum of 90% of the interrogated genes had a coverage of >20×. Variant classification was performed according to the ACMG guidelines<sup>45</sup> and CNVs were detected with ExomeDepth<sup>41</sup>.

Molecular karyotyping was performed by means of low-pass WGS (CNV-seq) on an Illumina NovaSeq 6000 with a genome-wide resolution of 100 kb (GRCh38).

WGS was performed by MacroGen on an Illumina NovaSeq 6000 platform. FASTQ files were aligned against the hg38 reference genome with BWA-MEM (v0.7.17). Files were converted to BAM format, duplicate marked and sorted and indexed using Samtools (v1.9) (ref. 46) and Picard (v2.21.6). Structural variants were then called using three different callers, namely DELLY (v0.8.3) (ref. 47), LUMPY (v0.2.13) (ref. 48) and Manta (v1.6.0) (ref. 49). Standard settings were used for all.

In family 2, genetic analysis was initiated by a clinical geneticist at Hôpital de Hautepierre in Strasbourg and included the analysis of neuropathy and lipodystrophy gene panels, WES and Sanger sequencing. Genetic studies in families 3 and 4 were initiated at Hôpital Pitié-Salpêtrière in Paris and consisted of the analysis of a gene panel including genes involved in lipodystrophic syndromes and WES followed by Sanger sequencing of the region encompassing the *PLAAT3* variant. WES was performed for the index case as part of the 'Neuromendeliome' study, including 26 patients with unsolved syndromes affecting the nervous system. Paired-end sequencing libraries were prepared using the Agilent SureSelectXT Human All Exon v7 Enrichment kit. Sequencing (2 × 100 bases) was performed on a HiSeq2500 (Illumina) on the GenomEast platform (IGBMC). Image analysis and base calling were performed using CASAVA v1.8.2 (Illumina). Reads were mapped onto the reference genome Hg19 using BWA (v0.7.5a) (ref. 50). Elimination of duplicate reads and base quality recalibration were performed using Picard (v1.122). Realignment around indels was performed with GATK (v3.2-2). Reads mapping to several positions in the genome were excluded using Samtools (v0.1.19) (ref. 46). Variant calling was done using GATK (v3.2-2) unified genotyper<sup>51</sup>. Variants were annotated using GATK v3.2-2 (ref. 51), SnpE\_v2.0.5 (ref. 52) and SnpSift v4.41 (ref. 53).

### Generation of *Plaat3*<sup>-/-</sup> mice

All mouse (C57BL/6J) experiments were approved by the Institutional Animal Care and Use Committee of the University of Tokyo (Medical-P17-084) and the Animal Care and Use Committee of the National Institute of Quantum and Radiological Science and Technology (1610111 and 1610121). *Plaat3*<sup>-/-</sup> mice (C57BL/6J) were intercrossed to obtain *Plaat3*<sup>+/-</sup> mice<sup>17</sup>. Genotyping of *Plaat3*<sup>+/-</sup> and *Plaat3*<sup>-/-</sup> mice was performed by PCR using primers flanking the target site (forward primer, 5'-GAAAGAAGCTGCTGTGCCATGTGGCC-3'; reverse primer, 5'-ACCTGATCACTCCGAGGAATCCATAGC-3') and subsequent acrylamide gel electrophoresis. Wild-type and mutated alleles were detected as 207-bp and 191-bp bands, respectively. Twenty-four inguinal WAT specimens were obtained from 6 male *Plaat3*<sup>-/-</sup>, 6 male *Plaat3*<sup>+/-</sup>, 6 female *Plaat3*<sup>+/-</sup> and 6 female *Plaat3*<sup>-/-</sup> at the age of 11 weeks. All mice were housed in a specific pathogen-free room maintained at a constant ambient temperature of 20–22 °C and 40–60% humidity under a 12-h light/dark cycle with free access to water and food. Mice were fed sterilized (30 kGy) CLEA Rodent Diet CE-2 standard pellet chow (CLEA Japan). No statistical methods were used to predetermine sample sizes, but our sample sizes are similar to those reported in previous publications<sup>17</sup>. Data collection and analysis were not performed blind to the conditions of the experiments.

### RNA sequencing

RNA sequencing was performed on 15 inguinal mouse WAT samples. Seven *Plaat3*<sup>+/-</sup> specimens (3 males and 4 females) were compared to 8 *Plaat3*<sup>-/-</sup> littermates (4 males and 4 females). Additionally, RNA sequencing was performed on an upper arm WAT sample of patient 1 and three healthy lean controls (females, between 30 and 58 years of age and of European Caucasian descent, BMI ranging between 19 and 22 kg m<sup>-2</sup>).



RNA paired-end sequencing was performed by Macrogen on an Illumina platform. Library preparation was performed using the SMARTer Universal Low Input RNA Kit (Takara Bio, 634936) and the TruSeq RNA Sample Prep Kit v2 (Illumina, RS-122-2101). Trimmed reads were mapped to the reference genome with HISAT2 (ref. 54). After read mapping, StringTie was used for transcript assembly<sup>55</sup>. The expression profile was calculated for each sample and transcript/gene as read count, fragment per kilobase of transcript per million mapped reads and transcripts per kilobase million. DEG analysis was performed on a comparison pair (test versus control) using edgeR<sup>56</sup>. Functional enrichment analysis of Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) pathways<sup>57</sup> within 100 most strongly down-regulated and upregulated genes was performed using the Metascape tool<sup>58</sup>. To identify potential TRs of the top 100 (mouse WAT) or 500 (human WAT) downregulated and upregulated genes, we performed LISA (<http://lisa.cistrome.org/>), which is designed to combine a comprehensive database of human and mouse DNase-seq, H3K27ac chromatin immunoprecipitation followed by sequencing (ChIP-seq) and TR ChIP-seq to identify TRs that regulate a query gene set<sup>19</sup>.

### Lipidomics and histopathology

Lipidomics analysis was performed on inguinal WAT of 12 *Plaat3*<sup>+/+</sup> (6 males and 6 females) and 12 *Plaat3*<sup>-/-</sup> (6 males and 6 females) mice and subcutaneous upper inner arm WAT specimens from two patients (patients 1 and 4) and 6 control individuals (all female, between 27 and 58 years of age and of European Caucasian descent, BMI ranging between 19 and 22 kg m<sup>-2</sup>) in collaboration with Lipometrix, the lipidomics core at KU Leuven in Belgium. The subcutaneous adipose tissue specimens were obtained through an open biopsy, and the site of sampling was identical for all patients and controls. After lipid extraction and sample normalization, hydrophilic interaction liquid chromatography (HILIC)-MS was performed, enabling the quantification of 1,800 lipid species across 18 different lipid classes (<https://www.lipometrix.be>).

Part of the mouse inguinal WAT biopsies were thawed and fixed in 10% neutral-buffered formalin for ~27 h, paraffin-embedded and stained with H&E. Light microscopic images were analyzed using ImageJ software<sup>9</sup>.

**Lipid extraction.** A sample containing 10 µg of protein was mixed with 800 µl 1 N HCl/CH<sub>3</sub>OH (1:8, vol/vol), 900 µl CHCl<sub>3</sub>, 200 µg ml<sup>-1</sup> of the antioxidant 2,6-di-tert-butyl-4-methylphenol (Sigma-Aldrich) and 3 µl of SPLASH LIPIDOMIX Mass Spec Standard (Avanti Polar Lipids, 330707). After vortexing and centrifugation, the lower organic fraction was collected and evaporated using a Savant SpeedVac SPD111V (Thermo Fisher Scientific) at room temperature and the remaining lipid pellet was stored at -20 °C under argon.

**MS.** Just before mass spectrometry analysis, lipid pellets were reconstituted in 100% ethanol. Lipid species were analyzed by liquid chromatography-electrospray ionization tandem mass spectrometry (LC-ESI/MS/MS) on a Nexera X2 UHPLC system (Shimadzu) coupled with a hybrid triple quadrupole/linear ion trap mass spectrometer (AB SCIEX; 6,500 + QTRAP system). Chromatographic separation was performed on an XBridge amide column (Waters; 150 mm × 4.6 mm, 3.5 µm) maintained at 35 °C using mobile phase A (1 mM ammonium acetate in water/acetonitrile (ACN) (5:95, vol/vol) and mobile phase B (1 mM ammonium acetate in water/ACN (50:50, vol/vol)) in the following gradient: (0–6 min: 0% B → 6% B; 6–10 min: 6% B → 25% B; 10–11 min: 25% B → 98% B; 11–13 min: 98% B → 100% B; 13–19 min: 100% B; 19–24 min: 0% B) at a flow rate of 0.7 ml min<sup>-1</sup>, which was increased to 1.5 ml min<sup>-1</sup> from 13 min onwards. Sphingomyelins (SM), cholesterol esters (CE), ceramides (CER), dihydroceramides (DCER), hexosylceramides (HCER) and lactosylceramides (LCER) were measured in positive ion mode with precursor scans of 184.1, 369.4, 264.4, 266.4,

264.4 and 264.4 respectively. Triacylglycerides (TAG), diacylglycerides (DAG) and monoacylglycerides (MAG) were measured in positive ion mode with a neutral loss scan for one of the fatty acyl moieties. Phosphatidylcholine (PC), lysophosphatidylcholine (LPC), phosphatidylethanolamine (PE), lysophosphatidylethanolamine (LPE), phosphatidylglycerol (PG), phosphatidylinositol (PI) and phosphatidylserine (PS) were measured in negative ion mode by fatty acyl fragment ions. Lipid quantification was performed by scheduled multiple reaction monitoring, and the transitions were based on neutral losses or typical product ions as described above. The instrument parameters were as follows: curtain gas = 35 psi; collision gas = 8 a.u. (medium); ion spray voltage = 5,500 V and -4,500 V; temperature = 550 °C; ion source gas 1 = 50 psi; ion source gas 2 = 60 psi; declustering potential = 60 V and -80 V; entry potential = 10 V and -10 V; collision cell exit potential = 15 V and -15 V.

The following fatty acyl moieties were taken into account for the lipidomic analysis: 14:0, 14:1, 16:0, 16:1, 16:2, 18:0, 18:1, 18:2, 18:3, 20:0, 20:1, 20:2, 20:3, 20:4, 20:5, 22:0, 22:1, 22:2, 22:4, 22:5 and 22:6, except for TGs, which considered 16:0, 16:1, 18:0, 18:1, 18:2, 18:3, 20:3, 20:4, 20:5, 22:2, 22:3, 22:4, 22:5 and 22:6.

**Data analysis.** Peak integration was performed with the MultiQuant software version 3.0.3. Lipid species signals were corrected for isotopic contributions (calculated with Python Molmass 2019.1.1) and were quantified based on internal standard signals and adheres to the guidelines of the Lipidomics Standards Initiative (LSI, level 2 type quantification as defined by the LSI). Unpaired *t* test *P* values and false discovery rate (FDR)-corrected *P* values (using the Benjamini–Hochberg procedure) were calculated in Python StatsModels version 0.10.1.

### Proteomics

**Sample preparation.** Proteomics analysis was performed on inguinal WAT of 9 *Plaat3*<sup>+/+</sup> (4 males and 5 females) and 9 *Plaat3*<sup>-/-</sup> (4 males and 5 females) mice. Inguinal WAT tissue in 1.5 ml tubes was mechanically disrupted by two cycles of grinding with a disposable micropestle and freeze-thawing. Then, 250 µl lysis buffer containing 5% sodium dodecyl sulfate (SDS) and 50 mM triethylammonium bicarbonate (TEAB), pH 8.5, was added to the tissue, and the samples were ground once more with a micropestle to disrupt any remaining intact material. The tubes were spun down shortly, each sample was divided over three wells of a PIXUL 96-well plate (Active Motif) and samples were sonicated with a PIXUL Multisample sonicator (Active Motif) for 30 min with default settings (Pulse 50 cycles, pulse-repetition frequency (PRF) 1 kHz, burst rate 20 Hz). After centrifugation for 5 min at 2,204g at room temperature, the bottom (nonlipid) layer was transferred to new tubes and the protein concentration was measured by bicinchoninic acid assay (Thermo Fisher Scientific). From each sample, 100 µg of protein was isolated to continue the protocol. Proteins were reduced by the addition of 15 mM dithiothreitol and incubation for 30 min at 55 °C and then alkylated by the addition of 30 mM iodoacetamide and incubation for 15 min at room temperature in the dark. Phosphoric acid was added to a final concentration of 1.2%, and subsequently, samples were diluted sevenfold with a binding buffer containing 90% methanol in 100 mM TEAB, pH 7.55. The samples were loaded on a 96-well S-Trap plate (Protifi), placed on top of a deep well plate, and centrifuged for 2 min at 1,500g at room temperature. After protein binding, the S-trap plate was washed three times by adding 200 µl binding buffer and centrifugation for 2 min at 1,500g at room temperature. A new deep well receiver plate was placed below the 96-well S-Trap plate and 50 mM TEAB containing trypsin (1/100, wt/wt) was added for digestion overnight at 37 °C. Using centrifugation for 2 min at 1,500g, peptides were eluted three times, first with 80 µl 50 mM TEAB, then with 80 µl 0.2% formic acid (FA) in water and finally with 80 µl 0.2% FA in water/ACN (50:50, vol/vol). Eluted peptides were dried completely by vacuum centrifugation.

TMTpro 18-plex labels (Thermo Fisher Scientific, 0.5 mg) were equilibrated to room temperature immediately before use and dissolved in 20  $\mu$ l anhydrous ACN. The dried peptides were resuspended in 80  $\mu$ l 100 mM TEAB (pH 8.5), the peptide concentration was determined on a Lunatic spectrophotometer (Unchained Labs)<sup>59</sup> and the peptide amount was adjusted to 40  $\mu$ g for each sample. Peptides were labeled for 1 h at room temperature using 0.5 mg of TMTPro label (labels used: 135N, 132C, 131N, 132N, 134N, 133C, 133N, 134C, 131C; *Plaat3*<sup>+/+</sup> heterozygous replicates; 130C, 128N, 126C, 127C, 129C, 129N, 128C, 130N, 127N; *Plaat3*<sup>-/-</sup> homozygous replicates). The reaction was quenched for 15 min at room temperature by the addition of 4.2  $\mu$ l of 5% hydroxylamine. The 18 labeled samples were combined, and 100  $\mu$ g labeled peptides were isolated, dried by vacuum centrifugation, redissolved in 100  $\mu$ l 0.1% trifluoroacetic acid (TFA) and desalted on a reversed-phase (RP) C18 OMI tip (Agilent). The tip was first washed 3 times with 100  $\mu$ l prewash buffer (0.1% TFA in water/ACN (20:80, vol/vol)) and pre-equilibrated five times with 100  $\mu$ l of wash buffer (0.1% TFA in water) before the sample was loaded on the tip. After peptide binding, the tip was washed three times with 100  $\mu$ l of wash buffer, and peptides were eluted twice with 100  $\mu$ l elution buffer (0.1% TFA in water/ACN (40:60, vol/vol)). The combined elutions were dried in a vacuum concentrator.

Vacuum-dried peptides were redissolved in 100  $\mu$ l loading solvent A (0.1% TFA in water/ACN (98:2, vol/vol)), and 95  $\mu$ l was injected for fractionation by reverse-phased high performance liquid chromatography (RP-HPLC) (Agilent series 1200) connected to a Probot fractionator (LC Packings). Peptides were first loaded in solvent A on a 4 cm precolumn (made in-house, 250  $\mu$ m internal diameter (ID), 5  $\mu$ m C18 beads; Dr. Maisch) for 10 min at 25  $\mu$ l min<sup>-1</sup> and then separated on a 15 cm analytical column (made in-house, 250  $\mu$ m ID, 3  $\mu$ m C18 beads; Dr. Maisch). Elution was performed using a linear gradient from 100% RP-HPLC solvent A (10 mM ammonium acetate (pH 5.5) in water/ACN (98:2, vol/vol)) to 100% RP-HPLC solvent B (70% ACN, 10 mM ammonium acetate (pH 5.5)) in 100 min at a constant flow rate of 3  $\mu$ l min<sup>-1</sup>. Fractions were collected every minute between 20 and 92 min and pooled every 24 min to generate a total of 24 samples for LC-MS/MS analysis. All 24 fractions were dried under vacuum in HPLC inserts and stored at -20 °C until further use.

**LC-MS/MS analysis.** Each fraction was solubilized in 20  $\mu$ l loading solvent A (0.1% TFA in water/ACN (98:2, vol/vol)) moments before analysis. In total, 15  $\mu$ l of the sample measured on Dropsense16 (Unchained Labs) was injected for LC-MS/MS analysis with an Ultimate 3000 RSLCnano system in-line connected to an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific). Trapping was performed at 20  $\mu$ l min<sup>-1</sup> for 2 min in loading solvent A on a 20 mm trapping column (Thermo Fisher Scientific; 300  $\mu$ m ID, 5  $\mu$ m beads). The peptides were separated on a 110 cm prototype  $\mu$ PAC column (Thermo Fisher Scientific). It was kept at a constant temperature of 50 °C. Peptides were eluted by a linear gradient reaching 26.4% MS solvent B (0.1% FA in water/ACN (2:8, vol/vol)) after 45 min, 44% MS solvent B at 55 min, 56% MS solvent B at 60 min, followed by a 10-min wash at 56% MS solvent B and re-equilibration with MS solvent A (0.1% FA in water). In the first 15 min, the flow rate was set to 600 nl min<sup>-1</sup> after which it was kept constant at 300 nl min<sup>-1</sup>.

The mass spectrometer was operated in data-dependent mode. Full-scan MS spectra (375–1,500  $m/z$ ) were acquired at a resolution of 120,000 in the Orbitrap analyzer after accumulation to a target AGC value of 400,000 with a maximum injection time of 50 ms. The precursor ions were filtered for charge states (2–7 required), dynamic exclusion (60 s;  $\pm$ 10 ppm window) and intensity (minimal intensity of 5E4). The precursor ions were selected in the quadrupole with an isolation window of 0.7 Da and accumulated to an automatic gain control (AGC) target of  $1 \times 10^4$  or a maximum injection time of 50 ms and activated using collision-induced dissociation (CID) fragmentation (35%

normalized collision energy (NCE)). The fragments were analyzed in the Ion Trap Analyzer at the turbo scan rate. The ten most intense MS2 fragments were selected in the ion trap using MS3 multinotch isolation windows of 2  $m/z$ . An orbitrap resolution of 60k was used with an AGC target of  $1 \times 10^5$  or a maximum injection time of 118 ms and activated using higher energy collisional dissociation (HCD) fragmentation (65% NCE). QCloud was used to control instrument longitudinal performance during the project<sup>60</sup>.

**Data analysis.** LC-MS/MS runs of all 18 samples were searched together using the MaxQuant algorithm (version 2.1.3.0) with mainly default search settings, including an FDR set at 1% at the peptide and protein levels. Spectra were searched against the mouse protein sequences in the Swiss-Prot database (database release version of January 2022), containing 21,986 sequences ([www.uniprot.org](http://www.uniprot.org)). The mass tolerance for precursor and fragment ions was set to 4.5 and 20 ppm, respectively, during the main search. Enzyme specificity was set to the C-terminus of arginine and lysine, also allowing cleavage at Arg/Lys-Pro bonds with a maximum of two missed cleavages. Variable modifications were set to oxidation of methionine residues and acetylation of protein N-termini, whereas carbamidomethylation of cysteine residues was set as a fixed modification. Only proteins with at least one unique or razor peptide were retained, leading to the identification of 4,036 proteins. MS3-based quantification using TMTpro labels was chosen as the quantification method, and a minimum ratio count of two unique peptides was required for quantification. Further data analysis of the shotgun results was performed with an in-house R script using the proteinGroups output table from MaxQuant. Reverse database hits were removed, reporter intensities were log<sub>2</sub> transformed and median-normalized and replicate samples were grouped. Proteins with less than three valid values in at least one group were removed, and missing values were imputed from a normal distribution centered around the detection limit (DEP package), leading to a list of 4,033 quantified proteins in the experiment, used for further data analysis<sup>61</sup>. To compare protein abundance between pairs of sample groups, statistical testing for differences between two group means was performed using the limma package<sup>62</sup>. Statistical significance for differential regulation was set by a raw  $P < 0.05$  and FC ( $|\log_2(\text{FC})| \geq 0.5$ ). The mass spectrometry proteomics data have been deposited to the ProteomeXchange consortium via the PRIDE partner repository with the dataset identifier PXD038815.

### CRISPR-Cas9-mediated deletion of *PLAAT3*

The lentiviral plasmid plentiCRISPRv2 was a gift from the Zhang Lab (Addgene plasmid 52961) and contains hSpCas9, a guide RNA (gRNA), and a puromycin resistance sequence. The gRNA targeting exon 2 of *PLAAT3* was designed and checked for efficiency (<http://cistrome.org/SSC>) and specificity (<http://crispr.mit.edu>). The web-based tool CRISPOR (<http://crispor.tefor.net>) was used to evaluate potential off-target sequences. Subsequently, the gRNA was cloned into plentiCRISPRv2, and lentiviruses were produced by the VVTG platform (SFR Necker). ASCs were infected with viral particles at a minimal titer of  $10^8$  transducing units per ml. Forty-eight hours postinfection, the cells were selected with 5  $\mu$ g ml<sup>-1</sup> puromycin dihydrochloride (Sigma-Aldrich, P9620). Surviving cells were propagated, and the heterogeneous cell pool was used for experiments. The percentage of on-target recombination including insertions and deletions (indels) in the genomic DNA from the KO cells was evaluated by Sanger sequencing followed by analysis using the Synthego web-based tool (<https://ice.synthego.com>). The gRNA sequence used in this study is provided in Supplementary Table 7.

### Cell culture

Human ASCs were isolated from surgical samples of subcutaneous abdominal adipose tissue from a control woman with normal BMI. Adipose tissue samples were enzymatically digested with collagenase

B (0.2%). After centrifugation, the stromal vascular fraction was filtered, rinsed, plated and cultured in minimum essential medium- $\alpha$  ( $\alpha$ -MEM) with 10% fetal calf serum (FCS),  $2 \text{ mmol l}^{-1}$  glutamine, 1% penicillin-streptomycin (P/S) ( $10,000 \text{ UI ml}^{-1}$ ), 1% HEPES and fibroblast growth factor-2 ( $145 \text{ nmol l}^{-1}$ ). After 24 h, only ASCs adhered to plastic surfaces, while other cells were removed after culture medium replacement. ASCs were maintained in an undifferentiated state in high-glucose ( $4.5 \text{ g l}^{-1}$ ) DMEM supplemented with 10% newborn calf serum and 1% P/S. All culture conditions were kept constant throughout the experiments. ASC differentiation was induced as described previously<sup>29,35</sup>. Briefly, 2-d postconfluent cultures were treated with high-glucose ( $25 \text{ mmol l}^{-1}$ ) DMEM supplemented with 10% FCS, 1% P/S,  $1 \mu\text{mol l}^{-1}$  dexamethasone,  $1 \mu\text{mol l}^{-1}$  rosiglitazone,  $250 \mu\text{mol l}^{-1}$  3-isobutyl-1-methylxanthine (IBMX) and  $0.17 \mu\text{mol l}^{-1}$  insulin for 10 d. The medium was then replaced with high-glucose DMEM supplemented with 10% FCS, 1% P/S,  $1 \mu\text{mol l}^{-1}$  rosiglitazone and  $0.17 \mu\text{mol l}^{-1}$  insulin and changed to fresh medium every 2 d until the 20th day.

#### Oil Red-O staining, image processing and quantification

Intracellular lipids were stained with Oil Red-O (Sigma-Aldrich, O0625). Cells were washed with PBS and fixed with 4% PFA in PBS for 10 min. Fixed cells were incubated with Oil Red-O solution for 1 h at room temperature and then with DAPI (Thermo Fischer Scientific, D1306) for 5 min. Fluorescence images were generated with an IX83 Olympus microscope, acquired with Cell-Sens VL6 and analyzed with Fiji software. Images of 8–10 different areas per condition were visualized by fluorescence microscopy using mCherry and DAPI filters, followed by computer image analysis using Fiji software. Briefly, analysis was performed by threshold converting the 8-bit red–green–blue image into a binary image, which consists only of pixels representing lipid droplets (that is, red). Notably, after separation, the binary image was manually compared with the original image for consistency and correct binary conversion. The area occupied by lipid droplets in the image was displayed by Fiji software as surface area in  $\mu\text{m}^2$  and normalized to cell number by semiautomated counting of DAPI-stained nuclei.

#### Quantification of intracellular TG content

Intracellular lipids were extracted from differentiated ASC using hexane/isopropyl alcohol (3:2). Cells were washed and incubated with hexane/isopropyl alcohol (3:2, vol/vol) using 500  $\mu\text{l}$  per well in six-well culture plates on a shaker ( $80 \text{ r.p.m. min}^{-1}$ ) at room temperature for 60 min. The content of each well was then transferred into a glass tube for nitrogen evaporation of the organic solvent. After evaporation, the lipids were resuspended in isopropyl alcohol and transferred into duplicate 96-well plates for analysis after drying. TGs were measured using the Infinity Triglyceride kit (Thermo Fischer Scientific) according to the manufacturer's instructions. The absorbance of each well was measured using a Tecan microplate reader (TECAN) and converted to concentration based on a standard curve. The results were normalized to the cell protein content.

#### Western blot

Cells were homogenized in NP-40 lysis buffer to obtain protein lysates. Thirty micrograms of protein extracts were separated by SDS–polyacrylamide gel electrophoresis, transferred to a polyvinylidene difluoride membrane and analyzed by immunoblotting. Adipose tissues were dissociated and homogenized with ceramic beads and NP-40 lysis buffer using a Bead Ruptor (OMNI International, 19-042E). Western blot quantification was performed in triplicate using Fiji software (Open source), and the results were normalized to the tubulin, ERK or AKT protein levels.

#### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

#### Data availability

All relevant data generated and analyzed in this study are included in the article. Mouse RNA-seq data have been deposited in the Gene Expression Omnibus at NCBI ([GSE233433](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE233433)). The mouse mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier [PXD038815](https://www.ebi.ac.uk/pride/archive/study/MSD000000000). The human and mouse lipidomics data are available as Supplementary Data 1–3 ([Plaatz mouse lipidomics.html](#); [Plaatz mouse mediator lipidomics.html](#); [PLAAT3 human lipidomics.html](#)). For reasons of privacy, clinical patient sequencing data are not publicly available. Source data are provided with this paper.

#### Code availability

All multi-omics-related software used in this study are published and cited either in the main text or Methods. No custom code was used for data processing or analysis of the transcriptomics, proteomics or lipidomics datasets. Data analysis approaches using published software packages are described in the Methods and Supplementary Notes. For patients' privacy reasons, the Seqplorer codes for variant calling and filtering won't be publicly available.

#### References

- Slifer, S. H. PLINK: key functions for data analysis. *Curr. Protoc. Hum. Genet.* **97**, e59 (2018).
- Richards, S. 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.* **17**, 405–424 (2015).
- Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
- Rausch, T. et al. DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics* **28**, i333–i339 (2012).
- Layer, R. M., Chiang, C., Quinlan, A. R. & Hall, I. M. LUMPY: a probabilistic framework for structural variant discovery. *Genome Biol.* **15**, R84 (2014).
- Chen, X. et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics* **32**, 1220–1222 (2016).
- Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
- DePristo, M. A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).
- Cingolani, P. et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* **6**, 80–92 (2012).
- Cingolani, P. et al. Using *Drosophila melanogaster* as a model for genotoxic chemical mutational studies with a new program, SnpSift. *Front. Genet.* **3**, 35 (2012).
- Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
- Pertea, M. et al. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).
- Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
- Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. & Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* **40**, D109–D114 (2012).

58. Zhou, Y. et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).
59. Maia, T. M. et al. Simple peptide quantification approach for MS-based proteomics quality control. *ACS Omega* **5**, 6754–6762 (2020).
60. Chiva, C. et al. QCloud: a cloud-based quality control system for mass spectrometry-based proteomics laboratories. *PLoS ONE* **13**, e0189209 (2018).
61. Zhang, X. et al. Proteome-wide identification of ubiquitin interactions using UbiA-MS. *Nat. Protoc.* **13**, 530–550 (2018).
62. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

### Acknowledgements

The authors would like to thank the patients and families who participated in this study and the GenomEast facility (IGBMC) for exome sequencing in patient 4. We thank M. Baetens for in-depth CNV-seq analysis. We also thank L. Müller and P. Pellet for their technical help in genetic analyses in families 3 and 4. J.G. is funded by the Fondation pour la Recherche Médicale (FRM; ARF20170938613 and EQU202003010517), the Société Francophone du Diabète (SFD; R19114DD), the Mairie de Paris (Emergences—R18139DD) and the Agence Nationale de la Recherche (ANR-21-CE18-0002-01). B.D. is supported by an Odysseus type 1 Grant of the Research Foundation Flanders (G0H8318N) and a starting grant from Ghent University Special Research Fund (01N10319). N.M. is supported by the Exploratory Research for Advanced Technology (ERATO) research funding program of the Japan Science and Technology Agency (JPMJER1702) and a Grant-in-Aid for Specially Promoted Research from the Japan Society for the Promotion of Science (22H04919). The Program for Undiagnosed Diseases (UD-ProZA) is supported by the Spearhead Research Policy Program and the Fund for Innovation from the Ghent University Hospital. The Neuromendeliome Study (to C.D., Strasbourg, France) was financially supported by Agence de la Biomédecine (France). C.V. and M.A. are supported by institutional funding from Inserm, Sorbonne Université, Assistance-Publique Hôpitaux de Paris, by the Fondation pour la Recherche Médicale (grant EQU201903007868), and by the Association Française des Lipodystrophies (AFLIP), through a donation to Association Robert-Debré pour la Recherche Médicale (ARDRM). D.H. would like

to thank S. van Sprang and the whole team at the Academia Belgica, Center for History, Arts and Sciences in Rome, Italy (<https://www.academibelgica.it>), for supporting the writing of this manuscript. The authors of this study are members of the European Reference Network for Rare Neurological Diseases (ERN-RND to D.H. and B.D.), the Solve-RD Consortium (N.S., D.H., B.P. and B.D.) and the European Reference Network on Rare Endocrine Conditions (Endo-ERN, Project ID 739527 to C.V.). For more information about the ERNs and the EU health strategy, visit <http://ec.europa.eu/health/ern>. For more information about Solve-RD, visit <https://solve-rd.eu>.

### Author contributions

N.S., D.H., J.G., S.E.C., C.D., I.J. and B.D. wrote the manuscript. N.S., J.G., I.J. and B.D. designed the study and performed the main analyses. N.S., S.E.C., D.H., M.V., S.N., M.T., C.V., N.V.D., F.R.C., J.A.U., J.C., W.T., B.D. and S.C. were involved in phenotyping and clinical follow-up of the patients. N.S., E.B., T.R., G.D., E.D., B.F., F.A., S.K., P.H., C.D., B.P., I.J. and B.D. were involved in genotyping of the patients. S.D. and F.I. performed the proteomics analysis. J.G. designed the CRISPR-Cas9-mediated PLAAT3 KO cellular model. M.A. provided technical support for the cell experiments. N.S. and B.D. collected adipose tissue biopsy specimens. F.O. and N.M. provided WAT from *Plaat3<sup>-/-</sup>* and *Plaat3<sup>+/+</sup>* mice. J.V.D. and C.V.H. performed histological analyses of human and mouse WAT. N.S., D.H., J.G., I.J. and B.D. edited the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41588-023-01535-3>.

**Correspondence and requests for materials** should be addressed to Bart Dermaut.

**Peer review information** *Nature Genetics* thanks Robert Semple, David Savage and Rebecca Brown for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

Salima EL CHEHADEH-DJEBBAR

## Exploration de familles non-résolues avec déficience intellectuelle présumée liée à l'X à l'ère du *Whole Genome Sequencing*

### Résumé

La déficience intellectuelle (DI) concerne 2% de la population et représente la première cause de consultation dans les centres de génétique pédiatriques. Plus de la moitié des causes de DI sont génétiques, avec près de 1500 gènes impliqués, dont 127 gènes situés sur le chromosome X, et 200 syndromes recensés. On estime qu'environ 21% des syndromes avec DI liée à l'X (DILX) restent irrésolus sur le plan moléculaire. Des gènes de DILX et des mécanismes mutationnels encore non décrits sur l'X restent donc probablement à identifier. Ce travail de doctorat a en partie consisté à séquencer le génome (WGS) de patients atteints de DI présumée liée à l'X dont les analyses génétiques antérieures étaient négatives, ce qui a permis de contribuer à la validation de nouveaux gènes candidats de DILX, tel que le gène *SLITRK2*. Il souligne les difficultés rencontrées lors de l'interprétation des données de WGS, en particulier l'identification des variants non-codants et le très grand nombre de variants à analyser. De plus, l'interprétation des variants faux-sens sur des gènes de l'X est délicate en l'absence de ségrégation familiale concluante et d'étude fonctionnelle disponible.

**Mots clés :** Déficience intellectuelle liée à l'X, séquençage de génome entier, gène *SLITRK2*.

### Résumé en anglais

Intellectual disability (ID) affects 2% of the population and is the leading cause of consultation in pediatric genetics centers. More than half the causes of ID are genetic, with almost 1,500 genes involved, including 127 genes on the X chromosome, and 200 syndromes identified. It is estimated that around 21% of X-linked ID (XLID) syndromes remain molecularly unsolved. Therefore, XLID genes and mutational mechanisms that have yet to be described, probably remain to be identified. Part of this PhD work consisted in performing whole genome sequencing (WGS) of patients with presumed XLID whose previous genetic analyses had been negative. This work contributed to the validation of new DILX candidate genes, such as the *SLITRK2* gene. It also highlighted the difficulties encountered when interpreting WGS data, in particular regarding the identification of non-coding variants and the very large number of variants to be analyzed. In addition, the interpretation of missense variants on X-linked genes is tricky in the absence of conclusive familial segregation and available functional studies.

**Key words:** X-linked intellectual deficiency, whole genome sequencing, *SLITRK2* gene