Identification de nouveaux gènes impliqués dans la survenue d'hémorragies intracrâniennes fœtales

Thibault Coste

Université Paris Cité

École doctorale Bio Sorbonne Paris Cité (ED 562)

Maladies neurodéveloppementales et neurovasculaires – Neurodiderot (UMR-S 1141)

Identification de nouveaux gènes impliqués dans la survenue d'hémorragies intracrâniennes fœtales

Par **Thibault COSTE**

Thèse de doctorat de Génétique

Dirigée par **Elisabeth TOURNIER-LASSERVE**
Et co-encadrée par **Anne-Louise LEUTENEGGER**

Présentée et soutenue publiquement le 24/10/2023

Devant un jury composé de :

Judith MELKI, Professeure Émérite, Université Paris Saclay, Rapportrice
Jean-Marie JOUANNIC, PU-PH, Sorbonne Université, Rapporteur
Nadia BAHI-BUISSON, PU-PH, Université Paris Cité, Examinatrice
Ferdinand DHOMBRES, MCU-PH, Sorbonne Université, Examinateur
Fabien GUIMIOT, MCU-PH, Université Paris Cité, Examinateur
Elisabeth TOURNIER-LASSERVE, PU-PH, Université Paris Cité, Directrice de thèse
Anne-Louise LEUTENEGGER, CR Inserm, Université Paris Cité, Membre invitée et Co-encadrante

Résumé

Identification de nouveaux gènes impliqués dans la survenue d'hémorragies intracrâniennes fœtales

Les hémorragies intracrâniennes (HIC) fœtales ont une prévalence estimée de 1 pour 10 000 grossesses et conduisent très souvent à une interruption médicale de grossesse. Il en existe des causes acquises, dont la principale est l'allo-immunisation plaquettaire fœto-maternelle, et des causes génétiques. Les variations pathogènes des gènes *COL4A1* et *COL4A2* sont la cause génétique la plus connue. Toutefois, la négativité du bilan génétique chez près de 80% des fœtus ne permet ni conseil génétique ni prise en charge thérapeutique adaptée pour éviter une récidive. L'objectif principal de cette thèse était l'identification des gènes manquants. Ce travail a été effectué dans le cadre d'un projet collaboratif financé par le NIH (R01NS096173) et approuvé par le comité éthique de l'INSERM (IRB00003888). Il s'est appuyé sur les données de séquençage d'exome entier d'une cohorte de 113 fœtus sans variation pathogène de *COL4A1* ou *COL4A2* dont 35 trios, deux quartets et une famille de six individus. Ces fœtus avaient été initialement référés dans le service de génétique neurovasculaire de l'hôpital Saint-Louis pour une suspicion de collagénopathie COL4A1/A2 dans un contexte d'HIC, ventriculomégalie, porencéphalie et/ou schizencéphalie. Un examen systématique des dossiers médicaux a été effectué afin d'exclure les fœtus présentant des causes identifiables acquises ou des facteurs de risque connus conduisant à une HIC. Plusieurs stratégies d'analyse des données d'exome entier avec des critères de filtres stricts ont été conduites, incluant une analyse sans a priori des variants apparus *de novo* et des variants transmis selon un mode récessif, une analyse basée sur des listes de gènes candidats comme ceux impliqués dans la mégacaryopoïèse et enfin une approche par test de charge en variants rares conduite sur l'ensemble de la cohorte.

Plusieurs variants pathogènes bialléliques ont été identifiés dans les gènes *COQ2, ESAM, PROC* et *ATP5PO.* Des variants causaux dans deux gènes impliqués dans la mégacaryopoïèse, les gènes *MPL* et *MECOM*, ont été mis en évidence dans deux familles à cas multiples d'HIC. D'autres gènes comme *DCAF5, EXD3, ITGA2B, NAXD, RNPEP, TIE1 et ZFYVE28* sont des candidats forts pour lesquels des explorations complémentaires sont nécessaires. Enfin, un test de charge en variants rares nous a permis d'identifier des variants causaux dans le gène *PDHA1*, gène impliqué dans le métabolisme mitochondrial, chez trois fœtus non apparentés dont le tableau radiologique et fœtopathologique avait fait suspecter une collagénopathie COL4A1/A2.

Notre étude montre la très grande hétérogénéité génétique des fœtus pour lesquels une HIC est suspectée et l'intérêt du criblage pangénomique dans cette indication. Elle montre également que des variants pathogènes dans des gènes impliqués dans le métabolisme mitochondrial peuvent entrainer des phénotypes mimant ceux rencontrés chez les fœtus porteurs de variants pathogènes de *COL4A1/COL4A2*, renforçant encore l'indication d'un séquençage exome entier chez les fœtus avec HIC. De nombreux fœtus restent toutefois encore sans anomalie identifiée, soulevant plusieurs questions qui seront l'objet d'explorations dans un futur proche.

Mots clés : hémorragie intracrânienne fœtale, hémorragie intracérébrale, séquençage exome entier, maladies monogéniques, test de charge, collagénopathie COL4A1/A2

Abstract

Identification of new genes associated with fetal intracerebral hemorrhage

Fetal intracranial hemorrhages (ICH) have an estimated prevalence of 1 per 10,000 pregnancies, and in most cases lead to medical termination of pregnancy. There are both acquired causes, the main one being fetal and neonatal alloimmunization thrombocytopenia, and genetic causes. Pathogenic variations in *COL4A1* and *COL4A2* genes are the best-known genetic cause. However, the negativity of molecular screening in almost 80% of fetuses makes impossible genetic counseling nor appropriate therapeutic management to prevent recurrence. The main objective of this thesis was to identify the missing genes. This work was carried out as part of a collaborative project funded by the NIH (R01NS096173) and approved by the INSERM ethics committee (IRB00003888). It was based on whole exome sequencing data from a cohort of 113 fetuses, including 35 trios, two quartets and one family of six individuals. These fetuses were initially referred to the neurovascular genetics department of Saint-Louis Hospital (Paris, France) for a suspected COL4A1/A2 collagenopathy in a context of ICH, ventriculomegaly, porencephaly and/or schizencephaly. A systematic review of medical charts was performed to exclude fetuses with identifiable acquired causes or known risk factors leading to ICH. Several strategies for analyzing whole exome data with stringent criteria were conducted, including an unbiased analysis of *de novo* and recessively inherited variants, an analysis based on lists of candidate genes such as those involved in inherited platelet disorders, and finally a gene-based burden test approach conducted on the entire cohort.

Several biallelic pathogenic variants were identified in *COQ2, ESAM, PROC* and *ATP5PO* genes. Causal variants in two genes involved in megakaryopoiesis, including *MPL* and *MECOM*, were identified in two families with multiple cases of ICH. Other genes such as *DCAF5, EXD3, ITGA2B, NAXD, RNPEP, TIE1* and *ZFYVE28* are strong candidates for further investigations. Finally, gene-based burden test allowed us to identify causal variants in *PDHA1*, a gene involved in mitochondrial metabolism, in three unrelated fetuses whose cerebral lesions had led to suspicion of COL4A1/A2 collagenopathy.

Our study highlights the strong genetic heterogeneity of fetuses with suspected ICH, and the value of genome-wide screening in this condition. It also shows that pathogenic variants in genes involved in mitochondrial metabolism can lead to phenotypes mimicking those seen in fetuses carrying pathogenic *COL4A1/COL4A2* variants, further strengthening the indication for whole exome sequencing in fetuses with ICH. However, many fetuses remain without any identified genetic anomaly, raising several questions that will be explored in the near future.

Keywords: fetal intracranial hemorrhage, intracerebral hemorrhage, whole exome sequencing, monogenic diseases, gene-based burden test, COL4A1/A2 collagenopathy

Remerciements

Tout d’abord, je tiens à remercier vivement le Professeur Judith Melki, le Professeur Jean-Marie Jouannic, le Professeur Nadia Bahi-Buisson, le Docteur Fabien Guimiot et le Docteur Ferdinand Dhombres d’avoir accepté d’évaluer mon travail au travers de leur expertise et de me faire l’honneur de participer à ce jury de thèse.

J’adresse mes remerciements les plus profonds à ma directrice de thèse, le Professeur Elisabeth Tournier-Lasserve, pour m’avoir accompagné durant toutes ces années. Ses conseils avisés et ses connaissances sont une source d'inspiration permanente, m'incitant à aller toujours plus loin dans ma démarche scientifique et médicale.

Je remercie infiniment le Docteur Anne-Louise Leutenegger d’avoir accepté de co-encadrer ce travail, pour tous ses conseils et remarques constructives qui m’ont permis d’avancer dans ma réflexion scientifique.

Ce travail n'aurait jamais été possible sans le soutien permanent de mes collègues INSERM et hospitaliers, c'est pourquoi je tiens à leur exprimer mes plus sincères remerciements. Je remercie particulièrement Chaker pour sa contribution, ses commentaires constructifs et suggestions qui ont permis d’améliorer mon travail, ainsi que les autres membres de l’équipe Neurodiderot notamment Gwénola, Françoise, Minh, Margot, Evaelle, Sidonie et Jessica pour les conseils et échanges toujours très enrichissants.

Un grand merci à mes collègues biologistes, Hélène et Florence, avec qui j’ai le plaisir de travailler quotidiennement et qui ont toujours été d’un grand soutien. Je remercie également Michaelle qui a été d’une aide technique précieuse dans ce travail, merci à Estelle, Audrey, Nathalie, Valérie, Terry et Aïcha pour leur aide et bienveillance.

Ce travail de thèse a été possible grâce au soutien du NIH et il est le fruit d'un effort collectif, je souhaite ainsi remercier toutes les personnes qui ont contribué à ce projet au travers de collaborations. Je tiens également à remercier les familles de patients impliquées dans ce projet.

Enfin, je n'oublie pas de remercier mes amis proches et ma famille pour leur soutien indéfectible tout au long de ces années. Vos encouragements, votre confiance en moi et vos mots de soutien ont été un moteur puissant dans les moments de doute.

Liste des principales abréviations

ACMG : *American College of Medical Genetics and Genomics*

ACP : Analyse en Composante Principale

BHE : Barrière hémato-encéphalique

CNV : *Copy number variation* / Variation du nombre de copies

DNV : *De novo variant* / Variant de novo

ExAC *: Exome Aggregation Consortium*

GATK : *Genome Analysis Tool Kit*

gnomAD : *Genome aggregation database*

HBD : *Homozygous by descent /* homozygote par descendance

HIC : Hémorragie intracrânienne

HIV : Hémorragie intraventriculaire

Indel : Insertion / délétion

LoF : *Loss Of Function* / Perte de fonction

MAF : *Minor Allele Frequency* / Fréquence allélique de l'allèle alternatif

NGS : *Next generation sequencing* / Séquençage massif parallèle

RE : Réticulum endoplasmique

SA : Semaines d'aménorrhées

SNP : *Single nucleotide polymorphism* / Polymorphisme nucléotidique

SNV : *Single nucleotide variant* / Variation d'un seul nucléotide

WES : *Whole exome sequencing* / Séquençage de l'exome entier

WGS : *Whole genome sequencing* / Séquençage du génome entier

WT : *Wild-Type*

Table des matières

Liste des Figures

Liste des Tableaux

Table des Annexes

Introduction

A- Les hémorragies intracrâniennes fœtales

1) Définition et données épidémiologiques

Une hémorragie intracrânienne (HIC) chez le fœtus est une hémorragie qui se produit en période anténatale par rupture d'un vaisseau sanguin dans les ventricules, l'espace sous-dural et/ou le parenchyme cérébral. Alors que l'hémorragie néonatale chez le prématuré est relativement fréquente, puisqu'elle touche 15 à 20 % des enfants nés avant 32 semaines de grossesse, l'HIC fœtale est un évènement rare (Szpecht *et al.*, 2016).

Les études sur la prévalence et l'incidence des HIC chez le fœtus sont toutefois limitées et les données varient en fonction de divers facteurs tels que la population étudiée, les critères diagnostiques et les méthodes de détection utilisées. Certaines études font état d'une incidence entre 0,5 et 1 pour 1000 grossesses (Vergani *et al.*, 1996 ; Elchalal *et al.*, 2005), d'autres d'une incidence de 17 à 35 pour 100 000 naissances vivantes (Dicuonzo *et al.*, 2008).

2) Classification

Une HIC fœtale peut se produire dans les ventricules, le parenchyme cérébral ou les espaces sous-arachnoïdiens, sous-duraux ou épiduraux et peut toucher les espaces supratentoriels et/ou infratentoriels (Figure 1).

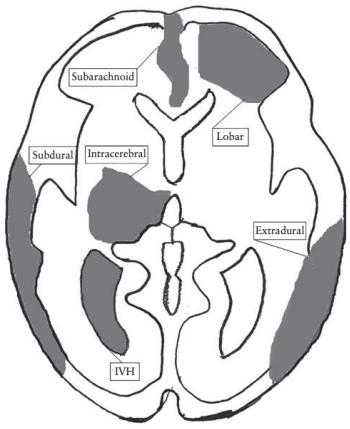


Figure 1. Schéma des différents types d'hémorragie intracrânienne. IVH = hémorragie intraventriculaire. Extrait de Sileo *et al.*, 2022.

Le sous-type le plus courant des HIC est l'hémorragie intraventriculaire (HIV), prenant naissance dans la matrice germinale périventriculaire, et pouvant rompre la ligne épendymaire pour gagner le système ventriculaire. Ces HIV sont subdivisées selon leur gravité en quatre grades (Figure 2), conformément à la classification couramment utilisée chez les nouveau-nés (Burstein *et al.*, 1979). Cette classification, modifiée par Ghi *et al.,* (2003), a également été proposée pour classer les HIC diagnostiquées pendant la période anténatale :

- Grade I : Hémorragie sous-épendymaire limitée à la matrice germinale

- Grade II : Hémorragie atteignant la lumière des ventricules latéraux sans dilatation ventriculaire

- Grade III : Hémorragie atteignant les ventricules latéraux associée à une dilatation des ventricules

- Grade IV : Hémorragie atteignant les ventricules latéraux et zone hyperéchogène dans le tissu périventriculaire (atteinte parenchymateuse adjacente)

Comme la gradation des HIC est reliée à leur sévérité, cette classification est largement utilisée en pratique clinique pour les décisions de prise en charge (Adiego *et al.*, 2019).

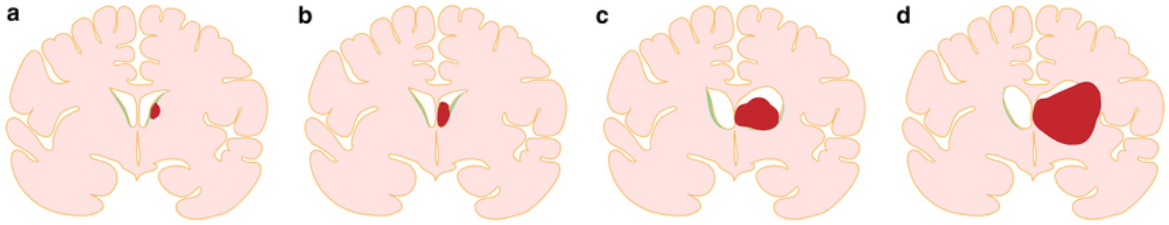


Figure 2. Les différents grades d'HIV. Grades I (a) à IV (d). Extrait de Luo *et al.*, 2019.

3) Diagnostic

Le diagnostic est posé le plus souvent par l'échographie qui permet d'identifier une ventriculomégalie et d'autres signes associées (Putbrese et Kennedy, 2016) tels que :

- Perte des repères cérébraux normaux

- Échogénicité variable, masse intracrânienne avasculaire

- Caillot aigu hyperéchogène adhérant à des plexus choroïdes volumineux/nodulaires

- Caillot hyperéchogène délimitant le cortex cérébral

- Épendyme nodulaire hyperéchogène
- Échogénicité accrue de la substance blanche périventriculaire
- Porencéphalie
- Hydranencéphalie

Un exemple d'HIC fœtale détectée à l'échographie du troisième trimestre associant ventriculomégalie et caillots dans les ventricules latéraux est illustré par la figure suivante (Figure 3).

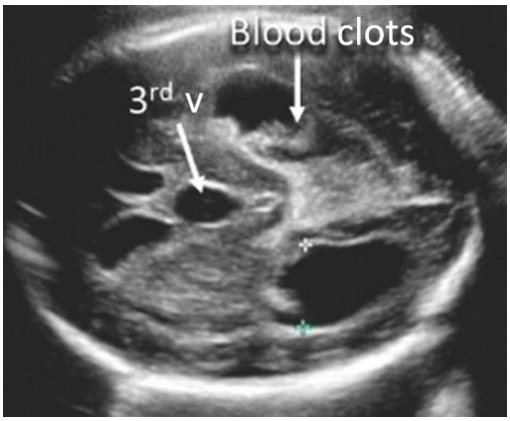


Figure 3. Exemple d'une HIC découverte à 35SA. Coupe axiale montrant une dilatation des ventricules latéraux et du troisième ventricule. Dans les ventricules latéraux, il y a présence d'un matériel échogène séparé des plexus choroïdes correspondant à des caillots sanguins. Extrait de Monteagudo, 2020.

L'imagerie par résonance magnétique (IRM) fœtale peut être utilisée comme complément pour le diagnostic lorsque des anomalies sont visualisées à l'échographie et peut avoir une plus grande précision diagnostique (Sanapo *et al.*, 2017). Dans certains cas, l'amélioration de l'évaluation par l'IRM fœtale par rapport à l'échographie permet de modifier le conseil parental prénatal et l'attitude thérapeutique chez le nouveau-né (Breysem *et al.*, 2003).

L'apport de l'IRM fœtale dans le diagnostic d'une HIC fœtale est illustré sur la figure suivante. Dans le cas présenté l'échographie initiale n'avait révélé qu'une ventriculomégalie unilatérale sur les clichés standard et la destruction parenchymateuse avec caillots et débris interne était masquée par un artéfact (Figure 4).

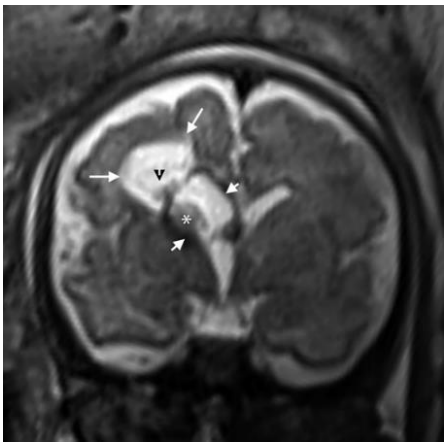


Figure 4. Exemple de porencéphalie sur IRM fœtale. IRM coronale T2 HASTE montrant une cavité porencéphalique (flèches longues) adjacente au ventricule latéral droit. La cavité contient des débris internes (pointe de flèche noire) et un caillot (astérisque) dans la corne frontale dilatée. Le dépôt d'hémosidérine provoque un artefact de déplacement chimique (flèches courtes), ce qui confirme que l'hémorragie intracrânienne est l'étiologie sous-jacente. Extrait de Putbrese et Kennedy, 2016.

4) Données cliniques

L'issue d'une grossesse avec HIC fœtale est généralement qualifiée de mauvaise avec un risque substantiel de décès du fœtus (Ghi *et al.*, 2003). Les conséquences exactes chez le fœtus peuvent varier en fonction de l'importance de l'hémorragie, de sa localisation dans le cerveau, du moment où elle se produit pendant la grossesse et d'autres facteurs individuels.

Les complications possibles chez un enfant ayant eu une HIC anténatale sont les suivantes :

- Une hydrocéphalie par une accumulation excessive de liquide dans les ventricules, exerçant une pression excessive sur le parenchyme cérébral nécessitant un drainage chirurgical
- Un retard de développement avec des retards dans les acquisitions motrices, cognitives et du langage, des troubles de l'apprentissage, des retards mentaux et des problèmes d'élocution
- Une déficience intellectuelle
- Un handicap neurologique tels qu'une infirmité motrice cérébrale ou paralysie cérébrale, des troubles de la coordination, une hémiplégie
- Une épilepsie
- Des séquelles ophtalmologiques

Les HIC de grade III et IV sont considérées comme sévères et sont un facteur de risque important de mortalité et de morbidité néonatales. Dans une méta-analyse Sileo et al., (2022) ont montré, dans le cas d'HIV fœtales de grades III et IV, que 52% des enfants présentaient un trouble neurodéveloppemental sévère et 43% une infirmité motrice cérébrale à la différence des enfants avec HIV de grades I et II (0 %). Devant ce pronostic sombre une interruption de grossesse est le plus souvent réalisée (Figure 5).

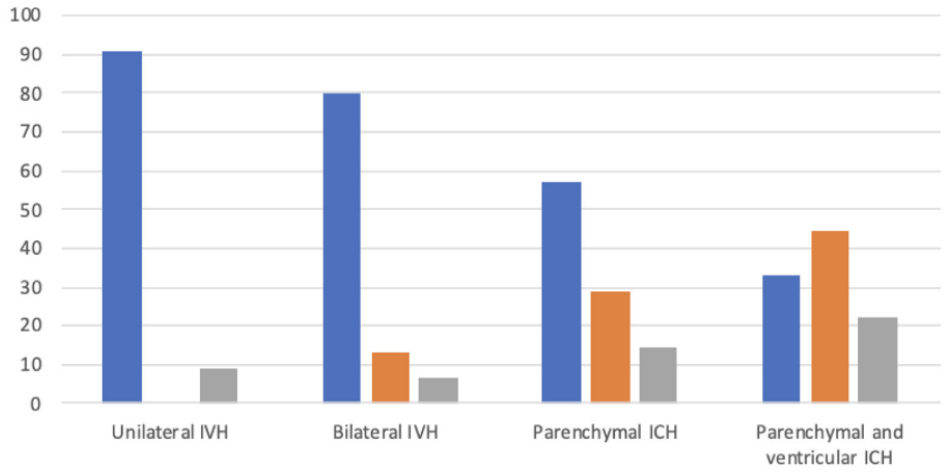


Figure 5. Issues de grossesses selon la localisation de l'HIC. Bleu = naissance, orange = IMG et gris = décès in utéro. Extrait de Gupta *et al.*, 2023.

B- Étiologies des HIC fœtales

1) Considérations physiopathologiques

Une HIC chez le fœtus se présente classiquement sous la forme d'une hémorragie de la matrice germinale avec hémorragie intraventriculaire. La matrice germinale sous-épendymaire est une couche transitoire hautement vascularisée située sur la face ventrolatérale des ventricules latéraux (Figure 6). Cette région est composée de vaisseaux endothéliaux fragiles, à risque d'hémorragies en raison d'une forte angiogenèse, d'une discontinuité de la barrière hémato-encéphalique, d'un faible nombre de péricytes et d'une lame basale immature. Le riche réseau microvasculaire se continue par un système veineux profond bien développé. Du fait de ces caractéristiques, la structure de la matrice germinale est ainsi vulnérable aux hémorragies. Par ailleurs, elle joue un rôle essentiel dans la production de neurones et de cellules gliales ; au troisième trimestre elle est la source de précurseurs gliaux qui deviennent des oligodendrocytes et des astrocytes. La matrice subit une diminution progressive de sa taille jusqu'à une involution presque complète entre 36 SA et le terme.

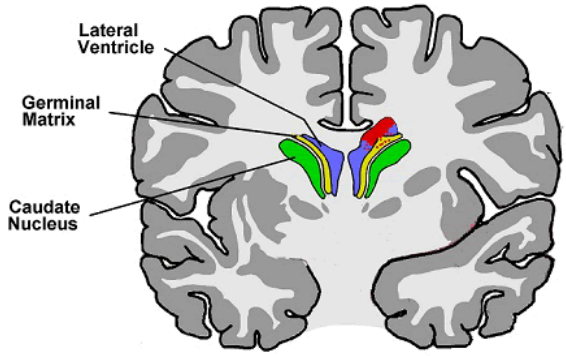


Figure 6. Localisation de la matrice germinale entre chaque ventricule latéral et le noyau caudé, avec illustration d'un saignement dans le ventricule droit. Extrait de Talbert *et al,* 2016.

Les hémorragies fœtales peuvent entraîner un infarctus veineux par l'occlusion des veines collectrices périventriculaires, la destruction des tissus qui s'ensuit et le développement d'une porencéphalie définie comme une cavité dans le parenchyme cérébral communiquant avec le ventricule latéral.

En cas d'hémorragie étendue précoce, la destruction peut s'étendre au cortex et conduire à des lésions destructrices majeures telles que l'hydranencéphalie.

2) Causes d'origine acquise

- **Thrombopénie par allo-immunisation fœtale ou néonatale (FNAIT)**

Son incidence est de 1/1000 naissances. Elle est la cause la plus fréquente d'HIC néonatale d'origine acquise, avec des HIC rapportées dans 10 à 30 % des cas (Toughza *et al.*, 2020). La majorité des HIC surviennent au cours des deuxième et troisième trimestres de grossesse. La mort fœtale intra-utérine est également relativement fréquente. Une hémorragie rétinienne peut être constatée après la naissance, avec une atrophie optique diagnostiquée lors du suivi.

Celle-ci résulte de l'immunisation maternelle contre les antigènes plaquettaires fœtaux hérités du père et absents chez la mère. Il existe au moins une trentaine d'antigènes plaquettaires humains (HPA) différents, et les hémorragies intracrâniennes périnatales sont plus fréquentes avec certains d'entre eux, notamment HPA1, HPA5b et HPA15b. L'antigène HPA1 est le plus courant chez les Caucasiens, et environ la moitié des HPA, y compris HPA1, sont des variants antigéniques de la sous-unité GPIIIa (intégrine sous-unité β3) (Lieberman *et al.,* 2019). L'allo-immunisation se produit si la mère est HPA négative (2 % de la population est HPA1 négatif c'est-à-dire homozygote pour l'allèle b (HPA1bb)) tandis que le père est HPA positif (HPA1ab ou HPA1aa), de sorte que le fœtus hérite d'au moins un allèle "a". Les allo-anticorps IgG maternels traversent le placenta et provoquent la destruction des plaquettes fœtales. La thrombopénie qui en résulte peut-être sévère (inférieure à 50 G/L) et à l'origine d'une HIC in utero ou au moment de l'accouchement. Par ailleurs, en utilisant des modèles murins, certains auteurs ont pu montrer qu'il existe également dans la FNAIT des preuves de perturbation de l'angiogenèse et de la fonction endothéliale (Yougbaré *et al.,* 2015). Le traitement actuel d'une FNAIT découverte au cours de la grossesse repose sur l'administration intraveineuse d'immunoglobulines associée à une surveillance échographique mensuelle à la recherche de signes d'une hémorragie cérébrale (Winkelhorst *et al.,* 2017).

Les recommandations pour le diagnostic biologique d'une FNAIT reposent sur un bilan d'allo-immunisation comprenant :

-Une mise en évidence dans le sérum maternel d'anticorps antiplaquettaires.

-Un crossmatch sur plaquettes paternelles.

-L'identification de l'antigène causal par typage HPA.

Tous ces tests nécessitent des compétences spécialisées de la part d'un laboratoire de référence et les résultats des examens FNAIT, en particulier les résultats négatifs, doivent être interprétés avec soin en tenant compte du contexte clinique (Petermann *et al.,* 2018).

- **Infections maternelles**

Il s'agit principalement de l'infection au cytomégalovirus (CMV) dont l'atteinte cérébrale peut se traduire par une microcéphalie, une hémorragie intraventriculaire, une ventriculomégalie, des lésions cérébelleuses, des calcifications périventriculaires, de nombreux kystes de germinolyse et une polymicrogyrie (Lazzarotto *et al.,* 2020). Le diagnostic de l'infection maternelle est basé sur la sérologie.

- **Grossesse monochoriale**

Le décès d'un jumeau au cours d'une grossesse monochoriale expose le jumeau survivant à un risque important de séquelles neurologiques. Le décès ultérieur du jumeau survivant survient dans 12% des cas. Des anomalies neurologiques du type hémorragie intracrânienne, ventriculomégalie, polymicrogyrie, kystes de germinolyse, peuvent être constatées par hypoperfusion intracrânienne du jumeau survivant au moment du décès du jumeau (Jelin *et al.,* 2008).

- **Déficit en vitamine K**

Pendant la grossesse, seul un tiers de la vitamine K présente dans le sang de la mère est transféré au fœtus, ce qui augmente le risque d'HIC fœtale si la mère prend certains médicaments qui interviennent dans le métabolisme de la vitamine K ou si elle souffre d'une maladie digestive avec malabsorption. Dans la plupart des cas rapportés, une période relativement longue de troubles de l'alimentation ou de troubles de l'absorption intestinale était à l'origine de la carence en vitamine K. Les recommandations incluent une supplémentation en vitamine K dans ces cas-là (Lee *et al.,* 2022).

- **Prises de drogues ou substances toxiques pour le fœtus**

Par exemple de cocaïne, de certains médicaments comme la Warfarine ou certains antiépileptiques (Ozduman *et al.,* 2004).

- **Traumatismes au niveau de l'abdomen**

3) Causes d'origine génétique

Après élimination des causes acquises, une origine génétique d'une HIC fœtale doit être suspectée. La découverte d'une cause génétique sous-jacente à une HIC fœtale est importante pour un conseil génétique adapté afin de prévenir une récurrence au sein de la même famille. De plus, l'identification des gènes impliqués est un apport considérable pour la compréhension des mécanismes sous-jacents aux HIC fœtales et le développement de nouvelles thérapeutiques.

Peu de données sont disponibles sur ce sujet et peu de gènes ont été rapportés comme étant responsables d'une HIC anténatale. Nous avons effectué une analyse de la littérature, sur la période allant de 1980 à juin 2023, par PubMed et EMBASE avec les mots clés « *fetal intracranial hémorrhage* », « *antenatal intracranial hemorrhage* », « *prenatal intracranial hemorrhage* » en combinaison avec les mots « *genetic* », « *gene* », « *exome* », « *molecular* » et avons identifié plusieurs gènes dont nous détaillons les caractéristiques dans les paragraphes suivants.

3-a) Gènes impliqués dans la coagulation

- **Généralités**

L'épidémiologie des déficits constitutionnels sévères en facteurs de la coagulation (hors hémophilie et maladie de Willebrand) est peu connue. Leur transmission est habituellement autosomique récessive (Tableau 1). La symptomatologie, c'est-à-dire la survenue des signes hémorragiques est variable et globalement corrélée avec le taux du facteur déficitaire.

Cliniquement, ces déficits rares sont caractérisés par un spectre large de signes hémorragiques allant de l'absence de symptômes à un syndrome hémorragique majeur.

Les profils cliniques des patients varient pour le même type de déficit, parfois même au sein d'une même famille. L'association entre la sévérité clinique et la sévérité biologique du déficit varie également selon le facteur concerné.

Les déficits en facteur de la coagulation en tant que cause d'une HIC sont connus chez les enfants et les adultes bien qu'à des fréquences très rares par rapport aux signes hémorragiques modérés. En revanche, on ne trouve dans la littérature que quelques études, principalement des cas isolés, qui traitent de l'association entre une HIC fœtale et les troubles de la coagulation. C'est le cas pour les déficits en facteur V, VII, X, XIII et pour la protéine C qui est un inhibiteur de la coagulation.

Facteur manquant	Prévalence	Corrélation biologique et clinique	Mode de transmission
Fibrinogène : Afibrinogénémie Hypofibrinogénémie Dysfibrinogénémie Hypodysfibrinogénémie	 1/1 000 000 Non connue Non connue Non connue	 Forte Forte Faible Faible	 Autosomique récessif Récessif ou dominant Récessif ou dominant Récessif ou dominant
Facteur II	1/2 000 000	Forte	Autosomique récessif
Facteur V	1/1 000 000	Faible	Autosomique récessif
Déficit combiné des facteurs V et VIII	1/1 000 000	Forte	Récessif
Facteur VII	1/500 000	Faible	Autosomique récessif
Facteur X	1/1 000 000	Forte	Autosomique récessif
Facteur XI	1/1 000 000	Nulle	Récessif ou dominant
Facteur XIII	1/2 000 000	Forte	Autosomique récessif
Déficit combiné des facteurs vitamino-K dépendants	Non connue	Faible	Autosomique récessif

Tableau 1. Caractéristiques des déficits rares en facteurs de la coagulation (tableau extrait du site de la filière des maladies hémorragiques constitutionnelles : https://mhemo.fr/)

- **Déficit en facteur V**

Le déficit héréditaire en FV, appelé également parahémophilie ou maladie d'Owren, est une maladie autosomique récessive très rare (prévalence de 1/1 000 000). Dans ce déficit, les tests courants de la coagulation (taux de prothrombine (TP) et temps de céphaline activée (TCA)) sont anormaux. Un dosage spécifique du facteur V est nécessaire pour confirmer le diagnostic. L'analyse moléculaire n'est pas nécessaire pour le diagnostic mais est utile pour le conseil génétique. Les patients homozygotes ou hétérozygotes composites ont en général des taux de FV plasmatiques inférieurs à 10 % de la normale, tandis que les porteurs hétérozygotes conservent des taux de FV plasmatique d'environ 50 % et sont généralement asymptomatiques.

Les manifestations courantes du déficit en facteur V sont des manifestations hémorragiques cutanéo-muqueuses (ecchymoses, épistaxis et saignements gingivaux), des hémorragies consécutives à des traumatismes mineurs et des ménorragies. Les hémorragies mettant en jeu le pronostic vital comme les HIC sont rarement observées chez les individus avec un déficit sévère en FV (Yang *et al.*, 2022).

La première description d'un déficit congénital en FV associé à une HIC concernait une petite fille qui avait présenté une hémorragie intraventriculaire à 32 semaines de grossesse (Whitelaw *et al.,* 1984). A la naissance elle présentait un taux de FV à 2 % mais le diagnostic moléculaire n'avait pas été effectué. Ses parents étaient probablement hétérozygotes avec des taux différents de FV (taux maternel de FV 52%, paternel 78%), tandis que son frère était apparemment normal avec une taux de FV de 100%.

Un autre cas d'HIC a été rapporté chez un fœtus qui présentait une anomalie du rythme cardiaque (Ellestad *et al.,* 2007). Une échographie à 36 semaines de grossesse montrait la présence de débris échogènes, une distorsion du parenchyme cérébral et un déplacement de la ligne médiane correspondant à une hémorragie cérébrale. Après césarienne en urgence, un bilan sanguin initial objectiva un taux de facteur V à 2,5%. Cet enfant reçut plusieurs lots de plasma frais congelé et bénéficia d'une chirurgie évacuatrice de l'hématome. Il décéda malheureusement quelques jours plus tard compte tenu de la sévérité de l'HIC. Les auteurs de ce cas soulignèrent à juste titre la question de la fréquence à laquelle un décès in utéro par HIC peut résulter d'un déficit non connu en facteur de la coagulation.

- **Déficit en facteur VII**

Le déficit héréditaire en FVII est une maladie de transmission autosomique récessive. C'est le plus fréquent des déficits rares et son expressivité clinique est variable avec des formes totalement asymptomatiques (30 % des patients) jusqu'à des tableaux extrêmement sévères. De plus, il semble exister une faible corrélation entre l'activité du FVII mesurée in vitro et les signes hémorragiques (Lou *et al.,* 2023). Une incidence accrue d'HIC a été signalée dans ce déficit ; les hémorragies surviennent le plus souvent chez des nourrissons et sont généralement attribuées à un traumatisme à la naissance. Pour ces tableaux sévères, il est souvent détecté des variations pathogènes de type perte de fonction sur les deux allèles du facteur VII (Giansily-Blaizot *et al.,* 2020).

Quelques cas d'HIC anténatales sont mentionnés dans la littérature. Mc Vey *et al.,* (1998) rapportèrent une famille dont le cas index consanguin décédé à l'âge de 12 jours présentait une hydrocéphalie, des caillots intraventriculaires et un œdème cérébral. Il avait une sœur décédée à l'âge d'un mois qui présentait également une hémorragie cérébrale et une hydrocéphalie. La temporalité de cette HIC n'était pas mentionnée (fœtale ou périnatale). L'activité du facteur VII était inférieure à 1 % chez les deux enfants et une variation homozygote dans un site d'épissage consensus du gène *F7* fut identifiée.

Enfin, Landau *et al.*, (2009) décrivirent deux cas de sexe féminin issus de parents apparentés. Le premier présentait en prénatal une hydrocéphalie et le second une HIC postnatale. Les deux avaient une activité du FVII inférieure à 1 %. Le criblage moléculaire révéla que les deux parents étaient hétérozygotes et les deux filles homozygotes pour une variation faux sens pathogène.

- **Déficit en facteur X**

Le déficit en FX est une pathologie de transmission autosomique récessive. Son prévalence est estimée à 1 / 1 000 000 d'habitants. Il est principalement observé dans les pays où la consanguinité est élevée. Le phénotype hémorragique est variable, les sujets atteints pouvant être asymptomatiques ou présenter des symptômes hémorragiques graves. Il existe une corrélation faible entre la symptomatologie hémorragique et la sévérité du déficit en FX.

Un seul cas d'HIC fœtale a été rapporté. Il s'agit d'un cas avec mise en évidence à l'échographie de 35 SA d'une collection liquidienne avec déplacement de la ligne médiane compatible avec une hémorragie sous-durale. Cette HIC fut confirmée à la naissance par un scanner cérébral montrant une importante hémorragie sous-durale gauche avec compression du ventricule latéral gauche. Un test de l'activité du facteur X retrouva un taux effondré (< 1%) chez le nouveau-né et diminué chez les parents présumés porteurs hétérozygotes (activité FX mère 56% et père 63%) (De Sousa *et al.*, 1988).

- **Déficit en facteur XIII**

Le facteur XIII de la coagulation (FXIII) est une protransglutaminase qui circule dans le plasma sous la forme la forme de deux sous-unités A potentiellement actives (FXIII-A) et de deux sous-unités B inhibitrices/transporteuses (FXIII-B). Le FXIII plasmatique activé joue un rôle essentiel dans l'étape finale de la cascade de la coagulation par stabilisation du caillot de fibrine. Le déficit congénital en FXIII est extrêmement rare avec une prévalence de 1/2 000 000 dans la population générale. La maladie est transmise selon un mode autosomique récessif et l'incidence des HIC dans le déficit en FXIII est d'environ 30 % des cas rapportés, ce qui est plus élevé que dans tout autre trouble congénital de la coagulation (Eshghi *et al.*, 2012).

Un cas d'une HIC anténatale a déjà été rapporté dans la littérature, il s'agit d'un fœtus issu d'une famille originaire du sud-est de l'Iran. Ce fœtus présentait une hydrocéphalie et un taux de facteur XIII < 5% (Tabibian *et al.*, 2018).

- **Thrombophilie héréditaire (Facteur V Leiden, variation 677C>T du gène *MTHFR* et variation 20210G>A de la prothrombine)**

Les polymorphismes les plus courants responsables d'une thrombophilie comprennent le facteur V de Leiden (remplacement d'une arginine par une glutamine en position 506), les polymorphismes du gène codant pour la méthylène-tétrahydrofolate réductase (MTHFR) et la variation 20210G>A du gène de la prothrombine (FII). Le rôle d'une thrombophilie héréditaire dans la survenue d'une HIC fœtale est controversé (Curtis *et al.*, 2017).

Dans la littérature, on trouve plusieurs cas d'HIC chez des fœtus porteurs d'un ou de plusieurs polymorphismes dans ces gènes. Notamment, Verdu *et al.*, (2005) rapportent un fœtus ayant présenté une hémorragie périventriculaire évoluant vers une porencéphalie à 28 SA et qui était hétérozygote pour le facteur V Leiden (R506Q). Ramenghi *et al.*, (2005) rapportent aussi la découverte d'une hémorragie périventriculaire chez un fœtus de 23 SA, hétérozygote pour le FV Leiden et homozygote pour le variant C677T de MTHFR. Enfin, Crespin et al., (2009) rapportent un cas d'HIC chez un fœtus à 22SA. L'examen neuropathologique montrait une hémorragie intraventriculaire et intraparenchymateuse au niveau de la matrice germinale et des lésions ischémiques secondaires (leucomalacie). Les auteurs émirent l'hypothèse de caillots sanguins in situ dans les petits vaisseaux de la zone germinative à l'origine de l'HIC mais aucun caillot fut retrouvé à l'examen neuropathologique. Ce fœtus était porteur hétérozygote du FV Leiden et de la variation 20210G>A de la prothrombine.

Dans tous ces cas mentionnés, le lien de causalité entre HIC fœtale et thrombophilie héréditaire n'a jamais été confirmé et reste difficile à démontrer puisqu'il s'agit de polymorphismes très fréquents dans la population générale. La thrombophilie à facteur V Leiden est la forme la plus courante de thrombophilie avec une prévalence variant selon les populations. La variation hétérozygote R506Q responsable du facteur V Leiden est présente chez 3 à 8 % individus dans les populations américaines et européennes (Kujovich, GeneReviews 2018). La variation C677T de la MTHFR est un polymorphisme commun dans la population générale avec une fréquence à l'état homozygote variant entre 5 et 15 % (Berrut *et al.*, 2003). Enfin la prévalence de l'allèle A pour le gène de la Prothrombine est de l'ordre 1 à 4 % dans la population générale (Trillot *et al.*, 2000).

- **Déficit en protéine C**

Le déficit en protéine C est connu à l'état hétérozygote comme étant responsable de thromboses macro- et microvasculaires. Il s'agit d'une maladie génétique rare causée par des altérations du

gène *PROC*. La prévalence du déficit hétérozygote en protéine C est de 0,2% tandis que le déficit homozygote en protéine C est très rare, avec une prévalence de 1 pour 500 000 à 750 000 individus (Song *et al.*, 2021). Les patients porteurs d'une variation pathogène hétérozygote souffrent de thromboses veineuses récurrentes d'apparition relativement tardive et d'une diminution des taux plasmatiques et de l'activité de la protéine C (Goldenberg et Manco-Johnson, 2008). Ils sont à risque pour une embolie pulmonaire mais peuvent aussi être asymptomatiques. En revanche, les patients homozygotes ou hétérozygotes composites ont un déficit sévère en protéine C et des symptômes plus graves incluant un purpura fulminans et une coagulation intravasculaire disséminée (CIVD) sévère pendant la période néonatale.

Le déficit sévère en protéine C peut se manifester en période anténatale par une HIC secondaire à la formation de microthrombi dans les capillaires des vaisseaux (Stutterd *et al.*, 2014) comme illustré dans le cas suivant d'un fœtus interrompu à 35 SA du fait d'une HIC sévère (Figure 7).

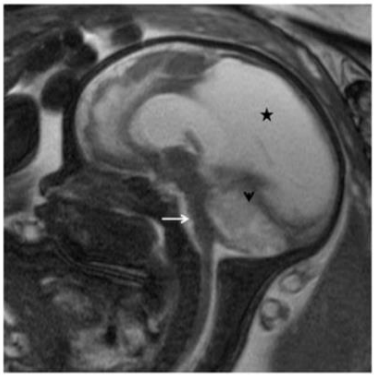


Figure 7. IRM fœtale d'un fœtus de 35SA avec déficit sévère en protéine C. Séquence sagittale T2 montrant un amincissement du tronc cérébral avec aplatissement du pons (flèche), une importante porencéphalie (étoile) et un œdème du cervelet (tête de flèche). Extrait de Stutterd *et al.*, 2014.

Bien que ces présentations fœtales/néonatales avec hémorragies intraparenchymateuses bilatérales liées à des variations pathogènes bialléliques de *PROC* soient connues, elles sont rarement recherchées en raison de leur très faible incidence (Fong *et al.*, 2010 ; Martin *et al.*, 2021).

3-b) Gènes codant pour une protéine de la membrane basale, *COL4A1* et *COL4A2*

- **Généralités**

Le collagène de type IV appartient à la superfamille des collagènes dont la structure des protéines est caractérisée par la présence de domaines composés d'une triple hélice alpha. A la différence des autres types de collagène tels que les type I et III qui forment des fibres (collagènes fibrillaires), le collagène de type IV a la capacité de s'autoassembler en réseaux organisés (Kaur et Reinhardt, 2015). Cette capacité est essentielle pour permettre l'assemblage de la membrane basale, structure hautement réticulée, et assurer sa stabilité. La membrane basale joue un rôle structural (ancrage des cellules dans le tissu conjonctif) mais elle intervient également dans de nombreux processus physiologiques (Pöschl *et al.*, 2004).

Les gènes *COL4A1* et *COL4A2* sont localisés sur le chromosome 13 en 13q34 et positionnés de manière « tète bèche » partageant ainsi le même promoteur (Figure 8). Le gène *COL4A1* est composé de 52 exons et *COL4A2* de 48 exons. Ces deux gènes codent pour les sous unité alpha 1 et alpha 2 du collagène de type IV, constituant essentiel des membranes basales de nombreux tissus dont l'endothélium vasculaire, l'épithélium de la cornée, la conjonctive de l'œil, les glomérules et les tubules rénaux.

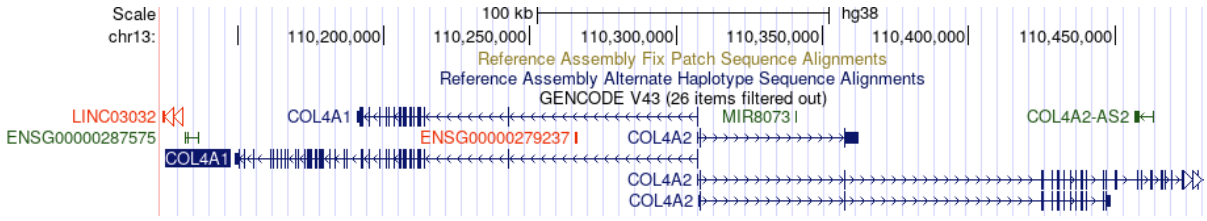


Figure 8. Localisation des gènes *COL4A1* et *COL4A2* sur le chromosome 13.
Source https://genome.ucsc.edu.

- **Structure des protéines COL4A1 et COL4A2**

Il existe 6 isoformes de collagène de type IV (α1 à α6) qui partagent une importante homologie de séquence et d'organisation structurale. Celles-ci permettent de produire trois types d'hétérotrimères α1α1α2, α3α4α5 et α5α5α6. L'hétérotrimère α1α1α2 est ubiquitaire dans l'organisme et apparaît précocement au cours du développement embryonnaire à la différence des autres hétérotrimères (Sundaramoorthy *et al.*, 2002).

Chaque isoforme est formée de trois domaines : un domaine 7S (*cross link* de 4 triples hélices), un large domaine collagénique formé de répétitions de triplets Gly-X-Y et comprenant plusieurs interruptions, et un domaine C-terminal (domaine NC-1) permettant l'assemblage moléculaire (Figure 9). L'hétérotrimère α1α1α2 comprend 2 chaines codées par le gène *COL4A1* et une chaine codée par *COL4A2,* chaines qui forment une triple hélice (Mao *et al.*, 2005).

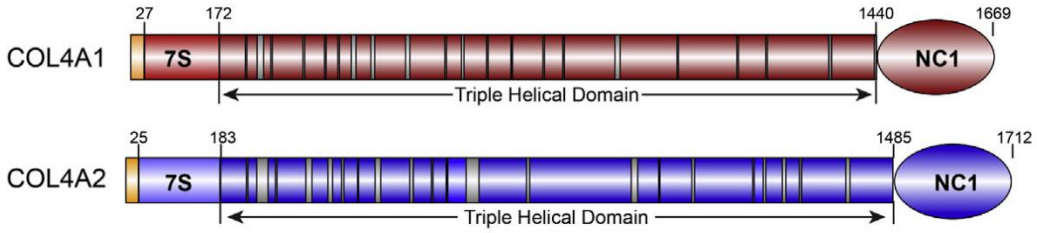


Figure 9. Structure des protéines COL4A1 et COL4A2. Les nombres indiquent les positions en acides aminés. Les cases grises dans le domaine de la triple hélice indiquent les interruptions des répétitions. Extrait de Mao *et al.*, 2015.

- **Manifestations cliniques des variations pathogènes de *COL4A1/COL4A2***

L'implication de *COL4A1* dans les hémorragies intracrâniennes a été décrite pour la première fois chez la souris en 2005 (Gould *et al.*, 2005). Dans ce modèle de souris hétérozygotes pour une délétion de l'exon 40 de Col4a1, une mortalité de 50 % était observée dans les premiers jours de vie à la suite d'une HIC et 18% des souriceaux survivants présentaient une porencéphalie. Ces souris Col4a1 +/DelEx40 présentaient au niveau histologique des anomalies de la membre basale des vaisseaux cérébraux. Pour déterminer si des variations de *COL4A1* chez l'homme pouvaient être responsable d'une porencéphalie, Gould *et al.*, investiguèrent deux familles présentant une porencéphalie autosomique dominante et identifièrent pour chacune d'entre elles une variation hétérozygote faux-sens affectant une Glycine dans le domaine de la triple hélice. Les années suivantes, l'analyse phénotypique de nombreux patients présentant une variation pathogène affectant une glycine de la triple hélice de *COL4A1* mit en évidence une grande variabilité dans l'expression phénotypique de cette affection que ce soit dans l'âge de début ou les symptômes (Breedveld *et al.*, 2006 ; Van der Knaap *et al.*, 2006 ; Gould *et al.*, 2006 ; Plaisier *et al.*, 2009, Meuwissen *et al.*, 2015, Zagaglia *et al.*, 2018). Les protéines COL4A1 et COL4A2 étant structurellement et fonctionnellement liées, il a également été démontré, chez la souris et chez l'homme, l'implication de variations de *COL4A2* dans la survenue d'une hémorragie intracérébrale (Jeanne *et al.*, 2012).

L'âge d'apparition des signes cliniques chez les individus porteurs de variants pathogènes de *COLA41/COL4A2* est très variable de la vie in utero à l'âge adulte (De Vries *et al.,* 2009 ; Meuwissen *et al.,* 2011 ; Lichtenbelt *et al,* 2012 ; Garel *et al.,* 2013 ; Zagaglia *et al.,* 2018, Guey et Hervé, 2022).

Chez l'enfant, les variations pathogènes de *COL4A1/COL4A2* peuvent être responsables d'une hémorragie cérébrale intra-parenchymateuse qui se manifeste par une hémiplégie infantile, diagnostiquée pendant les premiers mois de vie devant un retard de développement moteur. A l'imagerie cérébrale on peut retrouver (Yoneda *et al.,* 2012 ; Meuwissen *et al.,* 2015 ; Cavallin *et al.,* 2018 ; Itai *et al.,* 2021) :

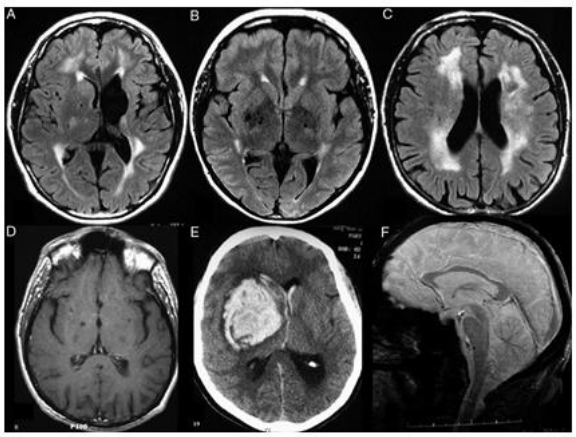
- Une porencéphalie
- Une dilatation ventriculaire avec hydrocéphalie
- Une schizencéphalie
- Des calcifications intracrâniennes
- Une polymicrogyrie
- Des hétérotopies nodulaires périventriculaires
- Une hypoplasie voire agénésie du corps calleux
- Des anomalies de la fosse postérieure de type hypoplasie cérébelleuse
- Une atteinte de la substance blanche périventriculaire (leucoencéphalopathie)

La porencéphalie est le signe le plus fréquemment observé chez l'enfant mais il existe potentiellement un biais de recrutement puisqu'il s'agit du principal signe d'appel conduisant à un criblage de *COL4A1/COL4A2.* Sur le plan anatomique, la porencéphalie est définie comme une lésion kystique ou une cavité remplie de liquide, présente dans le cerveau et communiquant avec le ventricule latéral. La porencéphalie serait secondaire à des lésions de dégénérescence focale de la substance blanche, notamment les tissus périventriculaires, à la suite d'un infarctus hémorragique par congestion et/ou occlusion veineuse.

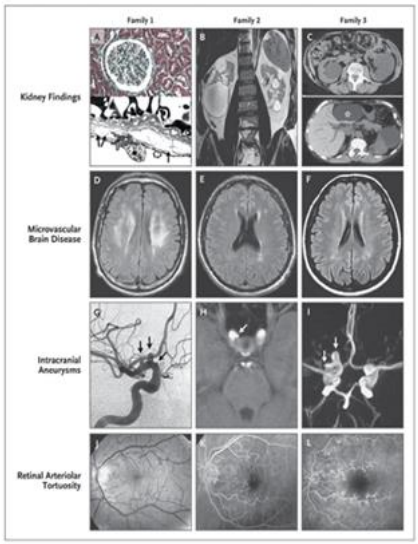
Chez le sujet adulte, les variations pathogènes de *COL4A1/COL4A2* sont responsables d'accidents vasculaires cérébraux (hémorragie intracérébrale, plus rarement accident ischémique transitoire ou infarctus lacunaire). La plupart des lésions vasculaires sont localisées dans les territoires profonds (noyaux gris centraux, substance blanche) et associées à une leucoencéphalopathie périventriculaire d'étendue variable et parfois à des microhémorragies

et/ou des espaces périvasculaires (Virchow Robin) dilatés témoignant d'une microangiopathie cérébrale. Certains patients peuvent avoir une leucoencéphalopathie asymptomatique avec des microhémorragies et/ou des espaces périvasculaires dilatés. Les atteintes neurovasculaires affectent le plus souvent les petits vaisseaux cérébraux mais sont aussi responsables d'anévrysmes intracrâniens, en particulier d'anévrysmes du siphon carotidien (Alamowitch *et al.*, 2009). Les lésions anévrysmales peuvent plus rarement affecter des artères périphériques, spléniques ou mésentériques.

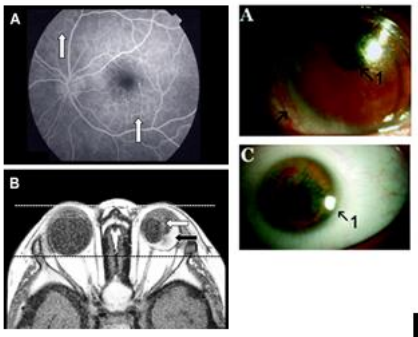
Outre les manifestations hémorragiques cérébrales qui font toute la gravité de cette affection, les atteintes cliniques peuvent être ophtalmologiques, rénales, et / ou musculaires. Les tortuosités artériolaires rétiniennes constituent un des principaux phénotypes oculaires de la maladie mais elles sont inconstantes. Il s'agit d'anomalies d'allongement des branches de l'artère centrale de la rétine limitées au réseau artériolaire. Elles peuvent se compliquer d'hémorragies rétiniennes responsables de pertes visuelles transitoires mais elles peuvent être totalement asymptomatiques et découvertes uniquement à l'examen du fond d'œil. Les autres anomalies oculaires décrites sont des dysgénésies congénitales du segment antérieur de l'œil et en particulier de l'iris et de la cornée (malformation d'Axenfeld-Rieger), une cataracte congénitale ou post-traumatique, une excavation du nerf optique et une hypertension intraoculaire. L'atteinte rénale se manifeste de façon variable par une hématurie macroscopique ou microscopique, une insuffisance rénale et des kystes rénaux. L'atteinte rénale est inconstante mais peut être sévère. Les autres manifestations cliniques possibles sont des anomalies cardiaques (troubles du rythme cardiaque supraventriculaire, phénomène de Raynaud, prolapsus de la valve mitrale), une anémie hémolytique et des kystes hépatiques (Meuwissen *et al.*, 2015). Des crampes musculaires avec élévation des CPK et des migraines avec aura ont été décrites dans certaines familles. Les principales manifestations clinico-radiologiques chez l'adulte sont illustrées dans la figure suivante (Figure 10).



Hématomes cérébraux

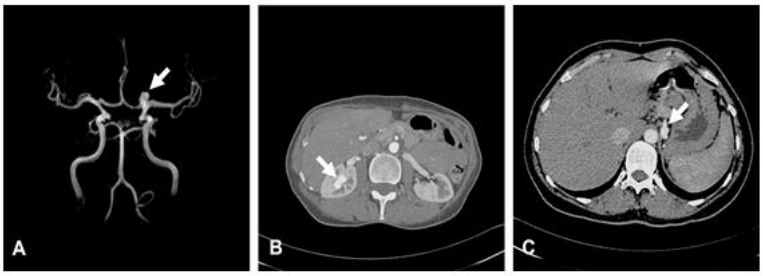


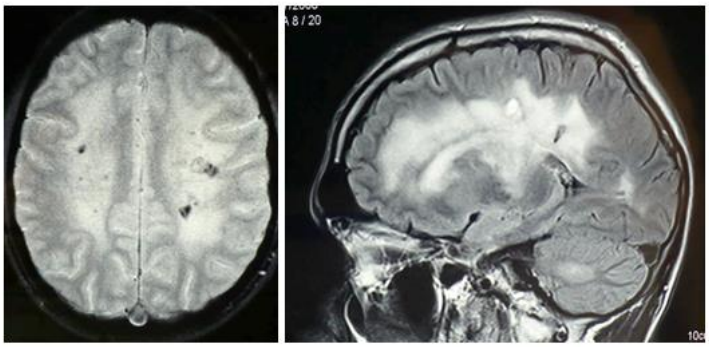
Hématurie
Kystes



Tortuosités artérielles rétiniennes
Anomalies du segment antérieur de l'œil
Cataracte, Microphtalmie

Anévrysmes intra et extra-crâniens





Leucoencéphalopathie et microbleeds

Figure 10. Manifestations cliniques chez l'adulte associées aux variations pathogènes de *COL4A1/COL4A2*. Images extraites de : Plaisier *et al.*, 2007 ; Sibon *et al.*, 2007 ; Vahedi *et al.*, 2007 ; Alamowitch *et al.*, 2009 ; Rouaud *et al.*, 2010 ; et communication du Dr F. Buffon.

Il existe une variabilité phénotypique intrafamiliale ainsi qu'une pénétrance incomplète. Dans une même famille il est possible d'observer des enfants atteints d'une porencéphalie avec des apparentés adultes asymptomatiques bien que porteurs de la variation pathogène, présentant ou non une leucoencéphalopathie vasculaires, comme illustré dans la figure suivante (Figure 11).

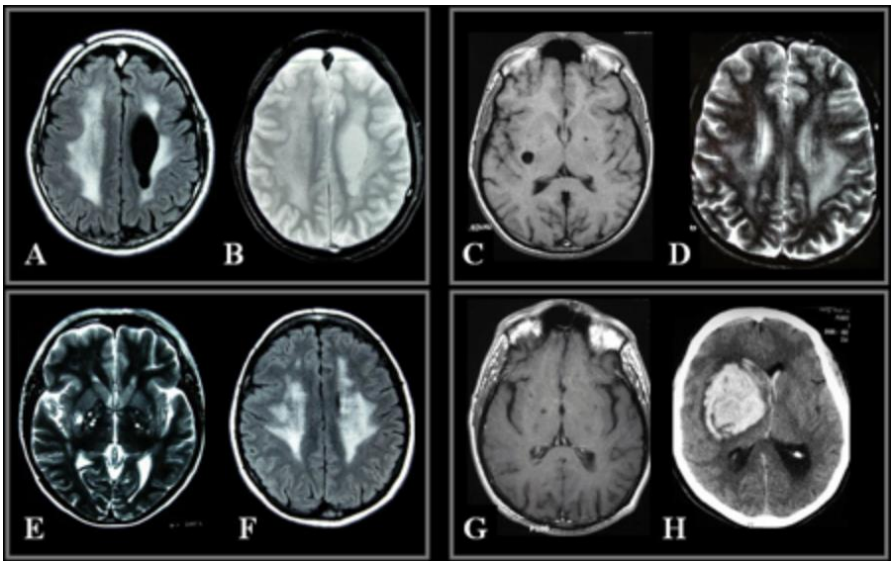


Figure 11. Exemple d'imageries cérébrales d'une fratrie de quatre sujets porteurs de la même variation pathogène de *COL4A1*. Images extraites de Vahedi *et al.*, 2007 et du site internet du *Centre de Référence pour les maladies rares des Vaisseaux du Cerveau et de l'œil (CERVCO).* En A et B, patient de 35 ans ayant comme antécédent une hémiplégie infantile avec une porencéphalie et une leucoencéphalopathie périventriculaire diffuse en FLAIR (A) et en écho de gradient (B). En C et D, patient de 33 ans asymptomatique présentant une dilatation des espaces périvasculaires en T1 (C) et une leucoencéphalopathie périventriculaire à prédominance postérieure en T2 (D). En E et F : patiente de 47 ans avec comme seul antécédent une migraine avec aura. Dilatation des espaces périvasculaires en T2 (E) et leucoencéphalopathie périventriculaire en FLAIR (F). En G et H : présence de quelques discrètes dilatations des espaces périvasculaires en T1 lors d'un examen systématique à 36 ans (G), décès du patient à 40 ans d'un volumineux hématome profond sous traitement par anticoagulant oral pour une thrombose veineuse (H).

Cette variabilité phénotypique peut également concerner des familles avec combinaison de formes adultes et foetales (Figure 12).

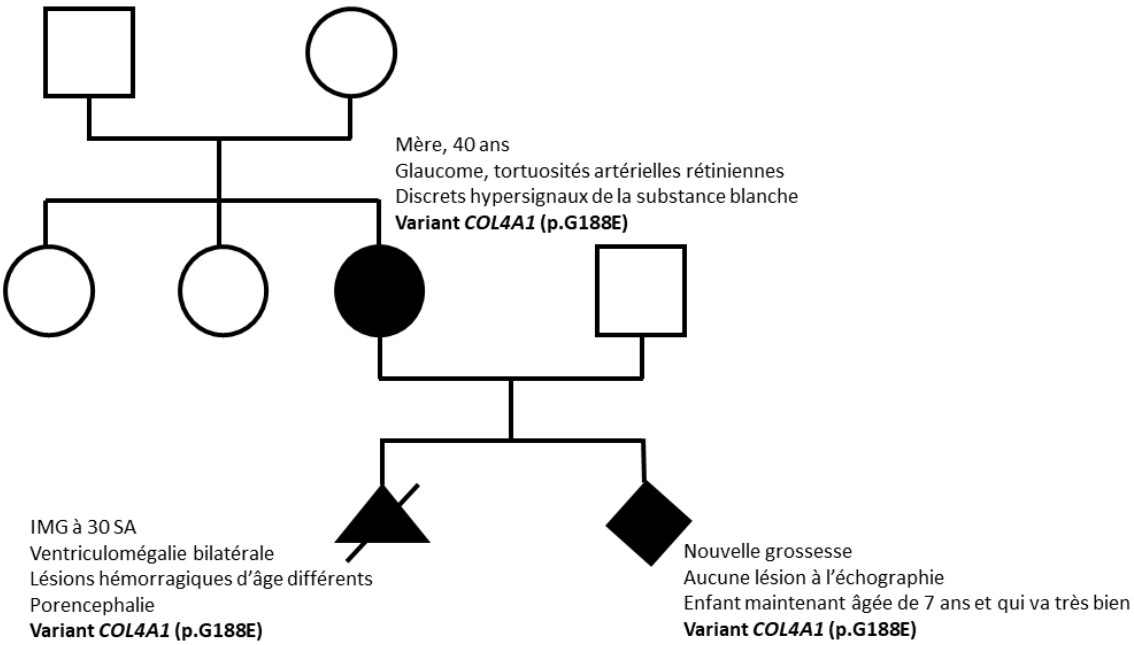


Figure 12. Illustration de la variabilité intrafamiliale associée aux variations pathogènes de *COL4A1/COL4A2 (Famille suivie au CERVCO).*

- **Spectre phénotypique chez le fœtus**

Chez le fœtus, l'atteinte est essentiellement cérébrale mais celle-ci peut s'associer à des anomalies du développement en particulier oculaire (microphtalmie, cataracte). Les facteurs contribuant de façon importante à cette atteinte vasculaire cérébrale sont le caractère très vascularisé et la fragilité de la matrice germinale. La plupart des cas de porencéphalie congénitale sont la conséquence d'une hémorragie de la matrice germinale en anténatal suivie d'un infarctus veineux et d'une destruction tissulaire.

Bien que l'apparition anténatale de lésions intracrâniennes associées au gène *COL4A1* ait été signalée dès les premières descriptions de cas, le diagnostic avait toujours été posé après la naissance. Ce n'est qu'en 2012, que le premier cas de fœtus porteur d'un variant pathogène de *COL4A1* a été décrit (Lichtenbelt *et al.*, 2012). Il s'agissait d'une découverte à l'échographie de 21 SA d'une hémorragie de la matrice germinale. L'échographie montra des parois hyperéchogènes au niveaux des ventricules latéraux et périventriculaire gauche, et un hémisphère cérébelleux gauche remplacé par une zone écholucide. L'IRM fœtale confirma les résultats de l'échographie.

Les années suivantes plusieurs cas de lésions cérébrales fœtales liées à des variants pathogènes de *COL4A1/COL4A2* ont été rapportés et ont permis d'affiner le spectre des anomalies

neurologiques observées chez le fœtus. Ces anomalies neurologiques décrites sont des lésions ischémiques-hémorragiques graves et/ou multifocales (Maurice *et al.,* 2021 / **Article en Annexe**) pouvant inclure une ventriculomégalie, une porencéphalie, une schizencéphalie, des calcifications, une agénésie ou une hypoplasie du corps calleux, ainsi que diverses anomalies de la fosse postérieure (Garel *et al.*, 2013 ; Colin *et al.*, 2014 ; Takenouchi *et al.*, 2015 ; Durrani-Kolarik *et al.,* 2017 ; England *et al.,* 2020 ; Maurice *et al.,* 2021 (**Article en annexe**) ; Hausman-Kedem *et al.,* 2021 ; Itai *et al.,* 2021). Il s'agit en fait de lésions similaires à celles décrites chez l'enfant diagnostiqué en période postnatale et qui sont associées chez ce dernier à un trouble neurodéveloppemental, une épilepsie et une hémiplégie infantile.

- **Caractéristiques anatomopathologiques**

Il existe peu de descriptions des lésions anatomopathologiques chez les fœtus porteurs de variants de *COL4A1/COL4A2*.

En 2021, Shannon *et al.,* rapportèrent quatre fœtus avec variants pathogènes de *COL4A1*. Au niveau morphologique, ces fœtus ne présentaient aucune anomalie viscérale en dehors des anomalies cérébrales et +/- oculaires. L'analyse histologique cérébrale montrait la présence de macrophages riches en fer au niveau de la zone ventriculaire, une gliose réactionnelle périventriculaire, une destruction et une vacuolisation des cellules musculaires lisses de la média, ainsi que des anomalies de la membrane basale des cellules musculaires lisses. De nombreux dépôts d'hémosidérine étaient également observés notamment aux niveaux épendymaire et sous-épendymaire. Un marquage immunohistochimique avec un anticorps anti-collagène IV montra une hypervascularisation sous-corticale et profonde et parfois la présence de petits vaisseaux très anormaux.

La même année une équipe Israélienne rapporta les premières données neuropathologiques d'un fœtus porteur d'un variant pathogène de *COL4A2* qui présentait une HIV avec atteinte parenchymateuse (grade IV). L'histologie révéla des hétérotopies neuronales, une polymicrogyrie et des anomalies de marquage anti-collagène type IV (Figure 13).

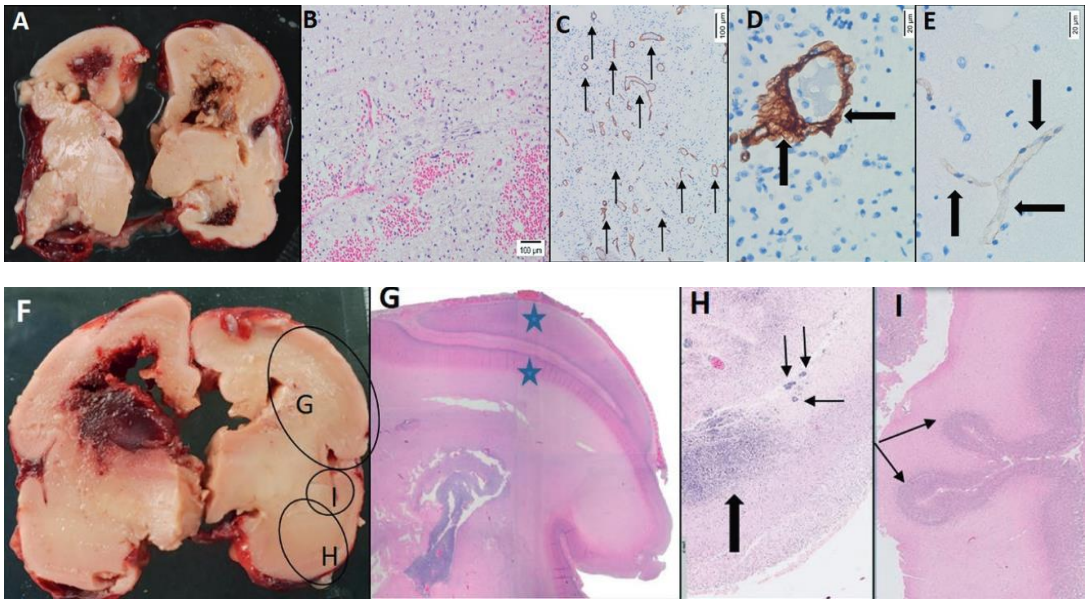


Figure 13. Examen neuropathologique d'un fœtus présentant une délétion intragénique de *COL4A2* à 26 SA. Extrait de Hausman-Kedem *et al.*, 2021. (A) Image macroscopique en coupe coronale des lobes temporo-pariétaux montrant une HIV bilatérale de grade IV. Présence de sang dans les ventricules latéraux, désintégrant la substance blanche périventriculaire. (B) Image histologique (coloration hématoxyline et éosine) montrant une désintégration de la substance blanche avec extravasation d'érythrocytes. (C à E : Immunomarquage du collagène de type IV) (C) Marquage proéminent au niveau de la membrane basale des vaisseaux périventriculaires. (D) Épaississement irrégulier du marquage dans un vaisseau périventriculaire. (E) Marquage fin dans la membrane basale d'un vaisseau de la substance blanche sous-corticale. (F) HIV de grade IV dans le lobe temporo-pariétal droit. Les cercles indiquent l'emplacement des images histologiques G à I. (G) Double cortex dysplasique. (H) Hétérotopies de la substance blanche. (I) Cortex polymicrogyrique.

Enfin, une des plus grandes séries de données fœtopathologiques rapportées est celle de Gubana *et al.*, 2022 (**Article en annexe**) portant sur huit fœtus avec un variant pathogène de *COL4A1*. Les caractéristiques les plus fréquentes à l'échographie initiale étaient une biométrie cérébrale diminuée, une ventriculomégalie modérée à sévère avec des zones hyperéchogènes et des irrégularités de la paroi ventriculaire, des hyperéchogénicités parenchymateuses et des lésions cavitaires. L'analyse neuropathologique mit en évidence une encéphalopathie hémorragique associée à des lésions ischémiques de la substance grise et blanche. En outre, il a été identifié des hémorragies pétéchiales disséminées d'abondance variable et d'âges différents, des calcifications périventriculaires et des macrophages chargés en fer au niveau de la zone sous-épendymaire. Une analyse morphométrique des vaisseaux à l'aide d'un immunomarquage anti-CD34 et anti-SMA suggérerait une augmentation anormale de l'angiogenèse, une dilatation des

vaisseaux, ainsi qu'une atténuation focale voire une perte du marquage des cellules musculaires lisses.

- **Mécanismes physiopathologiques**

Sur le plan moléculaire et cellulaire, deux hypothèses principales, non mutuellement exclusives, ont été émises (Figure 14). Dans la première de ces hypothèses, les hétérotrimères comprenant une/des isoformes mutées ne seraient pas sécrétés conduisant à une haploinsuffisance et à une toxicité cellulaire (atteinte pathogène primaire). Dans la seconde hypothèse la sécrétion dans la membrane basale d'hétérotrimères mutés pourrait directement ou indirectement modifier les interactions avec des molécules de signalisation ou des récepteurs à la surface cellulaire telles que les intégrines (Kuo *et al.*, 2012).

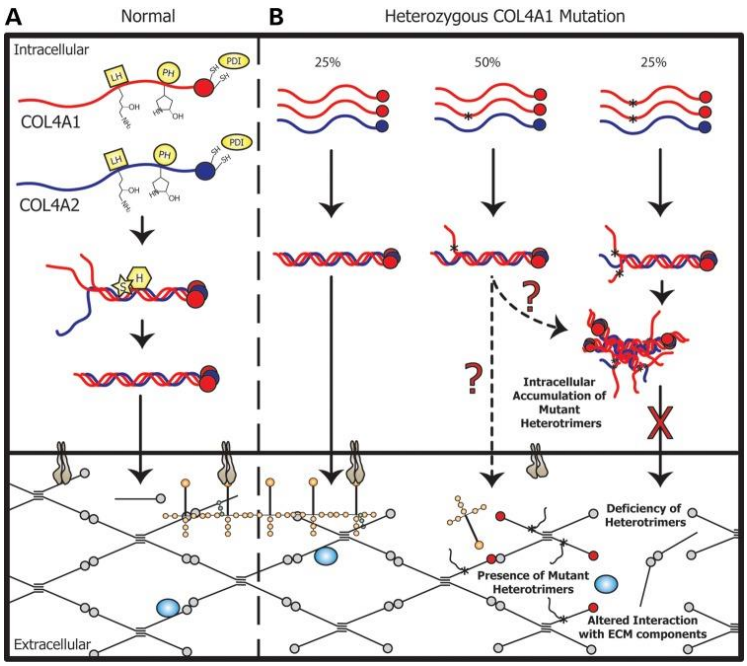


Figure 14. Représentation schématique de la biosynthèse du collagène de type IV en condition normale (A) et en présence d'un variant pathogène de *COL4A1* (B). LH, lysyl hydroxylase ; PH, prolyl hydroxylase ; PDI, protéine disulfure isomérase ; S, protéine sécrétée ; H, protéine de choc thermique 47 (HSP47). Extrait de Kuo *et al.*, 2012.

Sur le plan vasculaire, le mécanisme pathogène présumé est une altération de la membrane basale vasculaire. Gould *et al.*, (2005) ont initialement montré chez la souris que les variations de Col4a1 entraînaient des ruptures focales de la membrane basale vasculaire et un gonflement des cellules endothéliales vasculaires avec des vésicules proéminentes. Ces données ont été

confirmées chez l'humain, par Van der Knapp *et al.,* (2006) par microscopie électronique d'une biopsie de peau. Dans cette étude, les auteurs ont constaté des ruptures focales et une augmentation importante de l'épaisseur de la membrane basale vasculaire des capillaires cutanés (Figure 15).

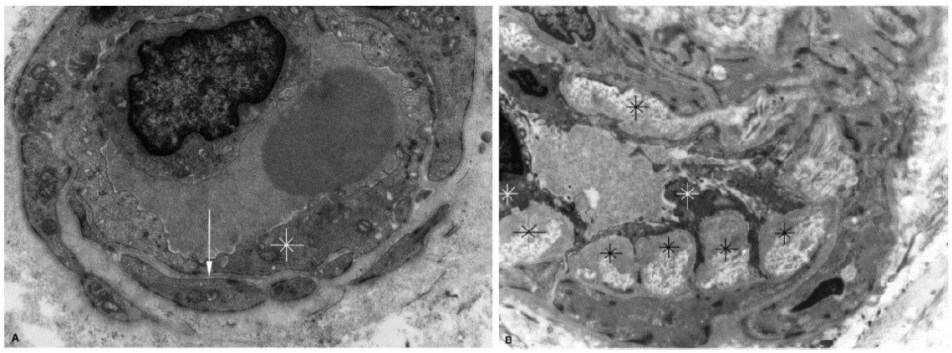


Figure 15. Microscopie électronique de capillaires cutanés chez un patient porteur d'un variant pathogène de *COL4A1*. L'image de gauche illustre la membrane basale d'un capillaire normal (grossissement 12 000), la flèche indique la membrane basale d'une épaisseur de 40 nm, l'astérisque indique une cellule endothéliale. L'image de droite est celle d'un capillaire anormal (grossissement 15 000). La membrane basale (astérisques noirs) a une épaisseur de 2 500 nm et est fragmentée. Extrait de Van der Knapp *et al.,* 2006.

Sur le plan des mécanismes des hémorragies cérébrales, Ratelade *et al.,* (2020) ont montré récemment que les souris mutantes Col4a1 présentent une hypermuscularisation (augmentation du nombre de cellules murales) des segments transitoires, qui sont situés à l'interface entre les artérioles et les capillaires, et une perte focale des cellules musculaires lisses des artérioles dans le cerveau et la rétine. L'hypermuscularisation des segments transitoires augmente la pression intravasculaire dans l'artériole en amont abimée par la perte des cellules musculaires lisses, conduisant finalement à une rupture de l'artériole (Figure 16).

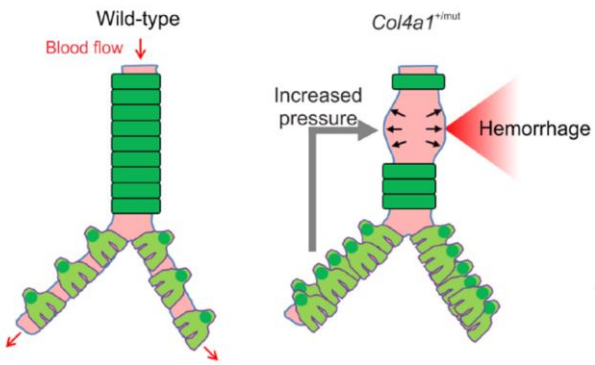


Figure 16. Mécanisme de rupture de la paroi artériolaire chez les souris mutantes Col4a1. Extrait de Ratelade *et al.*, 2020.

- **Criblage moléculaire à visée de diagnostic, conseil génétique et prise en charge des femmes enceintes**

Chez un sujet symptomatique, le criblage des gènes *COL4A1* et *COL4A2* est réalisé par séquençage haut débit des régions exoniques et introniques flanquantes de l'ADN extrait du sang circulant. Les variations pathogènes typiques sont des substitutions nucléotidiques responsables du changement d'un résidu de type Glycine dans le domaine de la triple hélice. Par ailleurs, des variations de type faux sens dans le domaine NC-1 ont été rapportées de même que des variations disruptives responsables d'une dégradation de l'allèle muté et donc d'une haploinsuffisance.

Chez un apparenté majeur asymptomatique, il existe un bénéfice à la réalisation d'un test de diagnostic présymptomatique lié à l'existence de mesures préventives et au conseil génétique. Les sujets souhaitant connaitre leur statut peuvent bénéficier d'un test après consultation de génétique. En cas de présence du variant pathogène familial, un bilan d'extension est alors proposé comprenant une IRM et ARM cérébrale, une échographie rénale, une consultation d'ophtalmologie avec réalisation d'un fond d'œil et un bilan biologique (dosage de la créatininémie, CPK, recherche de protéinurie et d'hématurie). Ce bilan est proposé après une consultation de neurologie visant à expliquer ce qui pourrait être mis en évidence lors de ce bilan et les mesures à mettre en œuvre.

Chez un apparenté asymptomatique mineur un test présymptomatique ne peut être réalisé qu'en cas de bénéfice direct et immédiat pour l'enfant. Ceci fait l'objet de discussions actuellement sans qu'une attitude claire ait été définie pour le moment (Programme de recherche clinique en cours associant généticiens, neuropédiatres et neurologues d'adultes dans la cadre du CERVCO et du Centre National de Référence de l'AVC de l'enfant).

Un diagnostic prénatal peut être proposé et réalisé après discussion du dossier auprès d'un Centre Pluridisciplinaire de Diagnostic Prénatal (CPDPN). Il est également possible d'effectuer un diagnostic préimplantatoire (DPI).

La découverte d'une variation pathogène chez un fœtus soulève une question additionnelle concernant le mode d'accouchement. Il a en effet été montré chez la souris que la césarienne permettait de prévenir les hémorragies périnatales survenant lors du passage de la voie vaginale chez le souriceau muté (Gould *et al.,* 2006). Par analogie, chez l'homme, la réalisation d'une césarienne pour l'accouchement des fœtus porteurs doit être discutée, tout en sachant qu'il n'existe pas de données à ce jour établissant le bénéfice de cette attitude. La réalisation d'une césarienne permettant l'évitement des efforts de poussée pour la mère fait également l'objet de discussions (Buffon, 2019).

Quel que soit le contexte, il est toutefois important de noter que l'expressivité variable et la pénétrance incomplète de cette affection sont des facteurs qui compliquent le conseil génétique et la prise en charge. De futures études sur le rôle de facteurs modificateurs sont nécessaires pour mieux comprendre cette variabilité.

3-c) Gène *COLGALT1* impliqué dans la biosynthèse de COL4A1/COL4A2

Le gène *COLGALT1* code pour la collagène bêtagalactosyltransférase 1 qui est une des deux enzymes qui transfèrent des groupements galactose aux résidus hydroxylysine du collagène. Ce gène est exprimé de façon constitutive et code pour une protéine soluble localisée dans le réticulum endoplasmique.

La connaissance de l'implication de *COLGALT1* en pathologie humaine est récente, puisque c'est en 2018 que Miyatake *et al.* rapportèrent deux patients non apparentés atteints d'une maladie des petits vaisseaux cérébraux de transmission autosomique récessive. Le premier patient était un garçon de 12 ans présentant un retard de développement sévère, une hypotonie, une quadriplégie spastique et des crises d'épilepsie. Son développement intellectuel était gravement perturbé. Un scanner cérébral réalisé à 5 mois de vie révéla la présence de calcifications et d'une dilatation du ventricule latéral gauche. Une IRM cérébrale réalisée à 21 mois montra une porencéphalie dans l'hémisphère gauche et une leucoencéphalopathie bilatérale. Le deuxième patient présentait un trouble du contact visuel dans la petite enfance,

associé à un léger retard développemental. L'IRM cérébrale à 4 mois a mis en évidence des hypersignaux de la substance blanche. A 8 ans son IRM montra une extension des hypersignaux avec présence de microsaignements au niveau des ganglions de la base. À l'âge de 9 ans, il contracta une infection par le virus de la grippe. Au cours de cette infection, une hémiplégie droite apparut puis une perte de conscience motivant la réalisation d'une IRM cérébrale. Celle-ci révéla de multiples foyers hémorragiques intra parenchymateux conduisant au décès quelques jours plus tard.

Au niveau génétique, des variants bialléliques du gène *COLGALT1* ont été identifiés par séquençage de l'exome entier. Les auteurs montrèrent que l'inactivation de *COLGALT1* dans des cellules humaines transfectées entraînait une diminution significative des quantités intracellulaires et extracellulaires de COL4A1 par rapport aux témoins. De plus, la protéine COL4A1 dans les cellules transfectées mutantes avait une distribution diffuse dans le cytoplasme avec un schéma suggérant une accumulation anormale de collagène dans la lumière du RE plutôt que dans les vésicules sécrétoires.

Le patient 1 était hétérozygote composite avec un variant faux sens p.(L151R) et un variant disruptif p.(E366Rfs*15). Les deux variants étaient absents de la base de données de contrôles ExAC et localisés dans le domaine catalytique de la glycosyltransférase (Figure 17). L'analyse de cellules dérivées du patient montra de faibles taux de protéine COLGALT1 par analyse en western-blot et aucune activité enzymatique détectable. Le patient 2 était également hétérozygote composite mais avec deux variants faux sens p.(A154P) et p.(G377R). Les deux variants étaient localisés dans le domaine catalytique de la glycosyltransférase. Le variant p.(A154P) était présent une fois dans la base de données ExAC (à l'état hétérozygote) et le variant p.(G377R) était absent d'ExAC. Les cellules du patient n'étaient pas disponibles pour l'étude, mais des études sur des cellules transfectées montrèrent une diminution des quantités de COLGALT1 par rapport aux contrôles. Les études de l'activité enzymatique montrèrent une activité nulle pour le variant p.(A154P) et diminuée pour le variant p.(G277R).

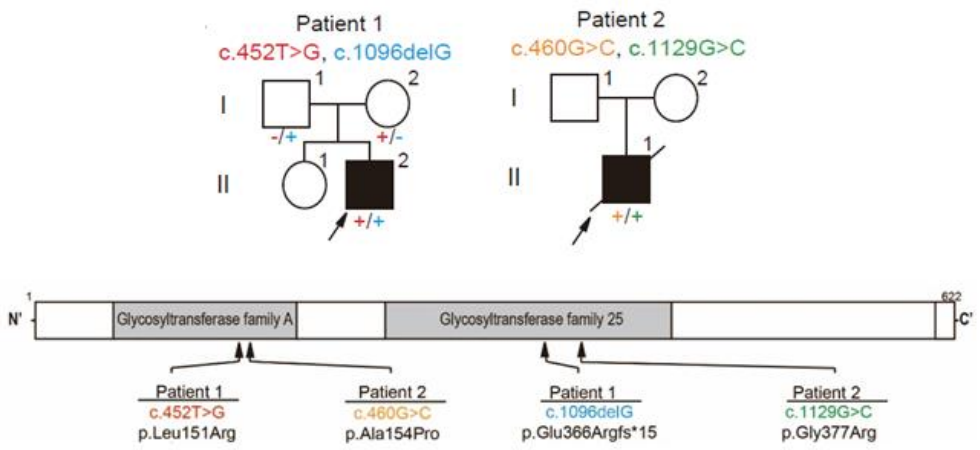


Figure 17. Arbre généalogique avec variants identifiés et représentation schématique de la localisation des variants sur la protéine COLGALT1. Extrait de Miyatake *et al.*, 2018.

Récemment, un nouveau cas a été rapporté avec découverte en anténatal à 20 SA d'une ventriculomégalie (Teunissen *et al.*, 2021). L'IRM fœtale à 21 SA confirma la ventriculomégalie avec hémorragie périventriculaire. Après naissance à terme, l'examen clinique montra une microcéphalie et l'IRM cérébrale montra une cavité porencéphalique et des anomalies de la substance blanche. Une analyse de l'exome entier permit d'identifier un variant homozygote au niveau d'un site canonique d'épissage (c.625-2A>C) du gène *COLGALT1*. Une analyse de l'ARN messager sur fibroblastes cutanés du patient mit en évidence deux transcrits, un transcrit majoritaire aboutissant à un codon stop prématuré et un transcrit minoritaire porteur d'une délétion de 12 paires de bases.

3-d) Gène *JAM3* impliqué dans la jonction des cellules endothéliales

L'endothélium, localisé sur la face interne de tous types de vaisseaux, est constitué d'une monocouche de cellules endothéliales jointives. Les protéines de jonctions intercellulaires endothéliales sont essentielles à l'adhésion intercellulaire, au contrôle de la perméabilité vasculaire aux fluides et aux molécules, et à la transmigration de cellules.

Deux catégories de molécules participent à la formation des jonctions intercellulaires :

- Des molécules transmembranaires : l'occludine, les claudines, les JAM (*junctional adhesion molecules*). Au sein de ces molécules transmembranaires, on peut distinguer les protéines d'adhérence qui assurent l'adhésion intercellulaire grâce à l'action des nectines et des cadhérines, et les jonctions serrées qui régulent le passage des ions et des petites molécules entre les cellules et établissent la polarité cellulaire (Figure 18).

- Des protéines d'attachement intracellulaire qui s'arriment sur les protéines transmembranaires telle que les protéines ZO (zonula occludens).

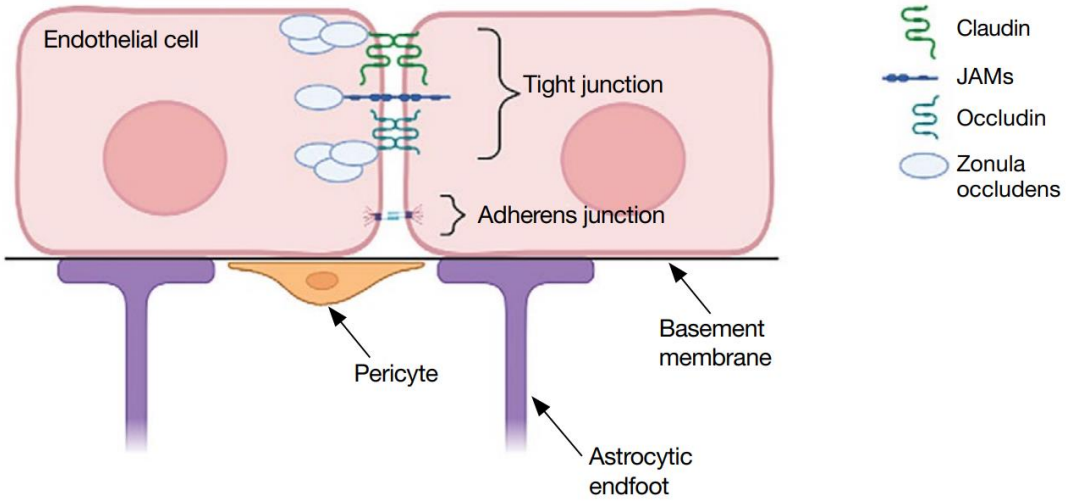


Figure 18. Schéma illustrant les molécules jonctionnelles impliquées dans la cohésion des cellules endothéliales de la barrière hématoencéphalique. Extrait de Schrenk *et al.*, 2022.

Les variations pathogènes de certains gènes codant pour des composants des jonctions comme l'occludine (OCLN) et les molécules d'adhésion jonctionnelle (JAM) ont été associés à des anomalies cérébrales chez l'enfant. On connaît trois types de JAM : JAM1, qui est exprimée à la fois dans les cellules endothéliales et épithéliales, et JAM2 et JAM3, qui sont exprimées dans les cellules endothéliales vasculaires. L'occludine et les JAM jouent un rôle majeur dans la régulation de la fonction de la barrière hémato-encéphalique, y compris la régulation de la perméabilité et de l'extravasation des leucocytes.

Les variations pathogènes de *JAM2* sont associées à une maladie autosomique récessive avec calcifications intracrâniennes appelée « Calcification cérébrale familiale primaire » (Cen *et al.*, 2020) sans manifestions cérébrales hémorragiques.

Les variations pathogènes bialléliques du gène de l'*OCLN* sont responsable d'un syndrome appelé Pseudo-TORCH en raison du phénotype cérébral ressemblant à celui d'une infection intra-utérine congénitale ; celui-ci a pour caractéristiques la présence d'une microcéphalie, d'une polymicrogyrie et de calcifications intracrâniennes (Briggs *et al.*, 2008 ; O'Driscoll *et al.*, 2010). Dans ce syndrome, il n'y a pas précisément d'hémorragie bien que les signes neuroradiologiques soient proches de ceux observés dans une collagénopathie COL4A1/COL4A2 (Thibault *et al.*, 2011 ; Meuwissen *et al.*, 2016).

Enfin, les variations pathogènes de *JAM3* provoquent un syndrome de destruction hémorragique multifocale du parenchyme cérébral avec cavités porencéphaliques, calcifications sous-épendymaires et cataracte congénitale. Ce phénotype est également chevauchant avec celui décrit pour les variations pathogènes de *COL4A1/COL4A2* (Mochida *et al.*, 2010).

Le mécanisme physiopathologique à l'origine de ces syndromes n'est pas connu mais l'hypothèse d'un dysfonctionnement de la barrière hémato-encéphalique est la plus plausible.

3-e) Gène *GATA1*

Le gène *GATA1* code pour un facteur de transcription à doigts de zinc lié à l'ADN qui joue un rôle essentiel dans la différenciation des lignées cellulaires hématopoïétiques. Ce gène est localisé en Xp11.23 et ses variations pathogènes sont responsables d'une thrombocytopénie et d'une anémie avec dyserythropoïèse. En 2018, Bouchghoul *et al.* ont rapporté une famille porteuse d'un variant de *GATA1* responsable d'une HIC fœtale diagnostiquée à 36 semaines de grossesse chez un fœtus de sexe masculin et de deux autres cousins mort-nés présentant une anasarque fœtale et un phénotype d'hémochromatose congénitale à la 37ème et 12ème semaines de grossesse. Le criblage moléculaire révéla la présence d'une variation hémizygote pathogène c.613G>A dans l'exon 4 du gène *GATA1* chez les trois fœtus. Il n'y a pas eu d'autres cas rapportés dans la littérature depuis cette date suggérant qu'il s'agisse d'une cause génétique d'HIC extrêmement rare.

3-f) Gène *USP18*

Le gène *USP18* code pour une protéine intracellulaire avec une double fonction, c'est une protéine inhibitrice de la voie de signalisation à l'interféron de type 1 et une isopeptidase membre de la famille des enzymes de déubiquitination (Honke *et al.*, 2016). Il a été montré que la protéine USP18 est présente dans la microglie (macrophages résidents du système nerveux central) de la substance blanche et qu'elle joue un rôle essentiel pour garder la microglie quiescente. En son absence, se produit une dysrégulation de la voie de signalisation JAK/STAT en aval du récepteur à l'interféron (IFNAR2) provoquant l'expression de nombreuses chémokines, cytokines et autre molécules antivirales ; responsable in fine d'anomalies de la microglie. L'analyse histologique chez la souris déficiente en Usp18 montre la présence pathologique d'amas de microglies dans la substance blanche qui ressemblent fortement à l'état

neuropathologique de plusieurs microgliopathies humaines dont le syndrome Pseudo-TORCH (Goldmann *et al.,* 2015 ; Schwabenland *et al.,* 2019). Par ailleurs, il a aussi été montré la présence d'une hémorragie cérébrale avec hydrocéphalie et nécrose des cellules épendymaires chez la souris Usp18 knock-out (Ritchie *et al.,* 2002).

Chez trois patients consanguins d'origine Turque atteints du syndrome pseudo-TORCH-2, Meuwissen *et al.* (2016) ont identifié une variation tronquante homozygote (p.Q218*) dans le gène *USP18* qui ségrégeait dans la famille. Cette variation a été trouvée par la combinaison d'une analyse de liaison à un séquençage de l'exome entier. Deux frères et sœurs allemands avec le même tableau clinique, précédemment rapportés par Knoblauch *et al.* (2003), se sont révélés être hétérozygotes composites avec la même variation (p.Q218*) à l'état hétérozygote et une large délétion dans la région 3' UTR d'*USP18* sur l'autre allèle. L'analyse des haplotypes de la région contenant la variation p.(Q218*) suggéra un effet fondateur. Les cellules des patients des deux familles montrèrent une absence totale de la protéine USP18 et une réponse aberrante à l'IFN de type I.

Les caractéristiques des individus affectés comprenaient une hémorragie cérébrale étendue, survenant soit avant la naissance, soit peu après la naissance, suivie d'un décès précoce. L'hémorragie était principalement localisée dans la région des ganglions de la base, du tronc cérébral, des pédoncules cérébelleux et des ventricules latéraux. Il y avait en plus des calcifications cérébrales et des anomalies de la substance blanche périventriculaire. Des signes systémiques étaient aussi présents chez les patients, incluant une thrombocytopénie, une insuffisance hépatique avec ascite et des problèmes respiratoires sévères.

L'incidence des variations de *USP18* reste encore inconnue mais la prévalence estimée du syndrome pseudo-TORCH 2 est de 1 / 1 000 000 (Sehrish *et al.,* 2022) dont une quinzaine de cas sont rapportés dans la littérature (Misk *et al.,* 2022).

3-g) Autres anomalies moléculaires

Pour ces anomalies suivantes, il n'y avait qu'un seul cas rapporté et peu d'arguments de causalité.

- **Gène *RANBP2***

Meroni *et al.,* (2023) ont décrit un cas de mort fœtal in utero à 35 semaines de grossesse chez une femme de 42 ans. L'examen fœtopathologique révéla la présence d'un hématome sus-

tentoriel et d'une hémorragie cérébrale intraparenchymateuse. Un WES a été réalisé et a mis en évidence chez le fœtus un variant de novo faux sens (p.S796F) dans le gène *RANBP2*. Les auteurs émirent l'hypothèse que le tableau de ce fœtus pourrait être lié au gène *RANBP2*, bien que ce gène n'ait jamais été associé auparavant aux HIC. Le gène *RANBP2* est connu pour être responsable d'une encéphalopathie aigue nécrosante (ANE), qui est une maladie pédiatrique autosomique dominante rare (âge médian de 5 ans) caractérisée par une encéphalopathie consécutive à une maladie fébrile (Sarigecili *et al.,* 2023). Les auteurs suggèrent que ce variant pourrait être lié à une nouvelle présentation phénotypique de la maladie associée à *RANBP2*. Cependant, ce variant n'a jamais été rapporté auparavant, n'est pas localisé dans un domaine fonctionnel de RANBP2 et aucune analyse fonctionnelle n'a été réalisée pour confirmer la pathogénicité du variant identifié. Davantage de cas similaires doivent être étudiés pour confirmer cette hypothèse.

- **Microdélétion 16p11.2**

Rodriguez *et al.,* (2022) ont rapporté un cas d'hypoplasie du corps calleux découvert à l'échographie de 22 SA, chez une femme de 39 ans, Une nouvelle échographie à 32 SA montrait également une ventriculomégalie asymétrique, des signes d'hémorragie au niveau de la matrice germinale et un kyste périventriculaire. L'IRM fœtale mit en évidence une ventriculomégalie asymétrique sévère. L'examen fœtopathologique fut refusé par les parents mais l'aspect externe du fœtus était sans anomalie. L'analyse sur puce à ADN mit en évidence une délétion de 394kb sur la bras court du chromosome 16, héritée de la mère asymptomatique. Dans la littérature, la microdélétion 16p11.2 est déjà connue comme responsable d'un retard de développement, une déficience intellectuelle et/ou à un trouble du spectre autistique, une obésité et une fréquence accrue de malformations cardiaques. Certains patients peuvent avoir des anomalies à l'imagerie cérébrale comme une dilatation ventriculaire modérée (Shimojima *et al.,* 2009), des malformations de la fosse postérieure et une hétérotopie périventriculaire mais sans signe hémorragique (Chung *et al.,* 2021). Ces patients présentent en outre des signes dysmorphiques (Shinawi *et al.,* 2010). Dans 80% des cas il s'agit d'une délétion apparue *de novo,* et la pénétrance de la maladie n'est pas complète (62,4%) (Rosenfeld *et al.,* 2013).

C- Problématique et objectifs de travail

La littérature scientifique et médicale est pauvre en informations sur les déterminants génétiques à l'origine d'une HIC chez le fœtus, même si plusieurs gènes ont été identifiés (cf. chapitre d'introduction). Une revue de la littérature réalisée par Cavalière et al en 2021, portant sur toutes les causes identifiées d'HIC anténatale sur la période 1980-2020, souligne le peu d'informations disponibles sur les causes génétiques dans ce domaine, avec en outre très peu de cas décrits pour chaque étiologie génétique.

La même année, une autre équipe s'est intéressée à l'identification de causes génétiques chez 26 cas (25 familles) d'hémorragies périnatales dont neuf cas diagnostiqués en anténatal (Hausman-Kedem *et al.,* 2021). Un séquençage exome entier avait été réalisé et un variant « causal » ou « probablement causal » identifié uniquement chez 3 cas index. Il s'agissait de variants dans *COL4A1, COL4A2* et *TREX1,* ce dernier étant peu convainquant ; les variations pathogènes de *TREX1* sont en effet associées à une vasculopathie rétinienne avec leucoencéphalopathie et manifestations systémiques (RVCL-S) chez des patients porteurs de variants disruptifs dans la partie C-terminale du gène (Richards *et al.,* 2007). Cette étude de Hausman-Kedem *et al.* correspond à la plus grande cohorte rapportée à ce jour avec analyse d'exome entier chez des patients présentant une HIC anténatale.

En résumé, bien que certaines causes génétiques aient été identifiées comme étant responsables d'une HIC chez le fœtus, chez la grande majorité des individus testés aucune de ces causes n'est mise en évidence. L'absence d'identification d'une cause génétique est préjudiciable pour les parents puisqu'elle ne permet pas d'anticiper une récurrence de l'évènement hémorragique lors d'une future grossesse ni de pratiquer un diagnostic prénatal ou préimplantatoire, ni de rassurer le couple sur le risque de récidive en cas de variation pathogène causale survenue *de novo*. Par ailleurs, l'identification d'une variation pathogène a aussi un impact pour les apparentés puisqu'ils pourraient bénéficier des mêmes conseils génétiques. Outre le conseil, l'identification d'une cause génétique sous-jacente peut aussi permettre d'établir des corrélations phénotype-génotype chez les enfants mutés et ainsi anticiper le pronostic et établir des recommandations sur la conduite thérapeutique. Enfin cette identification participe à une meilleure compréhension des mécanismes physiopathologiques des maladies cérébrovasculaires.

Le service de génétique Neurovasculaire de l'hôpital Saint Louis réalise le criblage moléculaire des gènes *COL4A1* et *COL4A2* pour l'ensemble des centres hospitaliers Français et dispose de

ce fait d'une cohorte importante d'échantillons d'ADN de patients et de fœtus adressés pour criblage de ces deux gènes dont beaucoup se sont révélés négatifs.

Notre objectif essentiel dans le cadre de cette thèse était d'identifier les autres gènes impliqués dans cette affection. Nous avons choisi pour cela une stratégie de WES combinant une analyse de listes de gènes impliqués dans des voies candidates et une analyse de variants candidats sans à priori. Cette stratégie a été appliquée à une cohorte de 113 fœtus dont le criblage des gènes *COL4A1/COL4A2* était négatif.

Matériel et Méthodes

A- Patients, familles et contrôles

1) Cohorte de fœtus utilisée pour l'identification des gènes candidats

Nous avons constitué une cohorte de travail de manière rétrospective incluant les fœtus adressés entre 2010 et 2021 au service de génétique Neurovasculaire de l'hôpital Saint-Louis (Paris, France). Sur les 216 fœtus adressés pour un criblage moléculaire des gènes *COL4A1* et *COL4A2* (séquençage ciblé des exons et des régions introniques flanquantes) devant un tableau d'HIC, 39 étaient porteurs d'une variation pathogène dans *COL4A1* ou *COL4A2* et 178 avaient un criblage négatif (Figure 19).

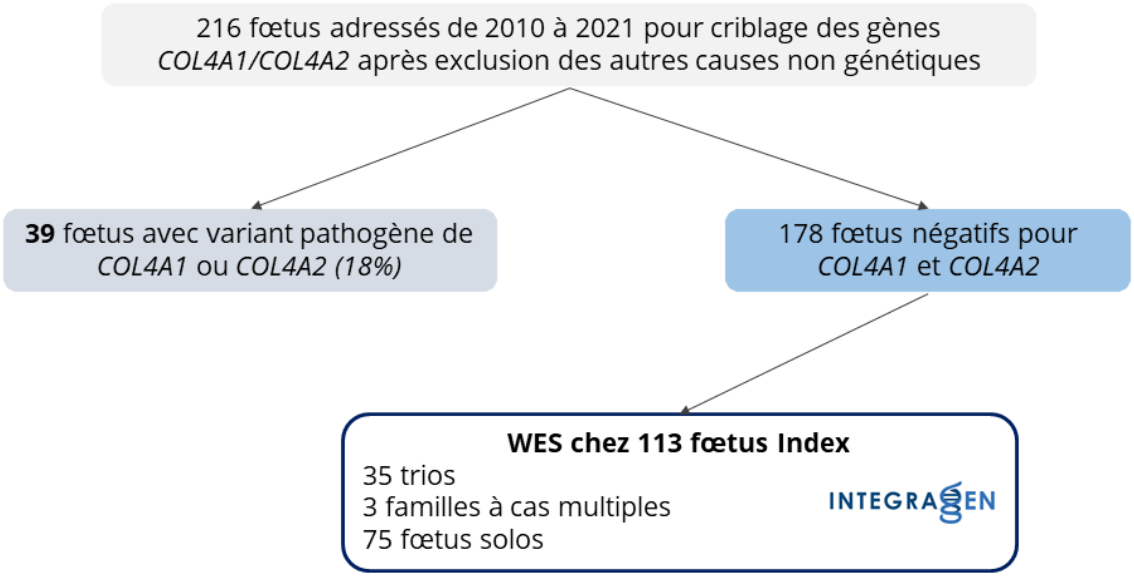


Figure 19. Cohorte fœtale utilisée pour cette étude. WES : *whole exome sequencing.*

Après relecture des informations transmises, obtention des consentements pour criblage pangénomique et vérification de la qualité de l'ADN pour les 178 fœtus négatifs, nous avons pu réaliser un WES chez 113 index comprenant 35 trios, 3 familles avec 2 fœtus atteints et 75 fœtus solos. Dans la majorité des cas il s'agissait de cas sporadiques d'HIC.

Les critères d'inclusion et de non-inclusion des fœtus pour un séquençage de l'exome entier ont été les suivants :

- Critères d'inclusion

- Fœtus adressés pour une HIC, une ventriculomégalie, une porencéphalie et/ou une schizencéphalie lors de l'imagerie prénatale et/ou de l'examen fœtopathologique
- Criblage négatif des gènes *COL4A1* et *COL4A2*
- Consentement écrit des parents pour une analyse de l'exome entier chez leur fœtus

- Critères d'exclusion

- Existence d'une allo-immunisation plaquettaire (présence d'anticorps antiplaquettaires chez la mère et positivité du cross match entre les plaquettes du père et le sérum de la mère)
- Cause acquise fortement suspectée (infection, traumatisme abdominal, usage de drogues durant la grossesse)
- Grossesse gémellaire monochoriale
- Anomalie cytogénétique au caryotype ou en puce à ADN
- ADN de mauvaise qualité

Les parents de tous ces fœtus ont été considérés, sur la base des comptes rendus détaillés des médecins prescripteurs, comme cliniquement asymptomatiques. Le sex-ratio (58 fœtus mâles / 55 fœtus femelles ; ratio = 1,05) n'a pas montré de biais statistique.

Dans la majorité des cas, les lésions cérébrales des fœtus ont été détectées lors du suivi échographique du troisième trimestre. L'âge moyen au moment du diagnostic d'HIC fœtale était de 28,3 +/- 5,2 SA. Une atteinte du parenchyme cérébral était observée dans 90% des cas et une ventriculomégalie était présente chez 84% des fœtus. En raison du mauvais pronostic de ces anomalies cérébrales, 90 % des grossesses ont abouti à une interruption médicale de grossesse et 10 % ont abouti à une mort fœtale in utéro.

2) Cohortes de contrôles

Nous avons utilisé plusieurs bases de données de contrôles publiquement disponibles sur internet. Ces bases de données ont servi pour le filtrage des variants candidats, la priorisation des gènes et l'analyse d'enrichissement en variants rares.

La base de données gnomAD (*Genome Aggregation Database*) est une base de données regroupant un grand nombre de variations génétiques provenant d'individus sains et de patients issus de projets de séquençage à grande échelle et propres à certaines pathologies (psychiatriques, cardiovasculaires, etc…). Les données issues de cohortes recrutées pour des maladies pédiatriques sévères, ainsi que leurs apparentés au premier et au second degré, ont été supprimées de gnomAD de sorte que cette base de données puisse servir de référence pour l'analyse des fréquences alléliques dans les études portant sur les maladies pédiatriques graves.

Trois versions différentes de gnomAD ont été utilisées pour nos analyses :

- Version 2.1 regroupant les données de 125 748 exomes et 15 708 génomes. Cette version comprend plus de 245 millions de variants nucléotidiques (SNVs) et 34,2 millions d'indels (insertions délétions) issus d'individus de différentes origines ethniques, en majorité d'origine Européenne (Figure 20). Les individus de cette version ont été utilisés pour filtrer les variants identifiés sur la base de seuils de fréquence allélique (MAF).

- Version 3.0 regroupant les données de 71 702 génomes. Elle comprend 526 millions de SNVs et 69 millions d'indels. Les individus de cette version 3.0 ont été utilisés comme contrôles pour les tests de charge ou « *burden test* ».

- Version 3.1 regroupant les données de 76 156 génomes. Elle comprend plus de 644 millions de SNVs. Les individus de cette version 3.1 ont été utilisés pour la priorisation des gènes sous l'hypothèse de transmission récessive.

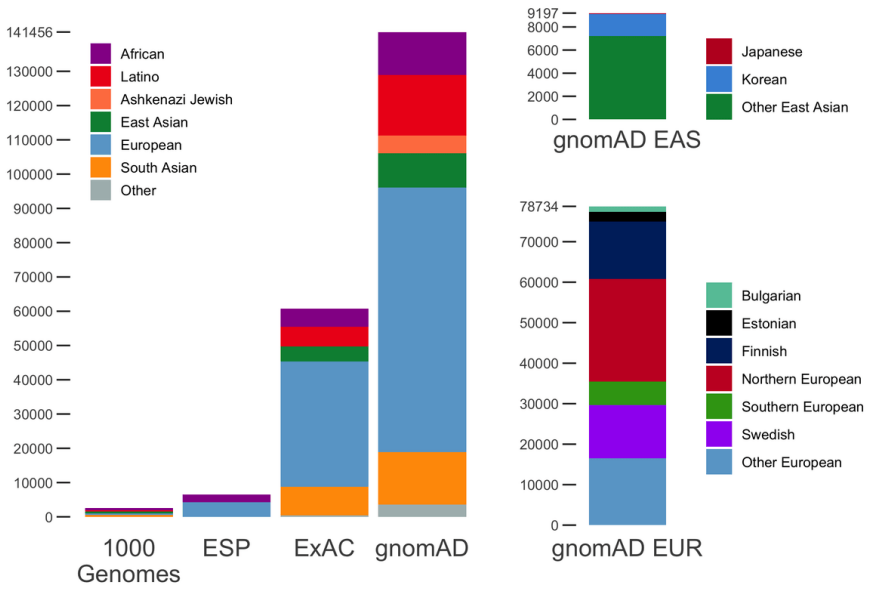


Figure 20. Origine géographique des échantillons dans la version 2.1 de gnomAD. Extrait du site de gnomAD https://gnomad.broadinstitute.org/news/2018-10-gnomad-v2-1/)

Enfin nous avons également utilisé la cohorte 1000 Genomes Project phase 3 composée de 2504 génomes d'individus sains de diverses origines ethniques. Cette cohorte a été utilisé pour le tri des variants candidats et pour l'analyse de l'origine géographique des fœtus.

B- Méthodes

1) Séquençage de l'exome entier (WES)

Le WES des 113 fœtus index et de leurs apparentés a été réalisé par la société Integragen (Evry, France) à partir d'ADN génomique extrait d'amniocytes, de tissus fœtaux, de sang de cordon ou de sang circulant pour les apparentés, selon les étapes ci-après :

- Préparation des librairies par capture avec le kit SureSelect Human All Exon V5 + UTRs (101 Index + 56 apparentés) et le kit Twist Human Core Exome Enrichment System (Twist Bioscience) + IntegraGen Custom V2 (12 Index + 24 apparentés).

- Séquençage *paired-end* (100 paires de bases) sur un séquenceur Illumina NovaSeq.

Après génération des données brutes l'analyse a été effectuée avec un pipeline bio-informatique développé par Integragen comprenant :

- Appel de bases (*base calling*) avec le logiciel Real-Time Analysis et génération des FastQ avec bcl2fastq

- Contrôle de qualité des FastQ avec FastQC

- Alignement sur le génome de référence (hg38) avec le logiciel BWA

- Marquage des « *reads* » dupliqués avec PICARD

- Appel de variants SNVs (Single Nucleotide Variations) et Indels de petite taille (jusqu'à 20bp) avec GATK Haplotype Caller

- Attribution de scores de qualité aux SNVs et Indels

- Annotation des variants avec l'outil *Ensembl Variant Effect Predictor* (VEP) 95.1 pour obtenir la localisation, les conséquences fonctionnelles, la fréquence allélique dans les bases de données de contrôles et les données des logiciels de prédiction in silico.

Le détail des versions des logiciels et des bases de données utilisés est présenté en annexe (Annexe 1).

Les filtres et la stratégie que j'ai utilisée pour l'analyse de ces données est exposée plus bas.

2) Analyse de l'origine géographique et de la consanguinité sur données de WES

Nous avons utilisé le package Peddy (Pederson et Quinlan, 2017) qui est un outil d'analyse avec plusieurs fonctionnalités. Il permet d'étudier la qualité des échantillons et des données de WES en fournissant une série de métriques et de graphiques, y compris des statistiques sur la couverture, les génotypes manquants, les pourcentages d'hétérozygotie et le nombre de duplicats. Il permet aussi de vérifier le sexe des individus et le degré d'apparentement entre chaque individu permettant ainsi un contrôle d'identitovigilance. Nous avons opté pour ce package car il permet de réaliser une analyse en composante principale de l'origine géographique des échantillons. C'est un programme écrit en Python, simple à mettre en œuvre puisqu'il utilise les fichiers de sortie générés par les outils de séquençage d'ADN couramment utilisés tel que GATK. Pour l'analyse biogéographique, Peddy utilise comme contrôles les données de génome des 2500 individus inclus dans le Projet 1000Genomes.

Dans les études sur les maladies rares, la cartographie de régions homozygotes par descendance (HBD) chez des cas consanguins et très utile pour identifier des variations pathogènes de transmission récessive. En effet un cas consanguin ayant des parents cousins est aussi informatif qu'une fratrie de 3 cas non consanguins germains (LOD score=1.2), cette dernière situation étant très improbable avec une taille de fratrie le plus souvent < 3 puisqu'on attend seulement 25% d'atteints dans une fratrie pour une maladie à transmission récessive. L'information sur la généalogie d'un individu n'est cependant pas toujours disponible. Le coefficient de consanguinité F a donc été estimé pour chaque fœtus à partir de ses données d'exome entier en mettant en œuvre une méthode de maximum de vraisemblance qui utilise une chaîne de Markov cachée (Leutenegger *et al.,* 2003). Un test du rapport de vraisemblance a été utilisé pour vérifier si le F estimé était significativement différent de zéro. Pour les fœtus consanguins, les segments comportant au moins 5 positions consécutives avec des probabilités HBD > 50 % ont été considérés comme des segments HBD. Toutes ces analyses ont été effectuées avec le package R Fantasio (https://github.com/genostats/Fantasio).

Avec controles des onnées selon les données décrites par les cliniciens.

3) Analyse des variants rares

Environ 30 000 variants par individu sont identifiés par WES. Afin de réduire le nombre de faux positifs, les variants sont d'abord filtrés sur la base de critères de qualité incluant la couverture et la balance allélique. Sont ensuite exclus les variants synonymes et les variants fréquents dans les bases de données de contrôles (fréquence dépendant du mode de transmission envisagé) pour aboutir à une liste de variants rares. Des stratégies supplémentaires sont appliqués pour identifier la variation causale parmi cette liste. Les principales stratégies sont résumées dans la figure suivante (Figure 21).

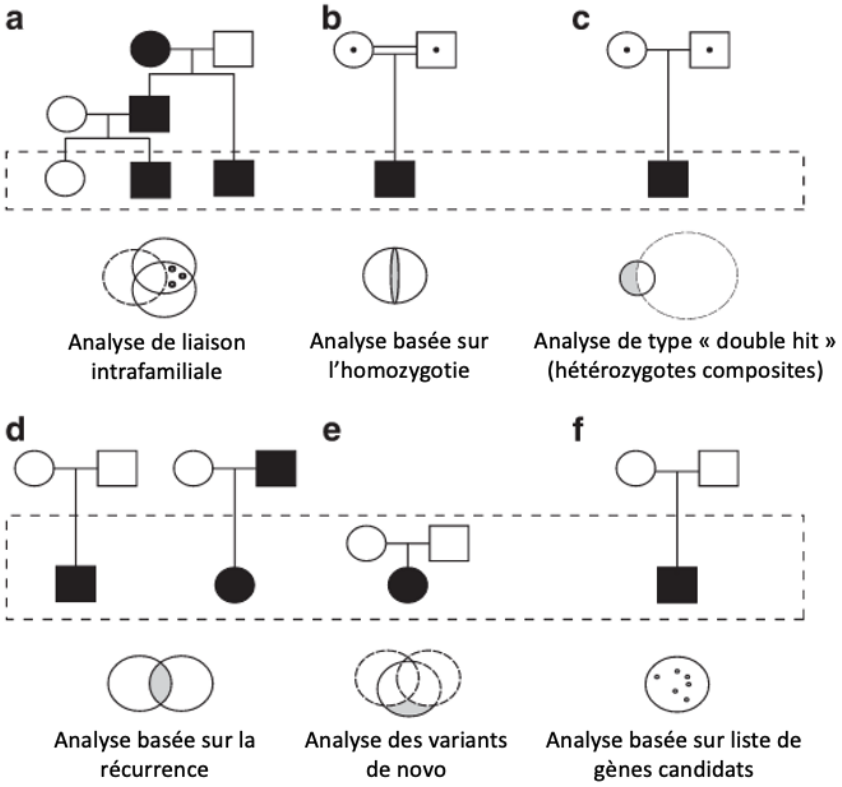


Figure 21. Stratégies d'identification de gènes responsables de maladies par WES. Adaptée de Gilissen *et al.*, 2012.

Nous avons appliqué ces stratégies (détaillées ci-après) pour l'analyse des données de WES des 113 fœtus à l'exception de l'analyse de liaison puisque nous ne disposions pas de grande famille informative.

3-a) Criblage par liste de gènes impliqués dans des voies candidates

- **Gènes impliqués dans l'homéostasie plaquettaire**

Basé sur le constat qu'une thrombopénie chez le fœtus peut induire une HIC fœtale comme observée dans l'alloimmunisation plaquettaire fœto-maternelle, nous avons émis l'hypothèse que des causes génétiques responsables d'une thrombopénie ou thrombopathie pouvaient être détectées dans notre cohorte de fœtus. Nous avons ainsi constitué une liste de gènes candidats en se basant sur les données de la littérature. Nous avons généré une liste de 64 gènes responsables d'anomalies héréditaires des plaquettes (Almazni *et al.,* 2019 ; Bastida *et al.,* 2018 ; Downes *et al.,* 2019 ; Heremans et Freson, 2018 ; Noris et Pecci, 2017 ; Nurden *et al.,* 2018 & 2020 ; Seo *et al.,* 2019). Sur les données de WES de 101 fœtus index (nombre d'index disponibles sur les 113 à la date ou été faite cette analyse) nous avons recherché la présence de variants rares et prédits délétères par PolyPhen-2 dans ces 64 gènes. Nous avons utilisé des seuils de fréquences alléliques de 1/100 sous l'hypothèse d'un mode de transmission autosomique récessif (variants homozygotes et hétérozygotes composites) et de 1/1000 sous une hypothèse dominante (fréquences dans gnomAD v2.1). Les variants ont ensuite été classés selon les recommandations de l'ACMG (Richards *et al.,* 2015).

- **Gènes intervenant dans la biosynthèse de COL4A1/COL4A2**

Devant l'association de variants perte de fonction dans les gènes *COL4A1* et *COL4A2* et la survenue d'hémorragie intracrânienne chez certains patients nous avons émis l'hypothèse que des variations pathogènes affectant les gènes impliqués dans la biosynthèse de ces collagènes pouvaient être impliqués dans les HIC fœtales. Nous avons généré une liste de 47 gènes candidats codant pour les protéines impliquées dans la biosynthèse de l'hétérotrimère de COL4A1 / COL4A2 basé sur les données de la littérature (Mao *et al.,* 2015) et sur des données expérimentales (protéomique) fournies par le Dr Roberto Vanacore dans le cadre d'un projet collaboratif financé par le *National Institutes of Health* (NIH). Sur les données de WES des 113 fœtus, nous avons cherché la présence de variants candidats dans cette liste de gènes. Nous avons pris comme filtre une fréquence allélique < 1 / 200 dans la population gnomAD en hypothèse récessive et 1 / 1000 en hypothèse dominante. Pour les variants de type faux-sens nous avons ajouté un critère de sélection supplémentaire, celui d'être prédit délétère par le logiciel de prédiction in silico PolyPhen-2

3-b) Analyse sans a priori des variants *de novo*

L'avènement des analyses pangénomiques au cours des dernières années a rendu possible l'identification de nombreux gènes causaux chez des sujets atteints d'affections sporadiques par l'approche dite en trios (Gillisen *et al.*, 2012). Cette analyse a été effectuée chez les 35 trios.

Les variants candidats ont été sélectionnés à l'aide des critères suivants :

- Qualité des variants détectés : un minimum de 10X était requis pour la profondeur de lecture et un minimum de 25% pour la balance allélique des variants hétérozygotes.

- Nature des variants : nous avons limité notre analyse aux variants disruptifs susceptibles d'entrainer une perte d'expression de l'allèle muté ou l'expression d'une protéine tronquée (variants non-sens, variants avec décalage du cadre de lecture, variants de sites d'épissage canoniques) et aux substitutions de type faux sens situés dans les régions codantes

-Variant présent chez le cas index et absent chez les parents (avec une bonne couverture de la région)

- Fréquence allélique : MAF (*minor allele frequency*) < 1/1000 dans les bases de données de contrôles (gnomAD v2.1.1 et 1000 Genomes Phase 3)

- Prédiction de la pathogénicité pour les variants faux-sens : prédits comme possiblement ou probablement délétère selon le logiciel PolyPhen-2 (Adzhubei *et al.*, 2015).

Cette stratégie de filtrage nous a permis de retenir plusieurs variants candidats. Pour rechercher d'autres variants candidats dans les gènes issus de l'analyse des 35 trios, une étude de récurrence avec les mêmes filtres mentionnés précédemment a été réalisée sur les données d'exome des 78 fœtus solos (variants pour lesquels le statut hérité ou non n'est pas connu) (Figure 22).

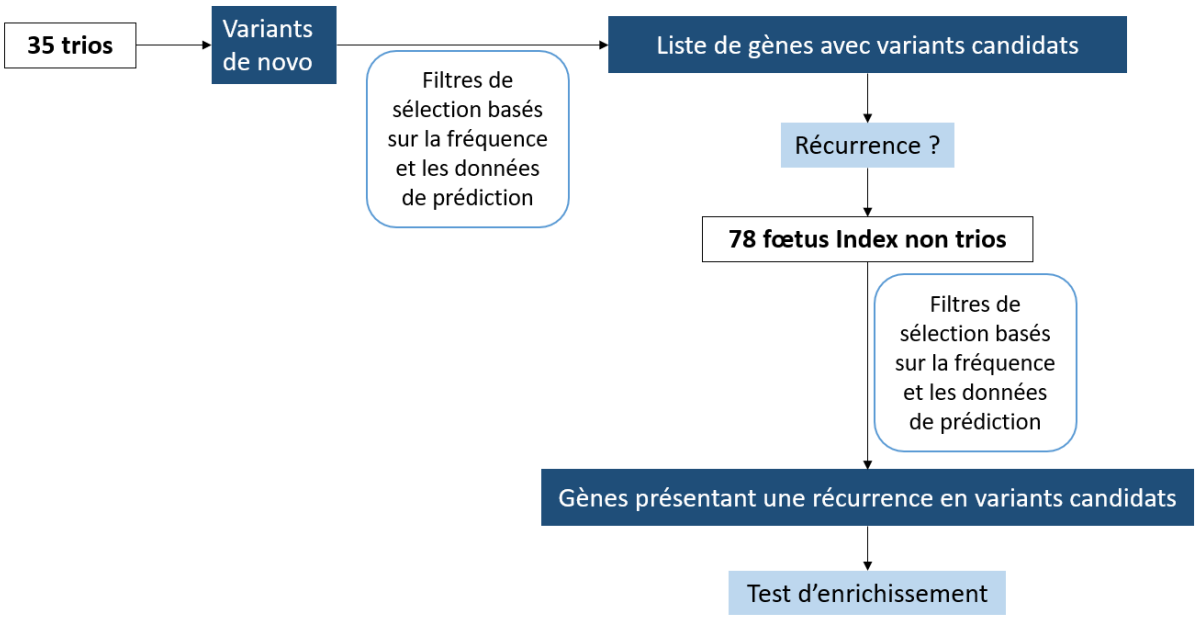


Figure 22. Stratégie d'analyse des variants *de novo*.

3-c) Analyse sans à priori des variants selon une hypothèse de transmission autosomique récessive ou récessive liée à l'X

L'analyse des hétérozygotes composites n'a pu être effectuée que chez les fœtus dont les parents ont été séquencés, à savoir les 35 trios et les 3 familles à cas multiples. La recherche de variants candidats homozygotes a été réalisée chez les 113 fœtus index.

Les variants ont été sélectionnés en utilisant les filtres suivants :

- Qualité des variants : un minimum de 10X était requis pour la profondeur de lecture et un minimum de 25% pour la balance allélique

- Nature des variants : variants disruptifs et substitutions de type faux sens situées dans les régions codantes

- MAF < 1/200 dans les bases de données externes (gnomAD v2.1.1 et 1000 Genomes Phase 3)

- Prédiction de pathogénicité pour les variants faux-sens : prédits comme possiblement ou probablement délétères selon le logiciel PolyPhen-2

La recherche de variants hémizygotes candidats a été réalisée chez les 58 fœtus index de sexe masculin. Les variants ont été sélectionnés en utilisant les critères suivants :

- Variants présents sur le chromosome X

- Qualité des variants : un minimum de 10X était requis pour la profondeur de lecture et un minimum de 25% pour la balance allélique

- Nature des variants : variants disruptifs et substitutions de type faux sens situées dans les régions codantes

- MAF < 1/10 000 dans les bases de données externes (gnomAD v2.1.1 et 1000 Genomes Phase 3)

- Prédiction de la pathogénicité pour les variants faux-sens : prédits comme possiblement ou probablement délétère selon le logiciel PolyPhen-2

3-d) Analyse des variants selon une hypothèse autosomique dominante à pénétrance incomplète

L'identification des causes génétiques reste difficile pour certaines maladies rares en raison de divers facteurs tels que l'absence de familles à cas multiples, l'hétérogénéité génétique et la pénétrance incomplète. Ces variants échappent aux approches classiques de type GWAS (variants trop rares pour être détectés) et de type analyse de liaison (impact clinique trop modéré). Par ailleurs, la rareté des variants à pénétrance incomplète imposerait de travailler sur de grandes cohortes de cas et de témoins pour obtenir la puissance statistique nécessaire à l'échelle d'un variant unique ce qui est impossible pour certaines maladies rares. Enfin, l'analyse des arbres généalogiques de notre cohorte de fœtus avec HIC montre une proportion importante de cas sporadiques sans données génotypiques parentales pour les 2/3 d'entre eux (ADN parentaux non disponibles), ne permettant pas d'établir la nature *de novo* ou héritée des variants rares identifiés. Il est très difficile de conclure pour les variants hétérozygotes hérités ou les variants pour lesquels le statut *de novo* ou hérité n'est pas connu.

Pour pallier à cette problématique, nous avons décidé d'utiliser une approche dite de test de charge à l'échelle du gène « *Gene-based burden test* » qui est une analyse d'enrichissement en variants rares et prédits délétères (candidats) pour chacun des 20 000 gènes codants connus chez l'humain selon le principe illustré dans la Figure 23. Un test exact de Fisher, comparant pour chacun des gènes le nombre de porteurs de variants candidats situés dans le transcrit canonique chez les cas et chez les témoins, est effectué sous l'hypothèse d'une transmission

autosomique dominante (MAF < 1/10 000). Le test donne une p-value pour chaque gène. Le seuil de significativité de cette analyse statistique est défini après ajustement du risque alpha par la méthode de Bonferroni, afin de corriger l'inflation de ce dernier par le nombre de tests réalisés. Le seuil de significativité pour cette analyse à l'échelle du gène est de 2,5 x 10-6 (0.05/20 000).

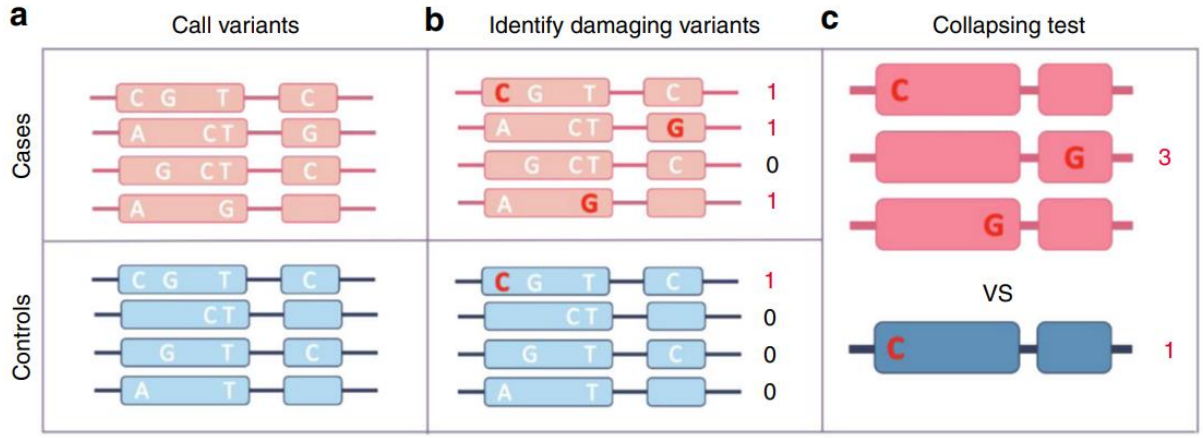


Figure 23. Schéma illustrant le principe du test de charge en variants candidats à l'échelle du gène. A) Identification des variants de chaque gène par séquençage. B) Sélection des variants candidats basée sur leur nature, leur rareté et les données de prédiction in silico. C) Comparaison du nombre de cas porteurs d'un variant candidat par rapport au nombre de témoins également porteurs d'un variants candidats au moyen d'un test statistique. Extrait de Cirulli *et al.*, 2020.

La réalisation de tests d'association impose une rigueur particulière puisque les cas et les contrôles doivent être issus de la même population afin d'éviter la détection de fausses associations résultant d'un biais de stratification (Cirulli, 2016). Cependant de nouvelles méthodes ont été développées pour s'affranchir de ce bais si l'on utilise un set de sujets contrôles externe, évitant ainsi la nécessité d'un séquençage de contrôles supplémentaires. C'est le cas de la méthode TRAPD (*Test Rare vAriants with Public Data*) qui permet d'effectuer des tests de charge en variants rares et prédits délétères en utilisant par exemple les données publiques de séquençage haut débit des individus de gnomAD (Guo *et al.*, 2018).

Sur les 113 fœtus index, un pipeline adapté de la méthode TRAPD par le Dr Chaker Aloui (Bio informaticien dans l'unité INSERM UMR 1141) a été lancé pour l'identification de gènes présentant un enrichissement en variants candidats dans la cohorte de fœtus. Les paramètres utilisés pour cette analyse ont été calibrés sur des données antérieures en utilisant les variants

synonymes rares (< 1%) pour identifier les meilleurs critères (filtres de qualité des variants (PASS, QD et VQSLOD) et seuil minimum de couverture pour les régions analysées entre les cas et les témoins) à utiliser pour rendre les données séquencées dans des plateformes différentes et traitées par des pipelines bioinformatiques différentes comparables.

Description de la méthode utilisée :

Les données agrégées des 71 702 individus de gnomAD v3.0 ont été utilisées comme cohorte de témoins. Pour chaque gène testé, il a été vérifié que les exons étaient couverts à plus de 10X dans au moins 90% des échantillons. Les données statistiques résumées sur la couverture des génomes ont été téléchargées à partir de https://gnomad.broadinstitute.org/downloads.

1) Ré-annotation

Pour les besoins des tests de charge, les données de génotypes des fichiers de variants bruts (VCF) pour les fœtus et les témoins ont été réannotés avec une version actuelle de *Ensembl Varaiant Effect Predictor* (VEP104) pour obtenir les données de MAF et de prédiction *in silico* de chaque variant.

2) Sélection des variants candidats

A l'aide de scripts bash et Python, les variants candidats ont été sélectionnés sur plusieurs critères :

-Site du variant avec une couverture minimale de 10X chez au moins 90% des fœtus et des témoins

-Qualité du variant (filtre PASS et score du QD>8)

-Nature du variant : variant disruptif ou faux sens

-Fréquence allélique : MAF < 1/10 000 dans gnomAD v2.1

-Données de prédiction pour les variants faux sens : prédit « *possibly damaging* » ou « *probably damging* » par PolyPhen-2

3) Test statistique

Pour chaque gène, le nombre d'individus porteurs d'au moins un variant candidat a été compté dans les 2 cohortes. Un tableau de contingence 2x2 a été généré pour chaque gène, représentant le nombre de cas et de sujets témoins porteurs et non porteurs d'un variant candidat dans ce gène. Les p-values ont été calculées à l'aide d'un test exact de Fischer bilatéral implémenté dans R (https://www.r-project.org), et le seuil de significativité a été établi à 2,5 x 10^{-6} selon la correction de Bonferroni.

4) Identification des variations du nombre de copies (CNV)

De nombreux outils ont été mis au point pour la détection des CNV à partir des données de WES et plusieurs études ont déjà évalué et comparé leurs performances (Zhao et al., 2020 ; Quenez et al., 2021). Certains outils sont plus sensibles mais produisent plus de faux positifs. D'autres détectent moins de CNV mais ont une plus grande précision. Nous avons opté pour une combinaison de deux outils d'analyse des CNV : le logiciel CANOES (Backenroth *et al.*, 2014) afin d'identifier de façon ponctuelle des CNV dans des gènes candidats et le logiciel HMZDelFinder pour rechercher des CNV homozygotes chez les fœtus consanguins (Gambin *et al.*, 2017).

4-a) Analyse avec l'outil CANOES

La détection des CNV avec cet outil est basée sur les profondeurs de lecture et les informations relatives au contenu en GC de chaque cible afin de réduire la variabilité observée dans les données de séquençage haut débit. La profondeur de lecture est calculée à l'aide de l'outil Bedtools (http://bedtools.readthedocs.io/en/latest/content/bedtools-suite.html), et le contenu en GC est déterminé à l'aide du logiciel GATK (https://software.broadinstitute.org/gatk). Conformément aux recommandations (Quenez *et al.*, 2021) les régions cibles proches (<30 paires de bases) ont été fusionnées afin de réduire le taux de faux positifs. Après comptage des lectures, les cibles mal couvertes (moins de 10 lectures chez 90% d'exomes) ont été éliminées. Pour réduire le nombre de faux positifs liés au bruit de fond, avoir un ensemble d'échantillons servant de contrôles et construire un modèle statistique valide, la méthode CANOES doit être effectuée sur des données brutes de WES les plus homogènes possible. Le logiciel CANOES a ainsi été utilisé par lots, en regroupant les exomes en fonction des kits de capture et des lots de séquençage utilisés. Les CNV ont été appelés en utilisant la fonction CallCNVs du logiciel CANOES. Cette fonction attribue à chaque CNV détecté un score de qualité Phred allant de 1 à 99. Elle attribue également un score "NA" aux CNV provenant d'échantillons comportant plus de 50 CNV. Les résultats relatifs à ces échantillons ont donc été considérés comme aberrants et exclus des analyses ultérieures.

Par défaut, l'algorithme du logiciel CANOES permet seulement d'analyser les CNV présents sur les autosomes (chromosomes 1 à 22) ; le script a été modifié par le Dr Chaker Aloui (Bio-informaticien dans l'unité INSERM UMR 1141) pour étendre l'analyse et inclure le chromosome X. L'algorithme a été lancé séparément chez les fœtus de sexe masculin et ceux de sexe féminin.

4-b) Analyse avec l'outil HMZDelFinder

Nous avons choisi cet outil car il a déjà fait ses preuves pour la détection de CNV homozygotes pathogènes (Mitani *et al.*, 2021 ; Gordeeva *et al.*, 2021 ; Chen *et al.*, 2022). Il est par ailleurs simple à utiliser (écrit intégralement dans R). C'est un algorithme mis au point par Gambin *et al.* (2017) afin d'identifier des délétions intragéniques homozygotes rares. Comme pour CANOES, il analyse les données de WES en se basant sur la profondeur de lecture mais peut aussi utiliser les données de zygotie provenant du fichier VCF de chaque échantillon ce qui le rend plus spécifique que CANOES pour la détection des CNV homozygotes. Les informations sur le nombre de lectures provenant des fichiers BAM sont appelées conjointement à partir de tous les échantillons de la cohorte, ce qui permet d'identifier les délétions exoniques potentielles homozygotes ou hémizygotes rares, tout en permettant d'exclure les exons dont la profondeur de couverture est faible.

Brièvement, chaque fichier BAM individuel des fœtus consanguins a été transformé en un fichier de profondeurs de lectures par exon (*reads per thousand base pairs per million reads* ; RPKM) à l'aide d'un script dans R développé par les auteurs et utilisant la fonction « featureCount » du package Rsubread et du fichier BED comportant toutes les régions exoniques séquencées. Les étapes suivantes comportaient l'utilisation de filtres permettant l'élimination des exons/régions d'intérêt avec un RPKM médian faible, une mauvaise qualité ou des CNV très fréquents dans la cohorte (nécessite au minimum 10 échantillons utilisés comme contrôles). Dans l'étape finale, les CNV provenant d'exons consécutifs ont été fusionnés et les CNV de moins de 50 paires de bases ont été exclus.

Cet outil n'a été utilisé que pour les fœtus consanguins pour identifier des CNV homozygotes ou hémizygotes dans les régions homozygotes par descendance.

5) Analyses complémentaires

Une visualisation systématique des variants retenus a été effectuée sur les données BAM via l'outil IGV (https://igv.org/) afin d'exclure ceux de mauvaise qualité.

- **Séquençage Sanger**

La validation des variations a été réalisée par séquençage capillaire (méthode de Sanger). Celui-ci est nécessaire pour s'assurer de la présence du variant chez le patient (identitovigilance) et d'écarter un possible artéfact (essentiellement pour les insertions et délétions, qui sont des faux

positifs fréquents avec les méthodes de séquençage à haut débit). Le choix des amorces de PCR et de séquençage a été réalisé grâce au logiciel Primer3web version 4.1.0 (http://primer3.ut.ee/) en se basant sur les séquences génomiques de références RefSeq NCBI. Le séquençage a été réalisé en utilisant le kit BigDyeTerminator v3.1, puis une migration électrophorétique a été faite sur un séquenceur Applied Biosystems 3130 ou 3500. Les séquences ont été analysées avec le logiciel SeqScape v2.5 (Applied Biosystems) par une inspection visuelle.

- **PCR allèle spécifique**

Pour la détermination de la position allélique en trans ou en cis de variants candidats dans un gène donné chez un même individu, une PCR allèle spécifique a été réalisée en utilisant des amorces de séquences spécifiques à l'aide du logiciel Primer3.

- **Analyse de microsatellites**

La nature *de novo* des variants a été confirmée par l'analyse de microsatellites sur l'ADN du fœtus et des parents à l'aide du kit d'amplification PCR AmpFISTR Profiler (Applied Biosystems) afin de confirmer la parentalité / le lien de parenté pour tous les trios et familles à cas multiples.

6) Implication des variants candidats

Une analyse détaillée a été effectuée pour chaque gène candidat identifié, en interrogeant diverses bases de données ainsi que la littérature. Cette analyse, incluait une analyse des données d'expression à la recherche notamment d'une expression dans les tissus vasculaires, et des données de structure et de fonction de la protéine pour étudier si les variants candidats étaient localisés dans un domaine important. En plus de PolyPhen-2, les données d'autres logiciels de prédiction de pathogénicité *in silico* dont SIFT et CADD (Ng et Henikoff, 2003 ; Kircher *et al.,* 2014) score ont été prises en compte. Un score de pathogénicité a été établi conformément aux recommandations de l'American College of Medical Genetics and Genomics (ACMG) (Richards *et al.,* 2015).

Une éventuelle implication dans des pathologies humaines a été recherchée en utilisant OMIM (Online Mendelian Inheritance in Man ; www.omim.org) et PubMed, de même que l'existence de modèles animaux via les bases de données MGI (Mouse Genome Informatics ; www.informatics.jax.org) et ZFIN (Zebrafish Information Network ; www.zfin.org) (Figure

24). Nous avons également utilisé GeneMatcher (www.genematcher.org) (Sobreira *et al.*, 2015) pour les gènes pour lesquels nous n'avons pas de récurrence et dont les informations phénotypiques chez l'humain étaient limitées. Enfin pour l'étude de réseaux de gènes nous avons utilisé la base de données STRING (https://string-db.org/) avec comme paramètres un score d'interaction de « high confidence » (> 0,7) et en prenant en considération toutes les sources d'interactions (sources expérimentales, bases de données, co-expression, proximité génomique, fusion de gène, co-occurrence).

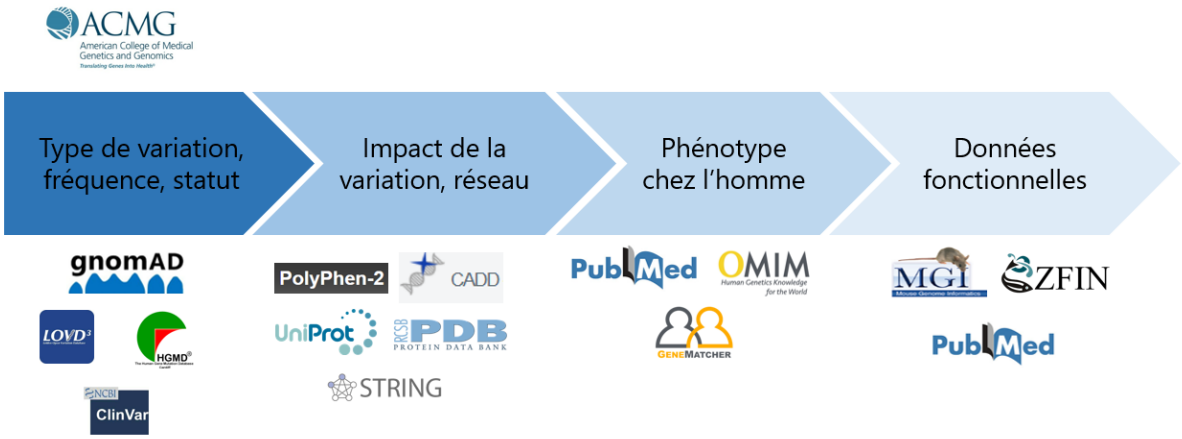


Figure 24. Illustration des grandes étapes de la collecte des arguments pour l'implication des variants / gènes candidats en utilisant divers outils disponibles publiquement en ligne.

Résultats

A- Variants pathogènes de *COL4A1/COL4A2* et variants dans les gènes impliqués dans des anomalies plaquettaires héréditaires chez les fœtus présentant une HIC

INTRODUCTION

Sur la période de 2010 à 2021, 216 fœtus index (194 index au moment de la publication de l'article) ont été référés au service de génétique moléculaire neurovasculaire de l'hôpital Saint Louis pour suspicion d'une collagénopathie COL4A1/A2. Un criblage moléculaire ciblé des gènes *COL4A1* et *COL4A2* a été réalisé et nous rapportons dans la première partie de cet article les données moléculaires des fœtus porteurs d'un variant pathogène dans *COL4A1* ou *COL4A2*.

Un séquençage de l'exome entier a ensuite été réalisé pour 113 de ces fœtus (101 au moment de la publication) sans variation pathogène de *COL4A1/COL4A2* et pour lesquels un consentement parental a été obtenu. En nous basant sur ces données d'exome entier, nous avons dans une première étape recherché des variants candidats dans une liste de gènes associés à des troubles plaquettaires héréditaires (IPD = *inherited platelet disorders*), groupe hétérogène de maladies rares caractérisées par des défauts dans la production et/ou la fonction plaquettaire. Certains gènes d'IPD codent pour des protéines jouant un rôle essentiel dans la production des plaquettes circulantes et en cas de variations pathogènes conduisent à une thrombocytopénie héréditaire. D'autres peuvent affecter les protéines du cytosquelette, la formation de granules, la réponse des plaquettes aux agonistes ou les interactions avec les protéines de la matrice extracellulaire (exposées par la rupture de la barrière endothéliale). Ces troubles plaquettaires héréditaires se traduisent par une tendance accrue aux hémorragies durant toute la vie. Nous avons émis l'hypothèse que certains fœtus pourraient être porteurs de variants pathogènes dans les gènes responsables d'IPD. Pour tester cette hypothèse, nous avons généré une liste de 64 gènes candidats par analyse des données de la littérature. Nous avons ensuite cherché la présence de variants candidats dans ces gènes sur les données de WES.

L'ensemble des résultats de cette étude a donné lieu à une publication dans *Prenatal Diagnosis* en 2022, présentée ci-après.

Received: 22 November 2021 | Revised: 12 January 2022 | Accepted: 9 February 2022

DOI: 10.1002/pd.6113

ORIGINAL ARTICLE

PRENATAL DIAGNOSIS WILEY

COL4A1/COL4A2 and inherited platelet disorder gene variants in fetuses showing intracranial hemorrhage

Thibault Coste[1,2] | Catherine Vincent-Delorme[3] | Morgane Stichelbout[4] | Louise Devisme[4] | Antoinette Gelot[5] | Igor Deryabin[5] | Fanny Pelluard[6] | Chaker Aloui[2] | Anne-Louise Leutenegger[2] | Jean-Marie Jouannic[7] | Delphine Héron[8] | Douglas B Gould[9] | Elisabeth Tournier-Lasserve[1,2]

[1]AP-HP, Service de génétique moléculaire Neurovasculaire, Hôpital Saint-Louis, Paris, France

[2]Université de Paris, INSERM UMR-1141 Neurodiderot, Paris, France

[3]CHU Lille, Service de génétique clinique Guy Fontaine, Lille, France

[4]CHU Lille, Institut de pathologie, Lille, France

[5]APHP, Service de fœtopathologie, Hôpital Trousseau, Paris, France

[6]University Bordeaux, INSERM, BaRITOn, U1053, Bordeaux, France

[7]APHP Sorbonne Université, Service de médecine fœtale, Hôpital Trousseau, Paris, France

[8]AP-HP, Service de génétique clinique, Hôpital de la Pitié-Salpêtrière, Paris, France

[9]Department of Ophthalmology, University of California San Francisco, San Francisco, California, USA

Correspondence
Elisabeth Tournier-Lasserve, INSERM UMR-1141 Neurodiderot, Hôpital R. Debré, Paris F-75019, France.
Email: tournier-lasserve@univ-paris-diderot.fr

Funding information
National Institutes of Health, Grant/Award Number: R01NS096173

Abstract

Background: Variants of *COL4A1/COL4A2* genes have been reported in fetal intracranial hemorrhage (ICH) cases but their prevalence and characteristics have not been established in a large series of fetuses. Fetal neonatal alloimmune thrombocytopenia is a major acquired ICH factor but the prevalence and characteristics of inherited platelet disorder (IPD) gene variants leading to thrombocytopenia are unknown. Herein, we screened *COL4A1/COL4A2* and IPD genes in a large series of ICH fetuses.

Methods: A cohort of 194 consecutive ICH fetuses were first screened for *COL4A1/COL4A2* variants. We manually curated a list of 64 genes involved in IPD and investigated them in *COL4A1/COL4A2* negative fetuses, using exome sequencing data from 101 of these fetuses.

Result: Pathogenic variants of *COL4A1/COL4A2* genes were identified in 36 fetuses (19%). They occurred de novo in 70% of the 32 fetuses for whom parental DNA was available. Pathogenic variants in two megakaryopoiesis genes (*MPL* and *MECOM* genes) were identified in two families with recurrent and severe fetal ICH, with variable extraneurological pathological features.

Conclusion: Our study emphasizes the genetic heterogeneity of fetal ICH and the need to screen both *COL4A1/COL4A2* and IPD genes in the etiological investigation of fetal ICH to allow proper genetic counseling.

Key points

What's already known about this topic?

- *COL4A1/COL4A2* pathogenic variants have been reported in several fetal intracranial hemorrhage (ICH) case reports but their prevalence and characteristics in a large series of fetal ICH is lacking.
- Fetal neonatal alloimmune thrombocytopeniais a well-known cause of thrombocytopenia and ICH in infants and fetuses but very little is known regarding the role in fetal ICH of variants of inherited platelet disorder genes leading to thrombocytopenia.

What does this study add?

- Fetal ICH is a highly heterogeneous condition with *COL4A1/COL4A2* pathogenic variants accounting for 19% of cases with a very high de novo rate.
- Albeit rare, pathogenic variants of megakaryopoiesis genes are associated with ICH and screening of these genes should be performed in fetal ICH etiological investigation.

1 | INTRODUCTION

Fetal intracranial hemorrhage (ICH) refers to bleeding that occurs antenatally from a blood vessel into the ventricles, subdural space or brain parenchyma. Whereas neonatal hemorrhage is a relatively common occurrence in preterm infants delivered before 32 weeks of gestation (WG), fetal ICH has an unclear incidence, although an estimate of 1 in 10,000 pregnancies has been suggested.[1] Prenatal diagnosis of fetal ICH is generally made during the second or third trimester by either ultrasound (US) or magnetic resonance imaging (MRI). Fetal ICH has been classified in five types according to the anatomic location of the bleeding: intraventricular (IVH), subarachnoid, intraparenchymal, cerebellar and subdural hemorrhage. IVH are further subdivided in four grades for which extensive hemorrhages are classified as grade III (hemorrhage with ventriculomegaly) and IV (hemorrhage within the cerebral parenchyma).[2] Both are associated with a very poor perinatal outcome including intrauterine fetal death (IUFD), or severe neurodevelopmental anomalies that may lead to a termination of pregnancy (TOP) decision.[3]

Fetal ICH can be associated with several acquired maternal and fetal factors including maternal trauma, infection, drug use and alloimmune thrombocytopenia. Among those acquired causes, fetal and neonatal alloimmune thrombocytopenia (fetal neonatal alloimmune thrombocytopenia [FNAIT]) is a well-established etiology of profound thrombocytopenia in fetuses or neonates and is responsible of severe bleedings especially in the brain[4]. On the other hand, variants of *COL4A1* and *COL4A2* genes, which encode basement-membrane proteins, have been reported in several fetal ICH case reports since 2011.[5] A recent meta-analysis conducted on PubMed and EMBASE databases ranging from 1980 to 2020 identified a total of 20 mutated fetal cases.[6] A few additional genes have been shown to be mutated in very rare ICH fetal cases including clotting factors V and VII, protein C and VWF.[6] In a recent study on fetal or neonatal ICH including 9 cases with a prenatal diagnosis and 17 with a postnatal diagnosis, Hausman-Kedem et al.[7] identified a causative/likely causative variant in 12% of them and emphasize the importance of these findings for genetic counseling. However, in most fetal ICH cases the underlying cause is not identified, precluding any genetic counseling.

In this study, we investigated the prevalence and characteristics of *COL4A1* and *COL4A2* gene variants in a cohort of 194 consecutive fetuses for whom TOP was performed or IUFD occurred due to fetal ICH and for whom acquired causes have been excluded. We also generated a list of 64 genes known to be involved in inherited platelet disorder (IPD) by manual curation of the literature and investigated the presence of pathogenic variants in those genes using whole exome sequencing (ES) data from 101 *COL4A1/COL4A2* negative fetuses.

2 | MATERIALS AND METHODS

2.1 | Case ascertainment and inclusion

From 2010 to 2020 a series of 194 consecutive fetuses interrupted following either TOP or IUFD were referred for *COL4A1* and *COL4A2* genes screening to our molecular genetics reference center in France (national reference center for inherited cerebrovascular disorders, Saint Louis Hospital). Fetal ICH was in most cases detected at systematic second and third trimester US examination and most cases were confirmed by fetal MRI and/or pathological examination. Systematic review of medical charts was performed in order to exclude fetuses with an identifiable cause or known risk factor for ICH. These included evidence of maternal trauma during pregnancy, cocaine or maternal drug use, maternal or neonatal infections and fetal alloimmune thrombocytopenia. Written informed consent for genetic investigation and research was provided by both parents in accordance with the declaration of Helsinki and the French law. This study has been approved by the Inserm Ethics Committee (INSERM IRB00003888).

Genomic DNA was extracted from post mortem fetal tissue and from peripheral blood leukocytes of both parents and relatives, when available.

2.2 | Postmortem examination

Complete postmortem examination was performed with the informed consent of both parents in accordance with the French law and following standardized protocols including X-rays, photographs, macroscopical and histological examination of all viscera. Brains were fixed with formalin for 5–12 weeks. Macroscopic analysis was performed allowing the selection and conditioning of samples (paraffin embedding, 7-micron slicing, hematoxylin and eosin stain [HES]) of brain tissue for histological analysis.

2.3 | Sequencing of *COL4A1* and *COL4A2* genes

Sanger sequencing (fetuses referred in 2010–2015) or targeted high throughput sequencing of *COL4A1* and *COL4A2* genes (fetuses referred in 2016–2020) was performed for the 194 fetuses included in this study. Library preparation was performed using the SureSelect QXT® capture kit (Agilent technologies) and sequencing was performed on a MiSeq® sequencer (Illumina). We used SeqNext software 4.4 and 5.0 versions (JSI Medical Systems) to analyze the data. Variants were confirmed by Sanger sequencing using the BigDye™ Terminator v3.1 (Applied Biosystems) on an ABI 3130 DNA sequencer with in house primers (available upon request). Glycine missense variants or in-frame deletions in the triple helical domain, loss of function variants and rare de novo variants located in the non-collagenous (NC1) domain were classified as pathogenic variants. The de novo nature of variants was established by microsatellite analysis of both fetus and parental DNA using the AmpFISTR Profiler PCR Amplification Kit (Applied Biosystems).

2.4 | Investigation of genes leading to inherited platelet disorders

Several genes have been reported to be associated with an IPD, a heterogeneous group of rare disorders characterized by defects in platelet production and/or function. Some IPD genes encode proteins that play a critical role in platelet production leading to an inherited thrombocytopenia (IT), others may affect cytoskeletal proteins, formation of granules, platelet responses to agonists or interactions with extracellular matrix proteins. We manually curated literature data to identify all genes currently known to be involved in IPD.[8-17] A list of 64 IPD genes was generated (Supporting information S1).

Exome sequencing (ES) being a cost-effective approach to screen a large number of candidate genes, we used this strategy to screen the 64 IPD genes on a series of 101 unrelated fetuses without *COL4A1* or *COL4A2* variants and for whom ES research consents were obtained from parents and for whom DNA with a sufficient quality and quantity was available for ES (Figure 1). Libraries were prepared at the IntegraGen platform (Evry, France) for fetuses and their relatives using the SureSelect Human All Exon V5-UTR (Agilent technologies). Libraries were sequenced on a NovaSeq system (Illumina Inc). Data analysis of single-nucleotide substitutions and small insertion-deletion (indel) variants were performed with an IntegraGen bioinformatic pipeline that included standard tools as BWA and Haplotype Caller GVCF tool (GATK 3.8.1). Ensembl's VEP (Variant Effect Predictor, release VEP 95.1)

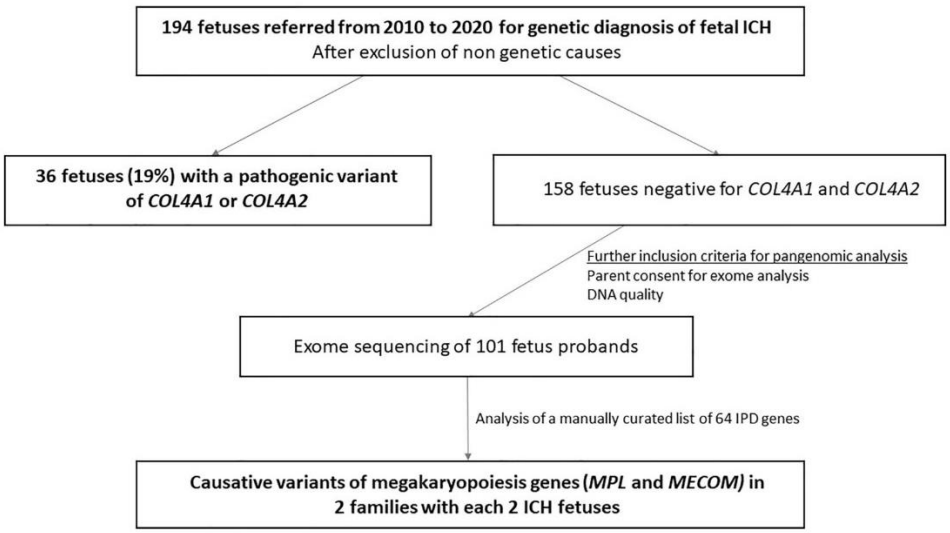


FIGURE 1 Study flowchart

program was used to process variants for further annotation. We used CANOES for the detection of copy number variation (CNV) on ES data.[18]

After removing variants with low quality score, variants in IPD genes were filtered out based on their frequency. We filtered out variants with a Minor Allele Frequency (MAF) ≥ 0.001 in the gnomAD v2.1.1 and 1000 Genomes phase 3 databases for genes known to be involved in dominant IPD and with a MAF ≥ 0.01 for genes with recessive inheritance. We restrained our analysis to high impact variants (nonsense, frameshift, and splice site variants) and missense substitutions located in coding regions. Missense candidate variants were scrutinized using the PolyPhen-2 in silico pathogenicity prediction software and we excluded variants predicted to be benign. All variants were finally classified using the consensus guidelines set out by the American College of Medical Genetics and Genomics (ACMG guidelines).[19] We considered as candidate variants only those classified as pathogenic or likely pathogenic. All candidate variants were confirmed by Sanger sequencing.

3 | RESULTS

3.1 | Identification of *COL4A1/COL4A2* pathogenic variants in 19% of fetuses

We identified a pathogenic variant in 36 out of the 194 unrelated fetuses referred for diagnosis. Thirty cases showed a pathogenic variant in *COL4A1* and six cases showed a pathogenic variant in *COL4A2*. A total of 35 distinct variants were identified in these patients (one *COL4A1* variant was shared by two fetuses). The vast majority of the 35 distinct pathogenic variants were located in the triple helical domain, including 27 glycine missense variants and 5 splice site variants. Two variants are located in the NC1 domain of *COL4A1* gene (Figure 2). *COL4A1* variants appeared to be clustered on the c-term half of the triple helix domain. We compared this distribution with the distribution of variants detected in consecutive adult proband patients referred to our laboratory for a diagnostic of vascular leukoencephalopathy (Supporting information S1). A significant difference in the location of the variants in the triple helix domain was observed (One-sided *t*-test, $p = 0.0028$; Supporting information S1).

For 32 fetuses out of the 36 with a *COL4A1/COL4A2* variant, parental DNA was available and molecular screening of parents' DNA showed that 22 of the 32 (70%) fetuses carried a de novo variant.

3.2 | Identification of ICH causative variants in inherited platelet disorder genes

Screening of genes involved in IPD disorders was performed on a series of 101 unrelated fetuses negative for *COL4A1/COL4A2*. Filtering on allele frequency and PolyPhen-2 in silico prediction allowed the identification of 68 candidate variants. We further applied ACMG criteria for each variant and classified 11 variants as pathogenic or likely pathogenic in 10 fetuses (Supporting information S1). Eight fetuses carried a heterozygous variant classified as likely pathogenic or pathogenic in *HPS3*, *HPS5*, *ITGA2B*, *ITGB3*, *LYST* and *VWF* gene. Inherited platelet disorders associated with those genes are mainly reported as autosomal recessive ones. Therefore, in the absence of any additional predicted pathogenic missense variant or CNV on the other allele, we could not formally establish the role of these variants in the ICH of these fetuses.

In the two remaining unrelated fetuses causative pathogenic variants were detected in 2 distinct genes. We identified two compound heterozygous variants in the *MPL* gene (c.305G>C [p.R102P] and c.1609C>T [p.R537W]) in the proband of family F1. This fetus belongs to a family with recurrent fetal ICH. Molecular screening of the second affected fetus detected both variants. Each variant was inherited from one distinct parent and none of the two healthy siblings carried both variants (Figure 3A). *MPL* encodes a transmembrane receptor belonging to the homodimeric class I receptor family, and its ligand is thrombopoietin (TPO). MPL plays a central role in the regulation of megakaryopoiesis by regulating platelet production in response to TPO.[20] The R102P variant is located in the EpoR ligand binding domain of MPL protein (Figure 3B). The variant R102P impairs cell surface expression of MPL protein preventing its interaction with TPO.[21] The R537W variant is located in the intracellular domain and is predicted to impair signal transduction of MPL.

Finally, we identified a heterozygous frameshift variant (c.1268del [p.S423Lfs*6]) leading to a premature stop codon in the *MECOM* gene in family F2. Sanger sequencing of his affected fetus sibling identified the same variant in *MECOM*. This heterozygous stop codon variant is located in exon 8 and expected to lead to a nonsense messenger RNA decay and a haploinsufficiency (Figure 4B). Parents' DNA was not available.

3.3 | Clinical and neuropathological features of fetuses with causative variants in megakaryopoiesis genes

Main pathological features of ICH fetuses in families F1 and F2 are summarized in Table 1. The proband (II-2) of family F1 was a male fetus from a second spontaneous pregnancy of clinically healthy nonconsanguineous parents (Figure 3A). His 35 years old mother had a splenectomy in childhood because of a hereditary spherocytosis. The first pregnancy was uneventful, and the mother gave birth to a healthy girl (II-1) although the mother underwent a post-partum hemorrhage. For the second pregnancy, US examination at 27 WG revealed a large bilateral ventriculomegaly consistent with an ICH diagnosis. Investigation by array comparative genomic hybridization analysis was normal but fetal blood cell count showed severe anemia and thrombocytopenia (hemoglobin 3 g/dl; platelets 15 Giga/L). FNAIT was suspected but there were no alloantibodies in the maternal serum. A TOP at 28 WG was performed at parental request. Fetal autopsy did not find any skeletal or visceral malformation.

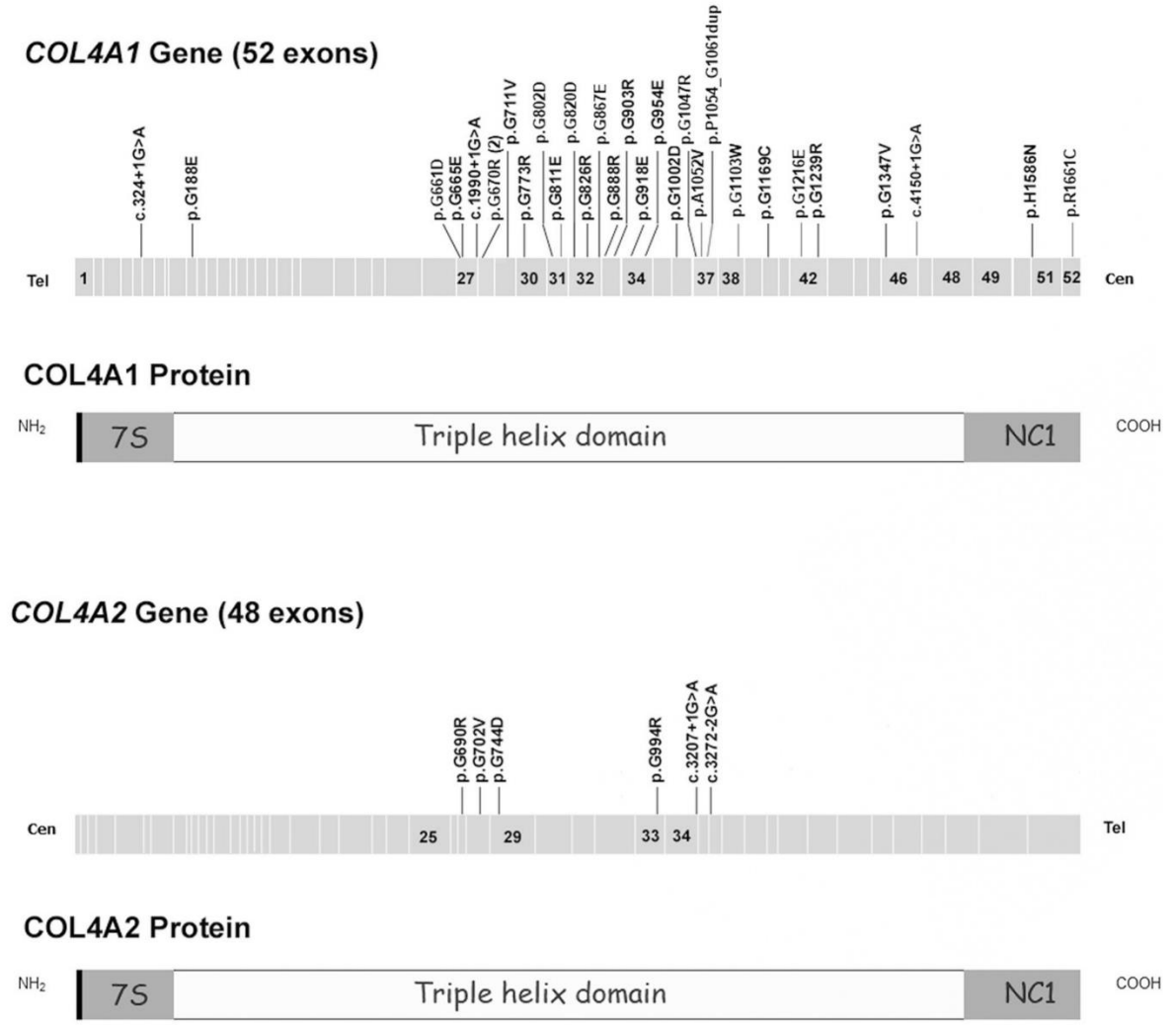


FIGURE 2 Distribution of *COL4A1* and *COL4A2* variants in fetal intracranial hemorrhage. Positions are annotated from the first coding nucleotide. Mature proteins are composed of three distinct domains: 7S, triple helix and non-collagenous (NC1). Cen: centromere side, Tel: telomere side

Neuropathological examination showed an asymmetry of the cerebral hemispheres and necrosis lesions in the right parietal region (Figure 3C *top-left*). Histological examination with HES showed suffusions in contact with the thick germinal zone; the edge of the ventricle ependymal layer was partially abraded and replaced by ependymal "pseudo-rosettes" (Figure 3C *top-right*). In association, we observed a diffuse ventricular and subarachnoid chronic hemorrhage, including hemosiderin pigments (Figure 3C *bottom-right*) and siderocalcium salts (Figure 3C *bottom-left*). Moreover, a microscopic right parietal parenchymal necrosis in the white matter and a focal parietal cortical polymicrogyria were noted, revealing inconspicuous ischemic lesions. There was not any sign of systemic hemorrhage. For the third spontaneous pregnancy, first trimester US examination performed at 11 WG revealed an increased nuchal translucency (3 mm). Comparative genomic hybridization analysis was normal. No fetal cell blood count could be performed at this early gestational age. IUFD occurred at 14 WG (II-3). An autopsy was performed, and the fetus presented advanced maceration. The brain was completely lysed; nevertheless, histological examination of the cerebral parenchyma identified a relatively bulky blood clot. Finally, in this family the fourth pregnancy was uneventful (II-4).

The proband of family F2 (III-1) was a male fetus from a first pregnancy of a non-consanguineous healthy couple (Figure 4A). Family history was uneventful apart the presence of a cleft palate in the sister's mother. Systematic US examination at 22 WG was normal except for a suspicion of a cardiac outflow tract anomaly. An US control at 27 WG finally revealed an IUFD. Fetal autopsy did not reveal any skeletal nor developmental abnormality except for a

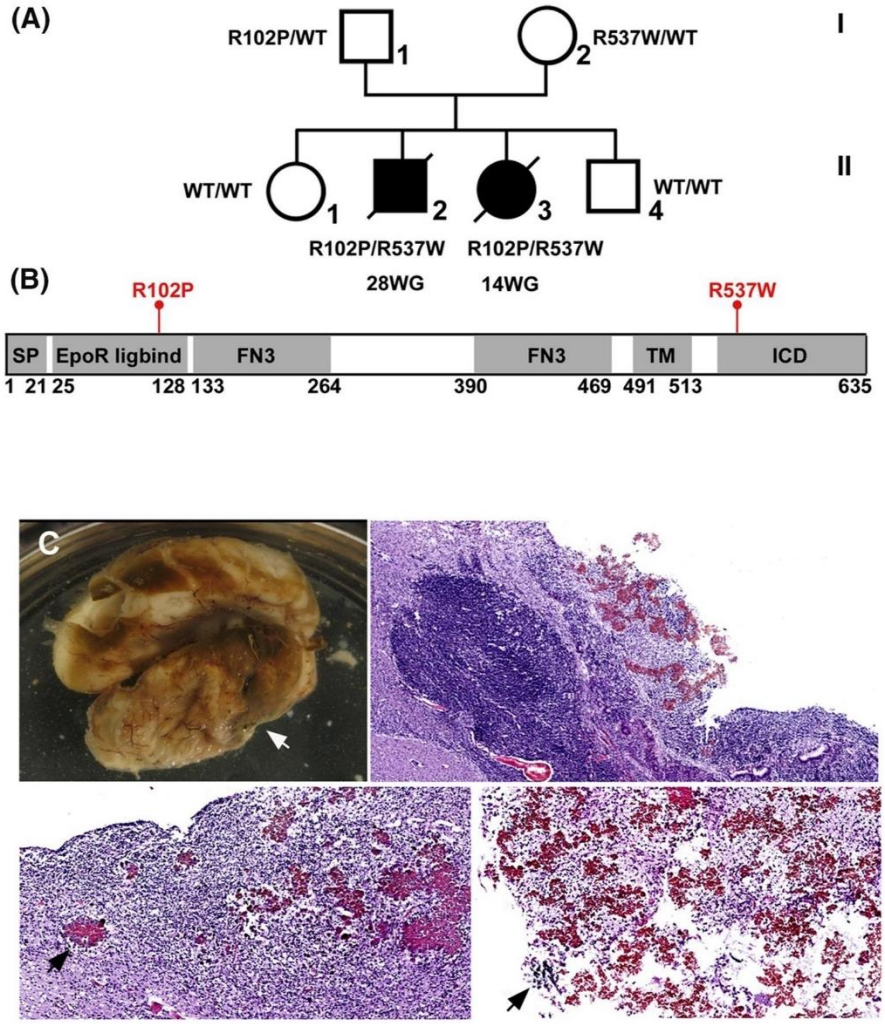


FIGURE 3 Family F1 with *MPL* pathogenic variants. (A) Family tree showing recessive inheritance of congenital amegakaryocytic thrombocytopenia (CAMT-MPL). Square = male; circle = female; diagonal black line = deceased individual; black filled symbol = clinically affected individual; empty symbol = clinically healthy relative. (B) Schematic representation of MPL protein and location of the two variants identified in the two fetuses. (C) Neuropathological findings in fetus II-2. Brain macroscopical aspect of Fetus II-2 (*top-left*). Histological examination with hematoxylin and eosin stain showing ependymal "pseudo-rosettes" (*top-right*, *original magnification* ×*40*), hemosiderin pigments (*bottom left*, *original magnification* ×*100*) and siderocalcium salts (*bottom-right*, *original magnification* ×*1000*)

pulmonary atresia with ventricular septal defect and an overriding aorta. Brain examination unveiled a hemorrhagic lesion in the right temporal lobe and showed necrotic and hemorrhagic reorganizations in the left temporal lobe. The second pregnancy was uneventful with the delivery of a healthy child (III-2). A recurrent IUFD occurred at 28 WG during the third pregnancy (III-3). In contrast with observations made for the fetus proband III-1, fetal autopsy revealed skeletal malformations with only 11 pairs of ribs, a rough draft of cervical ribs and a synostosis of radius and ulna. Neuropathological analysis of fetus III-3 of the brain showed a weight within the lower limit of normal, and there was a hematoma filling the left sylvian valley and extending to the frontal pole (Figure 4C). On section, the lateral ventricles were dilated while the left sylvian valley was ripped by a hematoma. The brain parenchyma was punctuated by diffuse and multiple petechiae in the cerebral hemispheres, the brainstem and cerebellum. Histological analysis revealed recent (no sign of resorption) and multifocal bleedings, very often centered by a capillary. There was no associated infectious sign.

4 | DISCUSSION

In a cohort of 194 consecutive ICH fetuses interrupted or stillbirth and referred for molecular screening of *COL4A1* and *COL4A2* we identified a pathogenic variant in one of these two genes in 19% of cases. These variants occurred de novo in 70% of these fetuses. In

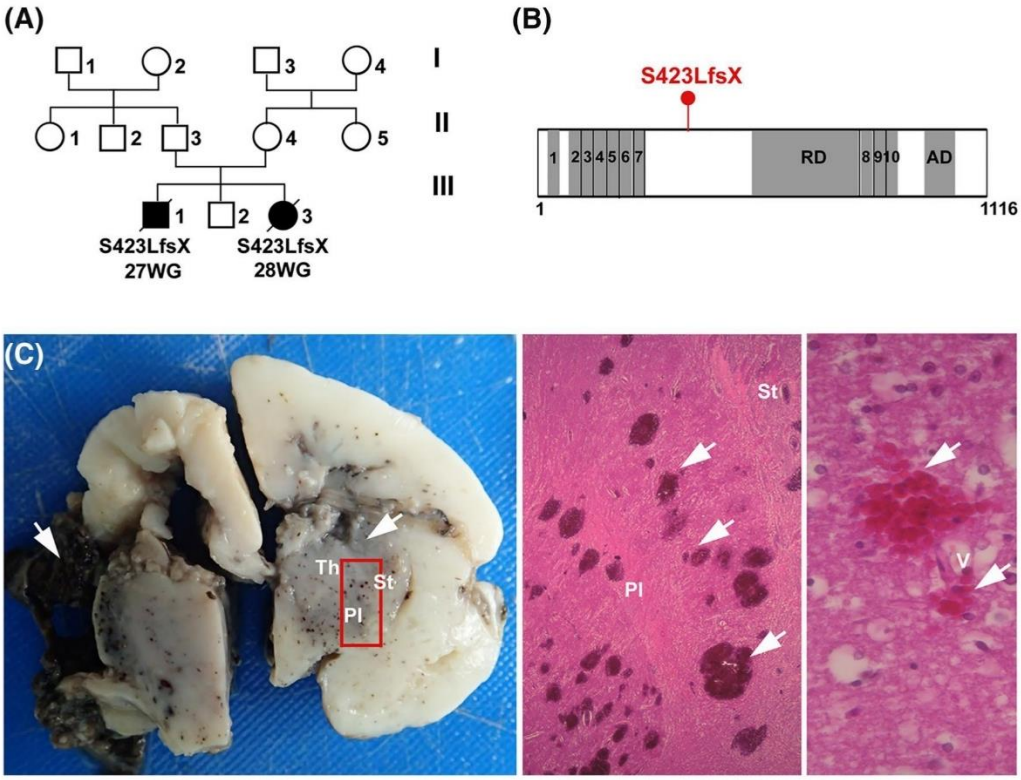


FIGURE 4 Family F2 with *MECOM* pathogenic variant. (A) Family tree showing dominant inheritance of *MECOM* variant. Square = male; circle = female; diagonal black line = deceased individual; black filled symbol = clinically affected individual; empty symbol = clinically healthy relative. (B) Schematic representation of EVI-I protein and location of the variant identified in the two fetuses. (C) Neuropathological findings in fetus III-3. Brain sectioning showed a disruption of the left sylvian valley parenchyma by a hematoma (*left arrow*). The brain parenchyma also appeared punctuated with multiple hemorrhagic foci (*right arrow*) and disseminated in the white and gray structures (*Th, thalamus; St, striatum; Pl, pallidum*). On a microscopic level (hematoxylin and eosin stain), these petechiae consisted of focal bleeding (*middle image*), mostly in contact with the vascular walls (*right image*)

addition, three pathogenic variants of *MPL* and *MECOM* genes were shown to cause recurrent ICH in two distinct families.

4.1 | *COL4A1* and *COL4A2* pathogenic variants are the most frequent genetic cause of fetal ICH

Several studies have reported the implication of *COL4A1/COL4A2* variants in fetal ICH with a broad phenotype spectrum including, in addition to ICH, porencephaly, cortical malformations and schizencephaly.[5,22-32] However, most of these studies were case reports of one or very few fetuses. To our knowledge, this is the sole large series of ICH fetuses screened for *COL4A1/COL4A2* genes. These data establish that (i) these two genes account for a large proportion of fetal ICH (ii) 70% of these variants occur de novo. These findings are important for genetic counseling.

The 70% de novo rate observed in this fetal cohort is much higher than in series including post-natal cases, either children or adult patients. Indeed, in their excellent review of their own cases and literature data Meuwissen et al.[29] reported a de novo rate of 40%. In small series limited to fetuses, the reported de novo rate is also very high, around 80%.[5,32] It was also close to 80% in the subgroup of patients with prenatal manifestations in Itai et al recent series.[33] In addition to this high de novo rate, we showed that the distribution of variants along the *COL4A1* gene is significantly different between fetuses and adult patients. Altogether, these data strongly suggest that the very early onset and severity of fetal ICH is at least partly explained by differences in mutations location. Interestingly, this observation is in adequation with mouse models in which severity is also correlated with mutations location.[34,35]

4.2 | Thrombocytopenia associated with megakaryopoiesis gene variants, albeit rare, is a cause of fetal ICH

The two etiologies evoked nowadays in fetal ICH are FNAIT, when associated with a thrombocytopenia, and *COL4A1/COL4A2*, when occurring in the absence of thrombocytopenia. IPD gene screening

TABLE 1 Main clinical and pathological features of ICH fetuses in F1 and F2 families

	Family F1		Family F2	
Case	II-2	II-3	III-1	III-3
Sex	Male	Female	Male	Female
Gestation at diagnosis (weeks)	28	14	27	28
Ultrasound findings	Bilateral ventriculomegaly	Fetal death Lymphedema	Fetal death Heart failure Hydrops fetalis	Fetal death
Outcome	TOP	IUFD	IUFD	IUFD
Fetal Brain examination	Bilateral hydrocephaly Diffuse ventricular and subarachnoid hemorrhage	Advanced maceration Liquefied cerebral parenchyma	Hemorrhagic lesion on the right temporal area	Intraparenchymal and subarachnoid lesions encompassing the temporal area
	Right parenchymal necrosis Parietal cortical polymicrogyria	Bulky blood clot indicating a recent hemorrhage	Necrotic and hemorrhagic reorganizations in the left temporal area	Diffuse intraparenchymal petechiae
Other autopsy anomalies	No sign of external or visceral malformation	No sign of external or visceral malformation	Pulmonary atresia Ventricular septal defect Overriding aorta	11 pairs of ribs Radio-ulnar synostosis

Abbreviations: IUFD, intrauterine fetal death; TOP, termination of pregnancy.

is most often not performed unless there is a strong familial history of hemorrhage. In the rare situations where hemostasis genes are tested, this screening is most often limited to clotting factors. Indeed, this was the case for the two families with *MPL* and *MECOM* variants reported herein. In addition, information on platelet count is often missing. For some fetuses, blood count might be missing either because fetal blood sampling has not been performed or the sample had been coagulated. The identification herein of causative variants in two megakaryopoiesis genes in two distinct families strongly suggests that IPD genes should however be screened in the absence of FNAIT and *COL4A1/COL4A2* variants, even though they are rare. Indeed, these data are of major importance for genetic counseling.

The reason why only two out of the 101 probands tested for IPD gene variants were found to be mutated in this cohort is unclear. We hypothesize that, among IPD genes, only gene variants leading to a profound thrombocytopenia would lead to fetal ICH. IPD gene variants responsible of mild to moderate thrombocytopenia, might not lead to fetal ICH or might need an additional triggering event. Several additional IPD genes were however reported to be potentially responsible for severe thrombocytopenia, including THPO, PRKACG, WAS, RBM8A, GATA1 or HOXA11 genes.[13,15,36] Pathogenic variants of genes leading to platelet dysfunction might also contribute to fetal ICH but have not been detected in this cohort owing to their rarity. None of these genes was mutated in our cohort suggesting that their prevalence is low in fetal ICH.

4.3 | Clinical features of *MPL* and *MECOM* associated platelet disorders

Biallelic *MPL* pathogenic variants are known to lead to classical congenital amegakaryocytic thrombocytopenia (CAMT-MPL).[37,38] CAMT-MPL is a rare autosomal recessive bone marrow failure syndrome leading to pancytopenia due to a depletion of hematopoietic progenitors. CAMT-MPL diagnosis is most often made at birth or later on within the context of an isolated thrombocytopenia and a pancytopenia later on. However, clinically affected patients showed wide phenotypic variability as illustrated in a recent series of 56 patients with CAMT-MPL.[39] In 13 of these 56 cases, no thrombocytopenia was detected in the neonatal period. Fetal ICH was observed in only seven cases (12.5%), including four cases diagnosed retrospectively. Despite its lower frequency, ICH was the most severe manifestation. The p.R102P variant found in our study is the most frequent one in CAMT-MPL. It is encountered in most cases as a compound heterozygous variant; it has been reported only once at homozygous state in one boy with an antenatal ICH.[40] This variant isoform is not addressed at the cell membrane.[21] The p.R537W variant is far less frequent and has been reported by two previous studies neither in an ICH context. One report was a series of patients from the Choctaw tribe affected by a thrombocytopenia suggesting being a founder variant.[41] Regarding heterozygous relatives, Germeshausen et al.,[39] observed that platelet counts of heterozygous parents and siblings of probands were within the normal range

except for one parent (p.R102P) with a mild thrombocytopenia. However, the possibility of a mild thrombocytopenia is contradictory with another study suggesting that germline *MPL* R102P heterozygous variant could lead to a thrombocytosis.[42] Altogether these data suggest that a clarification is needed about the effect of heterozygous *MPL* variants in carriers, including the father of proband F1, for proper clinical management and genetic counseling.

The second megakaryopoiesis gene *(MECOM)* plays an important role in the formation of hematopoietic stem cells and cell self-renewal. Somatic rearrangements of *MECOM* lead to an overexpression and have been reported in leukemia with a poor prognosis.[43] Germline missense or nonsense variants located in the zinc finger domain region[44] lead to bone marrow failure and extra-hematological events through a haploinsufficiency mechanism.[45] A recent study of 12 cases with *MECOM* variants revealed a wide clinical spectrum ranging from isolated radioulnar synostosis to severe bone marrow failure without skeletal abnormalities.[44] Authors observed an incomplete penetrance with the presence of unaffected relatives in 4 out of 11 family members. Other abnormalities have been described such as clinodactyly, hearing loss and cardiac malformations.[46] These findings were consistent with fetal examination features found in family F2 where we observed a clinical heterogeneity with one fetus having cardiac defects (a pulmonary atresia with ventricular septal defect) and no skeletal abnormality, and the other one a radio-ulnar synostosis with no cardiac defect.

Our study has some limitations. We focused our study on genes responsible of thrombocytopenia or platelet dysfunction based on physiopathological hypotheses. This study emphasizes the need for further pan genomic analyses in the next future.

To conclude, variants of *COL4A1* and *COL4A2* are the leading genetic cause of fetal ICH but genes associated with severe thrombocytopenia could be responsible of massive fetal ICH. Those genes should be investigated in ICH fetuses even in the absence of a platelet data count since the detection of such variants is crucial for genetic counseling and prenatal diagnosis. Altogether, these data emphasize the strong genetic heterogeneity of this condition and the need to perform advanced genetic analysis to broaden our knowledge on the biological mechanisms causing fetal ICH. We also emphasize the need to conduct pangenomic screening on large multicenter cohorts to better understand genotype-phenotype correlations and potentially serve as the basis for the development of targeted treatments.

ACKNOWLEDGMENTS

The authors thank families for their participation in this study. We also acknowledge clinicians, geneticists and pathologists who referred fetuses enrolled in this study especially Pr Tania Attie-Bittach, Dr Jelena Martinovic, Dr Paul Maurice, Dr Marie-Laure Moutard and Dr Stéphanie Valence. Finally, we are greatly indebted to Dr Florence Marchelli for her excellent figure editing. This work was supported by the National Institutes of Health (R01NS096173 grant).

CONFLICT OF INTEREST

Nothing to report.

ETHICS STATEMENT

This study was approved by the INSERM Ethics Review Committee (IRB00003888) and an informed and signed consent was obtained for parents and fetus probands included in the study.

DATA AVAILABILITY STATEMENT

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

ORCID

Thibault Coste https://orcid.org/0000-0001-5610-7411
Louise Devisme https://orcid.org/0000-0002-7412-7798
Jean-Marie Jouannic https://orcid.org/0000-0002-7890-3790

REFERENCES

1. Monteagudo A. Intracranial hemorrhage. *Am J Obstet Gynecol*. 2020;22:B34-B37.
2. Ghi T, Simonazzi G, Perolo A, et al. Outcome of antenatally diagnosed intracranial hemorrhage: case series and review of the literature. *Ultrasound Obstet Gynecol*. 2003;22:121-130.
3. Adiego B, Martínez-Ten P, Bermejo C, Estévez M, Recio Rodriguez M, Illescas T. Fetal intracranial hemorrhage. Prenatal diagnosis and postnatal outcomes. *J Matern Fetal Neonatal Med*. 2019;32:21-30.
4. Bussel JB, Vander Haar EL, Berkowitz RL. New developments in fetal and neonatal alloimmune thrombocytopenia. *Am J Obstet Gynecol*. 2021;225:120-127.
5. Meuwissen MEC, de Vries LS, Verbeek HA, et al. Sporadic COL4A1 mutations with extensive prenatal porencephaly resembling hydranencephaly. *Neurology*. 2011;76:844-846.
6. Cavaliere AF, Turrini I, Pallottini M, et al. Genetic profiling of idiopathic antenatal intracranial haemorrhage: what we know? *Genes*. 2021;12:573.
7. Hausman-Kedem M, Malinger G, Modai S, et al. Monogenic causes of apparently idiopathic perinatal intracranial hemorrhage. *Ann Neurol*. 2021;89:813-822.
8. Almazni I, Stapley R, Morgan NV. Inherited thrombocytopenia: update on genes and genetic variants which may be associated with bleeding. *Front Cardiovasc Med*. 2019;6:80.
9. Bastida JM, Lozano ML, Benito R, et al. Introducing high-throughput sequencing into mainstream genetic diagnosis practice in inherited platelet disorders. *Haematologica*. 2018;103:148-162.
10. Downes K, Megy K, Duarte D, et al. Diagnostic high-throughput sequencing of 2396 patients with bleeding, thrombotic, and platelet disorders. *Blood*. 2019;134:2082-2091.
11. Heremans J, Freson K. High-throughput sequencing for diagnosing platelet disorders: lessons learned from exploring the causes of bleeding disorders. *Int J Lab Hematol*. 2018;40:89-96.
12. Johnson B, Doak R, Allsup D, et al. A comprehensive targeted next-generation sequencing panel for genetic diagnosis of patients with suspected inherited thrombocytopenia. *Res Pract Thromb Haemost*. 2018;2:640-652.
13. Noris P, Pecci A. Hereditary thrombocytopenias: a growing list of disorders. *Hematology Am Soc Hematol Educ Program*. 2017;2017:385-399.
14. Nurden AT, Nurden P. High-throughput sequencing for rapid diagnosis of inherited platelet disorders: a case for a European consensus. *Haematologica*. 2018;103:6-8.

15. Nurden AT, Nurden P. Inherited thrombocytopenias: history, advances and perspectives. *Haematologica*. 2020;105:2004-2019.
16. Seo A, Gulsuner S, Pierce S, et al. Inherited thrombocytopenia associated with mutation of UDP-galactose-4-epimerase (GALE). *Hum Mol Genet*. 2019;28:133-142.
17. Shim YJ. Genetic classification and confirmation of inherited platelet disorders: current status in Korea. *Clin Exp Pediatr*. 2020;63:79-87.
18. Backenroth D, Homsy J, Murillo LR, et al. CANOES: detecting rare copy number variants from whole exome sequencing data. *Nucleic Acids Res*. 2014;42:e97.
19. Richards S, Aziz N, Bale 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*. 2015;17:405-424.
20. Hitchcock IS, Kaushansky K. Thrombopoietin from beginning to end. *Br J Haematol*. 2014;165:259-268.
21. Geddis AE. Congenital amegakaryocytic thrombocytopenia. *Pediatr Blood Cancer*. 2011;57:199-203.
22. de Vries LS, Koopman C, Groenendaal F, et al. COL4A1 mutation in two preterm siblings with antenatal onset of parenchymal hemorrhage. *Ann Neurol*. 2009;65:12-18.
23. Lichtenbelt KD, Pistorius LR, De Tollenaer SM, Mancini GM, De Vries LS. Prenatal genetic confirmation of a COL4A1 mutation presenting with sonographic fetal intracranial hemorrhage. *Ultrasound Obstet Gynecol*. 2012;39:726-727.
24. Colin E, Sentilhes L, Sarfati A, et al. Fetal intracerebral hemorrhage and cataract: think COL4A1. *J Perinatol*. 2014;34:75-77.
25. Garel C, Rosenblatt J, Moutard ML, et al. Fetal intracerebral hemorrhage and COL4A1 mutation: promise and uncertainty: Letter to the Editor. *Ultrasound Obstet Gynecol*. 2013;41:228-230.
26. Kutuk MS, Balta B, Kodera H, et al. Is there relation between COL4A1/A2 mutations and antenatally detected fetal intraventricular hemorrhage? *Childs Nerv Syst*. 2014;30:419-424.
27. Takenouchi T, Ohyagi M, Torii C, Kosaki R, Takahashi T, Kosaki K. Porencephaly in a fetus and HANAC in her father: variable expression of COL4A1 mutation. *Am J Med Genet*. 2015;167:156-158.
28. Matsumoto T, Miyakoshi K, Fukutake M, Ochiai D, Minegishi K, Tanaka M. Intracranial sonographic features demonstrating in utero development of hemorrhagic brain damage leading to schizencephaly-associated COL4A1 mutation. *J Med Ultrasonics*. 2015;42:445-446.
29. Meuwissen MEC, Halley DJJ, Smit LS, et al. The expanding phenotype of COL4A1 and COL4A2 mutations: clinical data on 13 newly identified families and a review of the literature. *Genet Med*. 2015;17:843-853.
30. Khalid R, Krishnan P, Andres K, et al. COL4A1 and fetal vascular origins of schizencephaly. *Neurology*. 2018;90:232-234.
31. Sato Y, Shibasaki J, Aida N, et al. Novel COL4A1 mutation in a fetus with early prenatal onset of schizencephaly. *Hum Genome Var*. 2018;5:4.
32. Maurice P, Guilbaud L, Garel J, et al. Prevalence of COL4A1 and COL4A2 mutations in severe fetal multifocal hemorrhagic and/or ischemic cerebral lesions. *Ultrasound Obstet Gynecol*. 2021;57:783-789.
33. Itai T, Miyatake S, Taguri M, et al. Prenatal clinical manifestations in individuals with COL4A1/2 variants. *J Med Genet*. 2021;58:505-513.
34. Kuo DS, Labelle-Dumais C, Mao M, et al. Allelic heterogeneity contributes to variability in ocular dysgenesis, myopathy and brain malformations caused by Col4a1 and Col4a2 mutations. *Hum Mol Genet*. 2014;23:1709-1722.
35. Jeanne M, Jorgensen J, Gould DB. Molecular and genetic analyses of collagen type IV mutant mouse models of spontaneous intracerebral hemorrhage identify mechanisms for stroke prevention. *Circulation*. 2015;131:1555-1565.
36. Bouchghoul H, Quelin C, Loget P, et al. Fetal cerebral hemorrhage due to X-linked GATA1 gene mutation. *Prenat Diagn*. 2018;38:772-778.
37. Ballmaier M, Germeshausen M. Congenital amegakaryocytic thrombocytopenia: clinical presentation, diagnosis, and treatment. *Semin Thromb Hemost*. 2011;37:673-681.
38. Germeshausen M, Ballmaier M. Congenital amegakaryocytic thrombocytopenia – not a single disease. *Best Pract Res Clin Haematol*. 2021;34:101286.
39. Germeshausen M, Ballmaier M. CAMT-MPL: congenital amegakaryocytic thrombocytopenia caused by MPL mutations - heterogeneity of a monogenic disorder - comprehensive analysis of 56 patients. *Haematologica*. 2021;106:2439-2448.
40. Eshuis-Peters E, Versluys AB, Stokman MF, van der Crabben SN, Nij Bijvank SWA, van Wezel-Meijler G. Congenital amegakaryocytic thrombocytopenia type II presenting with multiple central nervous system anomalies. *Neuropediatrics*. 2016;47:128-131.
41. Newman LA, Luter MA, Davis DB, Abdul-Rahman OA, Johnson JM, Megason GC. Congenital amegakaryocytic thrombocytopenia: a case series indicating 2 founder variants in the Mississippi band of Choctaw Indians. *J Pediatr Hematol Oncol*. 2017;39:573-575.
42. Bellanné-Chantelot C, Mosca M, Marty C, Favier R, Vainchenker W, Plo I. Identification of MPL R102P mutation in hereditary thrombocytosis. *Front Endocrinol*. 2017;8:235.
43. Yuan X, Wang X, Bi K, Jiang G. The role of EVI-1 in normal hematopoiesis and myeloid malignancies (Review). *Int J Oncol*. 2015;47:2028-2036.
44. Germeshausen M, Ancliff P, Estrada J, et al. MECOM-associated syndrome: a heterogeneous inherited bone marrow failure syndrome with amegakaryocytic thrombocytopenia. *Blood Adv*. 2018;2:586-596.
45. Maki K, Yamagata T, Asai T, et al. Dysplastic definitive hematopoiesis in AML1/EVI1 knock-in embryos. *Blood*. 2005;106:2147-2155.
46. Niihori T, Ouchi-Uchiyama M, Sasahara Y, et al. Mutations in MECOM, encoding oncoprotein EVI1, cause radioulnar synostosis with amegakaryocytic thrombocytopenia. *Am J Hum Genet*. 2015;97:848-854.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Coste T, Vincent-Delorme C, Stichelbout M, et al. *COL4A1/COL4A2* and inherited platelet disorder gene variants in fetuses showing intracranial hemorrhage. *Prenat Diagn*. 2022;1-10. https://doi.org/10.1002/pd.6113

Supplemental Data

COL4A1/COL4A2 and inherited platelet disorder gene variants in fetuses showing intracranial hemorrhage

Thibault Coste[1,2], Catherine Vincent-Delorme[3], Morgane Stichelbout[4], Louise Devisme[4], Antoinette Gelot[5], Igor Deryabin[5], Fanny Pelluard[6], Chaker Aloui[2], Anne Louise Leutenegger[2], Jean-Marie Jouannic[7], Delphine Héron[8], Douglas B Gould[9] and Elisabeth Tournier-Lasserve[1,2]

1 AP-HP, Service de génétique moléculaire Neurovasculaire, Hôpital Saint-Louis, France

2 Université de Paris, INSERM UMR-1141 Neurodiderot, Paris F-75019, France

3 CHU Lille, Service de génétique clinique Guy Fontaine, Lille, France

4 CHU Lille, Institut de pathologie, Lille, France

5 APHP, Service de fœtopathologie, Hôpital Trousseau, France

6 Univ.Bordeaux, INSERM, BaRITOn, U1053, F-33000 BORDEAUX, France

7 APHP Sorbonne Université, Service de médecine fœtale, Hôpital Trousseau, Paris, France

8 AP-HP, Service de génétique clinique, Hôpital de la Pitié-Salpêtrière, Paris, France

9 Department of Ophthalmology, University of California San Francisco, San Francisco, USA

Corresponding author: Pr Elisabeth Tournier-Lasserve

INSERM UMR-1141 Neurodiderot, Hôpital R. Debré, Paris F-75019, France

Email: tournier-lasserve@univ-paris-diderot.fr

Supplemental Data

Contents:

Supplemental Data

Supplemental Material and Methods.

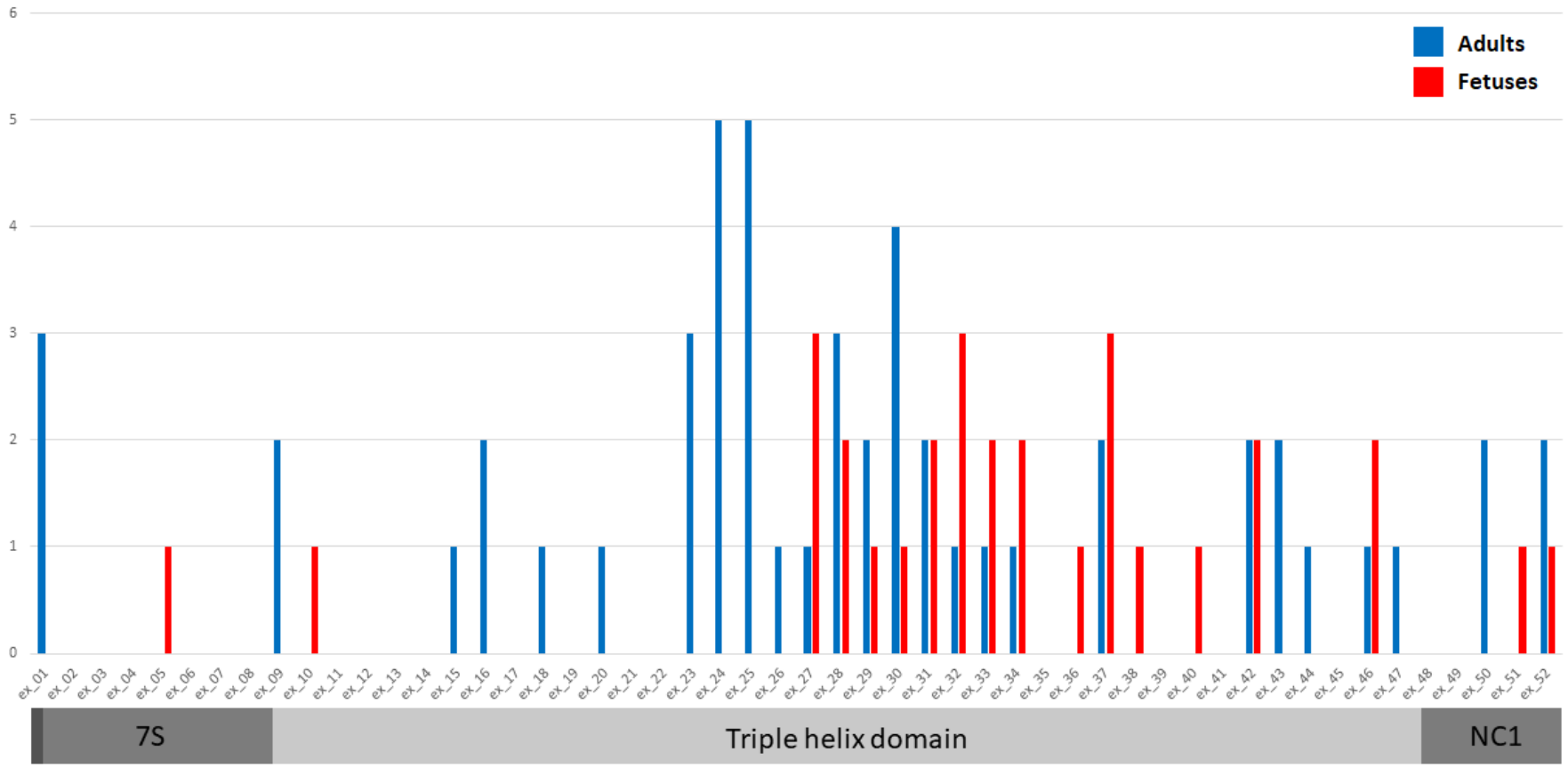
In order to compare the distribution of *COL4A1* variants identified in fetuses and in adult patients we included all adult patients referred to our laboratory for molecular screening of vascular leukoencephalopathy genes from 2010 to 2020. *COL4A1* and *COL4A2* were screened either by Sanger sequencing (from 2010 to 2015) or by NGS targeted sequencing. For targeted NGS, library preparation was made with the SureSelect QXT® capture kit (Agilent technologies) and sequencing was performed on a MiSeq® sequencer (Illumina). We used SeqNext software 4.4 and 5.0 versions (JSI Medical Systems) to analyze the sequencing data. All retained variants were confirmed by Sanger sequencing using the BigDye™ Terminator v3.1 cycle sequencer system (Applied Biosystems) and an ABI 3130 DNA sequencer with in house primers (available upon request).

The mean age at inclusion of adult patients with a pathogenic variant in *COL4A1* was 49.8 years +/- 16.3 years. Among the 51 adult patients with a pathogenic variant, the variant was inherited in four cases and was de novo in three cases. In the remaining cases (86 % of the probands) we did not have access to parental DNA (as usual in adult probands).

Statistical analysis was performed by comparing the mean of the cDNA position of variants in the triple helical domain of *COL4A1* in the two groups (Fetuses vs Adults). We used a one sided t-test with the hypothesis that the mean of cDNA position variants in fetuses was higher than in the group of adults. We used R software (version 4.0.5) and the t.test() function, including Welch correction for unequal variance. A p-value <0.05 was considered statistically significant.

Supplemental Data

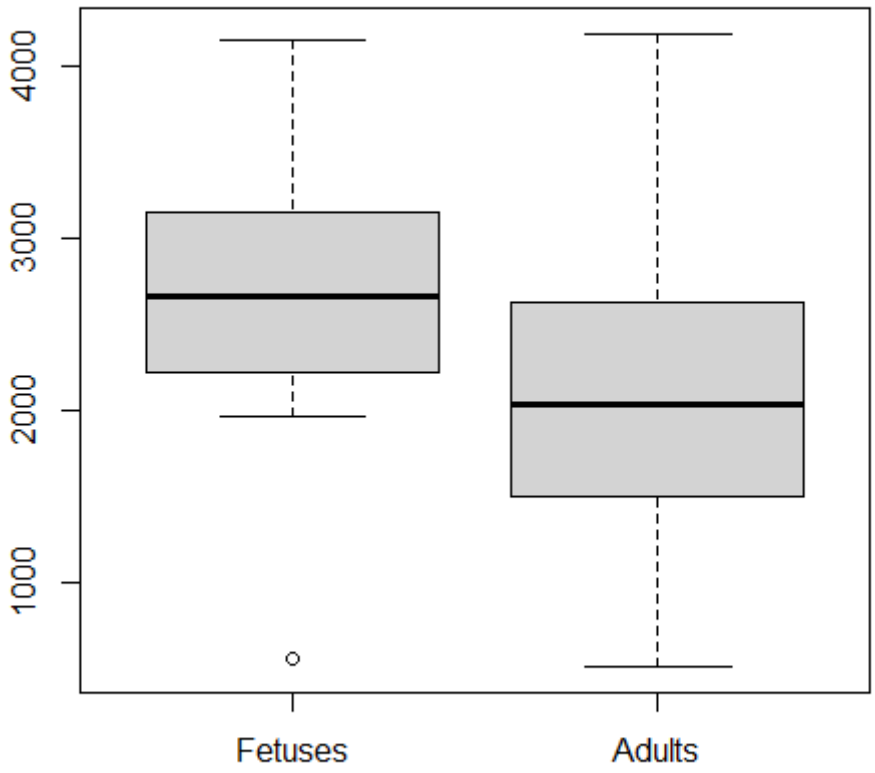
Figure S1. Location of *COL4A1* pathogenic variants identified in ICH fetuses vs adult probands



COL4A1 variants in the 30 ICH fetuses are shown in red. The 51 pathogenic variants detected in adult patients are shown in blue. The Y axis relates to the number of variant carriers in a given exon. The 52 exons are displayed on the X axis. Triple helical domain extends from exon 9 to 48 of the *COL4A1* gene.

Supplemental Data

Figure S2. Boxplot of the location of *COL4A1* variants in ICH fetuses vs adult probands



Y-axis represents cDNA positions of variants in the triple helix domain of *COL4A1*. The one sided t-test revealed a statistically significant increase of the *COL4A1* variant positions in fetuses (P = 0.0028).

Supplemental Data

Table S1. Manually curated list of genes involved in IPD that have been used to filter exome variants†

Dominant inheritance		Recessive inheritance				Recessive or dominant inheritance	
Gene	**Mode of inheritance**	**Gene**	**Mode of inheritance**	**Gene**	**Mode of inheritance**	**Gene**	**Mode of inheritance**
ACTN1	AD	*ABCG5*	AR	*LYST*	AR	*GP1BA*	AR; AD
ACTB	AD	*ABCG8*	AR	*MPIG6B*	AR	*GP1BB*	AR; AD
ANKRD26	AD	*ANO6*	AR	*MPL*	AR	*ITGA2B*	AR; AD
CDC42	AD	*AP3B1*	AR	*NBEAL2*	AR	*ITGB3*	AR; AD
CYCS	AD	*ARPC1B*	AR	*PLA2G4A*	AR	*TBXA2R*	AR; AD
DIAPH1	AD	*BLOC1S3*	AR	*P2RY12*	AR		
ETV6	AD	*BLOC1S6*	AR	*PRKACG*	AR		
FLNA	XLD	*DTNBP1*	AR	*PTPRJ*	AR		
GFI1B	AD	*FERMT3*	AR	*RASGRP2*	AR		
HOXA11	AD	*FLI1*	AR	*RBM8A*	AR		
IKZF5	AD	*FYB1*	AR	*RNU4ATAC*	AR		
MASTL	AD	*G6B*	AR	*SLC35A1*	AR		
MECOM	AD	*GALE*	AR	*STXBP2*	AR		
MYH9	AD	*GATA1*	XLR	*TBXAS1*	AR		
NBEA	AD	*GNE*	AR	*THPO*	AR		
PLAU	AD	*GP6*	AR	*VPS33B*	AR		
RUNX1	AD	*GP9*	AR	*VIPAS39*	AR		
SLFN14	AD	*HPS1*	AR	*VWF*	AR		
SRC	AD	*HPS3*	AR	*WAS*	XLR		
STIM1	AD	*HPS4*	AR				
TPM4	AD	*HPS5*	AR				
TRPM7	AD	*HPS6*	AR				
TUBB1	AD	*KDSR*	AR				

† List established on the basis of literature data (Ref. 7-16)

AD: autosomal dominant; AR: autosomal recessive; XLD: X-linked dominant; XLR: X-linked recessive

Supplemental Data

Table S2. Likely pathogenic and pathogenic variants identified in IPD genes according to ACMG criteria

ID	Gene	Refseq Transcript	cDNA position	Protein change	Status	Inheritance	ACMG attributed class
F9	*HPS3*	NM_032383.4	c.728_729ins	p.Ser244PhefsTer4	HTZ	Unknown	4
F8	*HPS5*	NM_181507.1	c.2A>T	p.Lys788Ter	HTZ	Maternal	5
F3	*ITGA2B*	NM_000419.4	c.748C>T	p.Gln250Ter	HTZ	Maternal	5
F4	*ITGA2B*	NM_000419.4	c.1612G>A	p.Glu538Lys	HTZ	Unknown	4
F5	*ITGB3*	NM_000212.2	c.1156C>T	p.Arg386Cys	HTZ	Maternal	4
F6	*LYST*	NM_000081.3	c.6122A>G	p.Asp2041Gly	HTZ	Unknown	4
F10	*LYST*	NM_000081.3	c.10235G>A	p.Arg3412His	HTZ	Unknown	4
F2	*MECOM*	NM_001105077.3	c.1268del	p.Ser423LeufsTer6	HTZ	Unknown	5
F1	*MPL*	NM_005373.2	c.305G>C	p.Arg102Pro	HTZ	Paternal	5
F1	*MPL*	NM_005373.2	c.1609C>T	p.Arg537Trp	HTZ	Maternal	4
F7	*VWF*	NM_000552.4:	c.7294G>A	p.Val2432Ile	HTZ	Unknown	4

HTZ: heterozygous; ACMG (American College of Medical Genetics) classification of variants: 4 = likely pathogenic, 5 = pathogenic

DISCUSSION

Nous avons identifié des variants pathogènes dans *COL4A1* et *COL4A2* chez 18% des fœtus référés. La majorité de ces variants était située dans le domaine de la triple hélice des protéines COL4A1 et COL4A2, et dans 83% des cas ces variations concernaient le gène *COL4A1*. Pour 32 des 36 fœtus mutés nous disposions des ADN parentaux. Chez 70% des fœtus porteurs de variants pathogènes de *COL4A1*/*COL4A2* il s'agissait de variants apparus de novo, avec une différence selon le gène. Pour *COL4A1*, les variations pathogènes sont dans 80% des cas *de novo* et dans 20% des cas héritées d'un parent ; pour *COL4A2*, 15% seulement des variants pathogènes étaient *de novo* (1/6). Nous avons également montré une localisation différente des variations pathogènes identifiées chez le fœtus par rapport à celles identifiées chez l'adulte atteint d'une leucoencéphalopathie vasculaire. Ces variations chez le fœtus sont localisées préférentiellement au niveau C-terminal du domaine de la triple hélice de COL4A1 (p=0.0028).

Le pourcentage élevé de variations *de novo* identifiées chez le fœtus est cohérent avec la sévérité du phénotype par rapport au spectre clinico-radiologique des collagénopathies COL4A1/A2 observé chez l'adulte. Ce taux élevé, particulièrement pour *COL4A1*, permet de rassurer un couple lors du conseil génétique quant au risque qu'un des deux parents soit porteur et du risque de récurrence pour une future grossesse. A noter que nous ne pouvons toutefois exclure le fait que certains parents puissent être porteurs de variants en mosaïque que nous n'aurions pas détectés dans les leucocytes du sang circulant. Par ailleurs, la localisation différente des variations pathogènes chez le fœtus et l'adulte pourrait expliquer une partie de la sévérité du phénotype fœtal. Il a été décrit dans la littérature que les variants en C-terminal provoqueraient plus d'accumulation intracellulaire de COL4A1 (Jeanne *et al.*, 2015), et une plus grande cytotoxicité cellulaire, et in fine une plus grande fragilité des vaisseaux. Néanmoins, certains fœtus ont aussi des variants dans la partie N-terminale de la triple hélice et certains adultes ont des variants dans la partie C-terminale, ce qui suggère que d'autres mécanismes doivent exister pour expliquer la sévérité et la variabilité du spectre phénotypique.

Concernant l'analyse des 64 gènes responsable d'IPD, chez 101 fœtus index, onze variants pathogènes ou probablement pathogènes ont été identifiés chez 10 fœtus.

Deux fœtus index étaient issus de familles à cas multiples et présentaient des variants dans des gènes impliqués dans la mégacaryopoïèse (production de mégacaryocytes et de plaquettes). Le premier fœtus présentait des variants biallélliques dans le gène *MPL*, qui est un gène majeur de la mégacaryopoïèse codant pour le récepteur à la thrombopoïétine et qui est associé, en cas de

variations pathogènes bialléliques, à une thrombopénie congénitale sévère appelée CAMT (*Congenital AMegakaryocytic Thrombocytopenia*) (Germeshausen et Ballmaier, 2020). Dans cette famille deux fœtus étaient atteints et mutés. Le premier fœtus présentait une ventriculomégalie détectée à l'échographie de 27 SA, et a été interrompu à 28 SA. L'analyse neuropathologique avait mis en évidence des lésions de nécrose dans la région pariétale droite, des suffusions au contact de la zone germinale épaissie, la présence de « pseudo-rosettes » épendymaires et la présence de pigments d'hémosidérine et de sels sidéro-calcaires. Une numération sanguine avait été faite et retrouvait une anémie et une thrombocytopénie sévère (Hb = 3 g/L, Plaquettes = 15 G/L) ; une FNAIT avait été suspectée mais la recherche d'anticorps maternels anti-plaquettes et le crossmatch étaient négatifs. Un deuxième fœtus de cette famille avait présenté une augmentation de la clarté nucale à 11SA puis une mort fœtale à 14SA. L'analyse de l'encéphale retrouva un caillot sanguin dans le parenchyme. La numération plaquettaire n'avait pas pu être faite. Ces deux fœtus étaient porteurs de deux variants hétérozygotes en position p.(R102P) et p.(R537W), hérités pour chacun d'un des 2 parents. La variation p.(R102P) est la plus fréquente dans le CAMT-MPL et à l'état homozygote est responsable d'une perte de fonction du gène. Néanmoins il a été rapporté des patients hétérozygotes pour ce variant p.(R102P) qui présentaient une thrombocytose (taux de plaquettes supérieur à 450 G/L) et donc un risque de thrombose (Bellanné-Chantelot *et al.*, 2017). Cette information est essentielle pour le conseil génétique car la variation p.R102P peut avoir un impact différent sur la production de plaquettes : thrombocytose ou thrombocytopénie selon la présence à l'état hétérozygote ou homozygote, respectivement.

La seconde famille présentait un variant disruptif hétérozygote dans le gène *MECOM*, entraînant une haploinsuffisance. *MECOM* est aussi un gène majeur de la mégacaryopoïèse qui code pour un facteur de transcription. En cas de variation pathogène, il est responsable d'une thrombocytopénie congénitale avec synostose radio-ulnaire (ou CTRUS-MECOM pour *Congenital amegakaryocytic Thrombocytopenia with Radio-Ulnar Synostosis* associée au gène *MECOM*) dont les signes cliniques sont très variables. Le premier fœtus de cette famille présentait une insuffisance cardiaque et une anasarque à 22 SA puis une MFIU à 27 SA. Il n'a pas été fait de numération sanguine. A l'examen fœtopathologique, il présentait une atrésie pulmonaire, une aorte dextroposée et au niveau cérébral un hématome sylvien gauche, de multiples foyers hémorragiques dans les structures blanches et grises correspondant à des saignements focaux au contact des parois vasculaires. Le second fœtus a bénéficié d'une surveillance rapprochée à 12, 18 et 22 SA sans particularité suivie de la détection d'une MFIU à 28 SA. Également, aucune numération sanguine n'avait été faite. Il présentait à l'examen un

hématome dans le lobe temporal et une synostose radio-ulnaire. Cette dernière est un élément particulier souvent décrit dans le CTRUS lié au gène *MECOM* et qui n'avait pas été remarquée à l'examen initial mais seulement après rediscussion du dossier suite à l'identification du variant causal (rétrophénotypage). Pour cette famille les parents n'ont pas été génotypés et sont a priori asymptomatiques. Ce statut asymptomatique pourrait s'expliquer par une pénétrance incomplète, ce caractère incomplet étant connu (Germeshausen *et al.*, 2018), ou par un mosaïcisme chez le parent porteur que nous n'avons hélas pas pu déterminer.

Pour 8 fœtus index nous avons également identifié un variant hétérozygote pathogène ou probablement pathogène (classe 4 et 5 selon l'ACMG) dans un gène responsable d'IPD (gènes *HPS3, HPS5, LYST, ITGA2B, ITGB3* et *VWF*). Pour ces gènes, la pathologie associée est de type récessive à l'exception des gènes *ITGA2B* et *VWF* pour lesquels il existe des formes dominantes. Pour ces deux gènes, les variants pathogènes n'entrainent pas une thrombopénie mais une anomalie fonctionnelle des plaquettes (thrombopathie). Pour *ITGA2B*, la maladie associée la plus courante est la thrombasthénie de Glanzmann, maladie autosomique récessive ; cependant des formes dominantes avec macrotrhombocytopénie congénitale sont rapportées mais sans saignement sévère de type HIC (Nurden *et al.*, 2011). Le gène *VWF* est associée à la maladie de Willebrand, maladie hémorragique héréditaire rare caractérisée par un défaut d'adhésion plaquettaire et un défaut de coagulation (par défaut de transport du facteur VIII). Il en existe trois sous-types principaux : déficit partiel (type 1) ou total (type 3), et anomalies qualitatives/fonctionnelles (type 2). L'hérédité est le plus souvent autosomique dominante, mais la maladie est transmise de manière autosomique récessive dans le cas du type 3 et dans certains cas du type 2. Pour le fœtus porteur du variant hétérozygote probablement pathogène p.(V432I) du gène *VWF* nous n'avons pas pu conclure car il s'agissait d'un fœtus issu d'un couple sans antécédents notables avec un bilan d'hémostase normal et qui avait par ailleurs trois autres enfants en bonne santé.

Pour ces 8 fœtus nous avons cherché la présence d'un CNV sur l'autre allèle du gène en question à l'aide de l'outil CANOES et cette recherche s'est avérée négative. Cependant l'existence d'un CNV ne peut être complètement exclue en raison des limites analytiques bien connues de la recherche de CNV sur les données de WES. De même, on ne peut exclure la présence d'un variant intronique sur l'autre allèle qui pourrait avoir un effet sur l'épissage. Des analyses complémentaires par un séquençage génome entier sont envisagées chez ces fœtus pour conclure définitivement sur l'implication des gènes identifiés.

Au final, ces résultats mettent en évidence l'importance des variants pathogènes de *COL4A1*, *COL4A2* et d'autres gènes liés à la mégacaryopoïèse dont *MPL* et *MECOM* dans la survenue d'HIC fœtales. Ces résultats ont des implications très importantes pour le conseil génétique d'une part et la compréhension des mécanismes sous-jacents aux hémorragies cérébrales chez le fœtus d'autre part. Les variations dans les gènes responsables de pathologies plaquettaires héréditaires sont peu rapportées chez le fœtus et donc non investiguées de façon systématique devant un tableau d'HIC même en présence d'une thrombopénie quand la numération plaquettaire est disponible.

Dans cet article nous avons présenté une analyse par liste de gènes candidats qui a l'avantage de restreindre l'étude à un nombre de variants candidats plus gérable qu'une analyse sans à priori, mais qui ne permet pas d'incriminer des gènes dans des voies biologiques non suspectées. L'analyse sans à priori sur la fonction des gènes est présentée dans les travaux suivants.

B- Le séquençage exome entier de fœtus référés pour HIC met en évidence une extrême hétérogénéité génétique

INTRODUCTION

Pour la totalité des fœtus de notre cohorte, les parents sont décrits comme asymptomatiques, et bien qu'il y ait quelques cas de formes familiales avec récidive d'une HIC sur plusieurs grossesses, la majorité des cas d'HIC fœtales est de type sporadique. Ce constat est plutôt en faveur de l'existence de variants pathogènes apparus *de novo* et /ou de variants bialléliques. Ainsi, après une analyse a priori d'une liste de gènes candidats, nous avons analysé les variants hétérozygotes apparus de *novo* chez 35 trios, les variants bialléliques homozygotes détectés chez nos 113 fœtus index et les variants bialléliques hétérozygotes composites pour les trios et familles à cas multiples pour lesquels les ADN des parents étaient disponibles.

Nous avons tout d'abord réalisé une analyse en composante principale pour mieux caractériser l'origine géographique des fœtus et une analyse de la consanguinité sur les données de WES. Nous avons ensuite utilisé différents filtres de sélection des variants candidats (basés sur la qualité, la fréquence allélique dans les bases de données de contrôles, les données de prédiction in silico) et de priorisation des gènes (nombre d'individus porteurs de variants disruptifs dans gnomAD). Enfin, pour les variants hétérozygotes de l'ensemble de la cohorte, nous avons réalisé un test de charge en variants rares et prédits délétères à l'échelle du gène afin d'identifier des gènes causaux ou candidats associés à une pénétrance incomplète, ou des gènes pour lesquels les variants hétérozygotes seraient apparus *de novo* mais dont le génotype des parents n'est pas connu (ADN génomiques des parents non disponibles pour 2/3 des fœtus de la cohorte).

Les résultats de ces analyses sont présentés dans l'article ci-après qui est en cours de soumission.

Exome sequencing of fetuses referred for intracerebral hemorrhage reveals an extreme genetic heterogeneity

Authors:

Thibault Coste[1,2], Chaker Aloui[1], Jelena Martinovic[3], Tania Attie-Bitach[4,5], Florence Petit[6], Delphine Héron[7], Alexandre G. de Brevern[8], Ragousandirane Radjasandirane[8], Michaelle Coprechot[2], Hélène Morel[1,2], Rachel Petermann[9], Anne-Louise Leutenegger[1], Elisabeth Tournier-Lasserve[1,2]

Affiliations:

1 Université Paris Cité, Inserm, NeuroDiderot, F-75019 Paris, France
2 Service de génétique neurovasculaire, AP-HP, Hôpital Saint-Louis, F-75010 Paris, France
3 Unité de Fœtopathologie, AP-HP, Hôpital Antoine Béclère, Groupe Hospitalo-Universitaire Paris Saclay, Clamart, France.
4 Service de Médecine Génomique des Maladies Rares, Hôpital Universitaire Necker-Enfants Malades, Paris, France
5 INSERM UMR 1163, Institut Imagine, Université Paris Cité, Paris, France
6 CHU Lille, Clinique de Génétique Guy Fontaine, Lille, France
7 UF de Génétique Médicale et CRMR « Déficience intellectuelle », Département de Génétique médicale, Groupe Hospitalier Pitié-Salpêtrière, AP-HP Sorbonne Université, Paris, France
8 INSERM UMR_S 1134, Univ. Paris Diderot, Sorbonne Paris Cité, INTS, GR-Ex
9 Platelets Immunology Laboratory, National Reference Centre for Perinatal Hemobiology (CNRHP), Paris, France

Correspondence:

Thibault Coste
Service de génétique neurovasculaire
Hôpital Saint-Louis
1 Avenue Claude Vellefaux
75010 Paris, France
Email : thibault.coste@aphp.fr

Key words:

Fetal intracerebral hemorrhage, exome sequencing, genetic heterogeneity

Number of figures: 5

Number of tables: 3

ABSTRACT

Introduction: Fetal intracerebral hemorrhage (FICH) is a serious event that occurs in approximately 1/10,000 pregnancies and is associated, for the severe forms, with a very poor prognosis. Two main causes of FICH are well known: the fetal and neonatal alloimmune thrombocytopenia (FNAIT) and the pathogenic variations in genes encoding type IV collagen (*COL4A1* and *COL4A2*), an essential component of the basement membrane. *COL4A1/A2* pathogenic variations are detected in 20% of fetuses referred for FICH after exclusion of acquired causes. No cause is identified in 80% of fetuses referred for molecular screening rendering genetic counseling or prenatal diagnosis impossible. Our main objective was to identify pathogenic variants in novel genes associated with FICH.

Methods: A whole exome sequencing was performed on 113 unrelated fetuses including 35 trios and 3 families with two affected fetuses. Fetuses were referred to our neurovascular genetic department for a suspected *COL4A1/A2* collagenopathy in a context of FICH, ventriculomegaly, porencephaly and/or schizencephaly. Fetuses with pathogenic variations in *COLA4A1/A2* were excluded. In addition, a systematic review of medical charts was performed to exclude fetuses with known risk factors leading to ICH. Several data analysis strategies with rigorous filtering criteria were conducted including an analysis of *de novo* and biallelic variants as well as a collapsing gene-based burden test approach.

Results: Several disease-causing genes have been identified, affecting diverse biological pathways. Some genes were involved in platelet metabolism and hemostasis (*MPL, MECOM* and *PROC* genes) We also identified a gene involved in endothelial cell adhesion (*ESAM*) and four genes involved in mitochondrial metabolism (*ATP5PO*, *PDHA1*, *COQ2* and *NAXD)*. Pathogenic variants in these four latter genes lead to phenotypes mimicking phenotypes encountered in FICH. In addition, our study identified several strong candidate genes requiring further analysis, including *TIE1, ITGA2B* and *DCAF5*.

Discussion: In addition to basement membrane genes such as *COL4A1/COL4A2*, genes involved in either platelet metabolism, clotting or endothelial cell adhesion might be mutated in an ICH fetus. However, FICH is characterized by an extreme genetic heterogeneity with very few fetuses mutated for each incriminated gene. Altogether, these data strongly support the

need for exome-wide analyses of large cohorts of affected fetuses in order to characterize the full genomic architecture of FICH.

INTRODUCTION

Intracranial hemorrhage (ICH) is frequent in premature infants. It also occurs in utero with an estimated incidence of 1/10,000 pregnancies (Monteagudo *et al.,* 2020). Most of the time the diagnosis is established during a routine ultrasound. However, the diagnosis is sometimes made by postmortem examination in a context of in utero fetal death (IUFD).

Fetal intracerebral hemorrhage (FICH) most commonly originates in the germinal matrix and may extend to the ventricles. When extensive, germinal matrix hemorrhage can lead to venous infarction due to the occlusion of the periventricular collector veins, subsequent tissue destruction and development of a porencephaly. FICH can be subdivided in four grades with extensive hemorrhages classified as grade III (hemorrhage with enlarged ventricles (ventriculomegaly) or grade IV (hemorrhage within the cerebral parenchyma) (Ghi *et al.,* 2003). Both grades are associated with a very poor perinatal prognosis and in most cases result in severe motor disability and developmental delay (Adiego *et al.,* 2019).

Several non-genetic factors contribute to FICH, such as maternal trauma, maternal drug exposure, infectious diseases, fetal or neonatal alloimmune thrombocytopenia (FNAIT) or defects in fetal hemostasis (Bussel *et al.,* 2021). After excluding these possible causes, a condition arising from a genetic disorder should be taken into consideration.

The most frequent Mendelian cause of FICH is associated with pathogenic variations of *COL4A1* and *COL4A2*. Collagen type IV alpha 1 (COL4A1) and 2 (COL4A2) are extracellular matrix proteins that together constitute a major component of nearly all basement membranes. COL4A1/COL4A2 pathogenic variations lead to severe and/or multifocal lesions that can be supratentorial or infratentorial with multifocal ischemic and hemorrhagic lesions of different ages, associated with schizencephaly or porencephaly. Other signs have also been reported such as leukomalacia, brain calcifications, corpus callosum anomalies and cerebellar hypoplasia (Yoneda et al., 2012; Meuwissen et al., 2015; Itai et al., 2021). We previously showed that *COL4A1/A2* pathogenic variations occurred *de novo* in 70% of fetal cases (Coste *et al.,* 2022 a). Other known genetic etiologies for FICH include pathogenic variants of inherited platelet disorder genes, hemostatic genes (responsible of factors V or VII deficiencies for example) or pathogenic variations in *GATA1* gene, which is an extremely rare X-linked condition (Bouchghoul *et al.,* 2018; Cavalière *et al.,* 2021). However, the underlying cause cannot be

identified in 80% of fetuses referred for molecular screening of a suspected ICH, precluding any appropriate therapy, genetic counseling for future pregnancies, and family counseling.
The purpose of this study was to identify pathogenic variants in novel genes and decipher the genetic causes of FICH using whole exome sequencing data from 113 unrelated fetuses referred for *COL4A1/COL4A2* gene screening which turned out to be negative.

MATERIAL AND METHODS

Material and ethics statement

A total of 113 unrelated fetuses including 35 trios (fetus and his/her unaffected parents) were part of this study. All fetuses deceased as a result of intrauterine fetal death (IUFD) or were interrupted by medical termination of pregnancy (TOP). They were initially referred for *COL4A1/COL4A2* screening in a context of fetal intracerebral hemorrhage, ventriculomegaly, porencephaly and/or schizencephaly, and they turned out to be negative for this *COL4A1/COL4A2* screening.
A systematic review of medical charts was performed to exclude fetuses with acquired identifiable causes or known risk factors leading to ICH. A written informed consent for genetic investigation and research was provided by parents in accordance with the declaration of Helsinki and the French law. In addition, this study was approved by the Inserm Ethics Committee (INSERM IRB00003888).
Genomic DNA was isolated from umbilical cord blood or from post-mortem fetal tissue, and from peripheral blood leukocytes for parents and relatives when available. Prior to whole exome sequencing, a pathogenic variant in *COL4A1* and *COL4A2* genes had been excluded in each fetus by targeted high throughput sequencing using Agilent SureSelect QXT® capture kit (conditions available upon request).

Whole exome sequencing (WES)

Exon capture was performed at the IntegraGen platform (Evry, France) for fetus probands and their relatives using the SureSelect Human All Exon V5-UTR (Agilent technologies) or the Twist Human Core Exome Enrichment System (Twist Bioscience), followed by a 100 base paired-end sequencing using an Illumina NovaSeq platform. Sequencing data were annotated by a previously described pipeline (Coste *et al.*, 2022, b).

Ancestry prediction and consanguinity analysis

Each individual's ancestry was predicted on WES data using *Peddy* that implements a support vector machine trained on the 1000 Genomes Project phase 3 individuals (Pedersen *et al.*, 2017).

The consanguinity analysis was performed using the 1000 Genomes allele frequencies corresponding to each individual's inferred ancestry from *Peddy*.

The inbreeding coefficient F was estimated for each fetus from her/his exome data using maximum likelihood with a hidden Markov model approach (Leutenegger *et al.*, 2003 & 2011). A likelihood ratio test was performed to evaluate whether the estimated F was significantly different from zero. For inbred fetuses, segments with at least 5 consecutive positions with posterior HBD probabilities > 50% were considered as homozygosity by descent (HBD) segments. All these analyses were performed with the R package *Fantasio* (https://github.com/genostats/Fantasio).

Filtering and prioritization strategies for de novo, biallelic and X-linked variants

Analysis of de novo variants (DNVs) was performed on the 35 trios and is summarized in figure 1. Briefly, DNVs were filtered based on the following criteria: (1) inferred loss of function (LoF) variants (stop gain, stop loss, frameshift and canonical splice site) and missense variants predicted deleterious using PolyPhen-2; (2) minimum read depth of 10X in the proband and both parents; (3) proband alternative allele ratio >25%; (4) global minor allele frequencies (MAF) < 1/1 000 in public databases (1000 Genomes phase 3 and gnomAD v2.1).

Genes with qualifying DNVs were then investigated in the 78 solo fetuses, looking for additional qualifying variants in the whole fetus cohort. Finally, an enrichment analysis at gene level of recurrent qualifying variants was performed using Fisher's exact test (implemented in R) and gnomAD v3.0 as a control cohort (N=71,702).

Recessive variants flowchart analysis is summarized in figure 2. Identification of homozygous qualifying variants was performed on the 113 proband fetuses, whereas analysis of compound heterozygous variants was conducted on the 35 trios, 2 quartets and 1 family of 6 members with 2 affected fetuses. Criteria to filter qualifying variants included: (1) inferred LoF variants and missense variants predicted deleterious using PolyPhen-2; (2) minimum read depth of 10X and

alternative allele ratio >25%; (3) global MAF < 1/200 in public databases (1000 Genomes phase 3 and gnomAD v2.1).

Regarding X-linked variants, 58 male fetuses were included in this analysis (Figure 3). Same filter criteria were used apart from the MAF (< 1/10 000).

Due to the small number of genes with recurrent qualifying variants in at least 2 fetuses, additional criteria to prioritize genes without recurrence at gene level were used. In a context of a severe condition, recommendation is that for a novel candidate gene with proposed recessive inheritance, no unaffected individuals should be observed with homozygosity for LoF variants in the candidate gene (MacArthur *et al.*, 2014). Therefore, we only prioritized genes that did not show any homozygous LoF variant in the control database gnomAD v.2.1.

Gene-based collapsing test

Aggregated data from the 71,702 individuals without severe Mendelian condition included in the gnomAD v3.0 were used as control data. Combined VCF files for fetuses and controls were reannotated with VEP104. Rare and high-confidence likely pathogenic variants were selected using several filters. We limited our study to inferred LoF and missense variants predicted damaging by PolyPhen-2. Rare variants were defined as having a MAF <1/10 000 or being absent in gnomAD exome samples (version 2.0.1). For each tested gene, we checked that exons were covered at more than 10x in at least 90% of samples. Genome coverage summary statistics were downloaded from https://gnomad.broadinstitute.org/downloads. In addition, to filter out common local variants, we used a cutoff of 1% either across the case samples or across the control samples.

Finally, for each gene, the number of individuals carrying at least one qualifying variant was compared between cases and controls using Fisher's exact test implemented in R (https://www.r-project.org). Bonferroni-corrected exome-wide statistical significance threshold was set to 2.5 x10-6.

Evaluation of candidate genes/variants

Pathogenicity of qualifying variants was established according to the American College of Medical Genetics and Genomics (ACMG) guidelines (Richards *et al.*, 2015). Selected genes carrying qualifying variants were systematically evaluated using OMIM (Online Mendelian Inheritance in Man; www.omim.org), possible involvement in FICH based on mouse knockout

models and functional information available in literature. GeneMatcher was used (Sobreira *et al.*, 2015) for genes with limited phenotypic information.

Genes were finally classified into three categories:
1/ *FICH causing gene:* well-known gene function or animal model associated with ICH or a similar phenotype.
2/ *Candidate gene:* gene with evidence suggesting a possible involvement in ICH (including expression, pathway, nature of identified variants) but for which causality is not currently established; or currently unknown gene function and not reported in the literature, therefore not excluded from having a potential role in FICH.
3/ *Likely non-causal gene:* well-known gene function that is not directly associated with the disease phenotype.

Data visualization and Sanger confirmation

In silico visualization of BAM files was performed for all filtered variants to remove false-positive calls. Variants were further confirmed by standard PCR amplification and Sanger sequencing using the BigDye™ Terminator v3.1 (Applied Biosystems) on an ABI 3130 DNA sequencer. The *de novo* nature of variants was confirmed by microsatellite analysis of both fetus and parental DNA using the AmpFISTR Profiler PCR Amplification Kit (Applied Biosystems).

Data previously reported

Data of six index fetuses from the 113 fetus cohort have already been reported using different strategies than analysis of *de novo* and recessive variants. Reported genes were *MPL*, *MECOM, ESAM* and *PDHA1*. Molecular data of these previously reported fetuses are summarized in Table S1.

RESULTS

Cohort characteristics

We enrolled 113 *COL4A1/COL4A2* negative ICH fetuses probands in this study, including 35 trios, 2 quartets and a family of six relatives with two affected fetuses and two unaffected children. All parents were self-reported as clinically asymptomatic (Figure S1). The sex ratio (58 male fetuses / 55 female fetuses; ratio = 1.05) did not show any bias. Fetuses were referred for a suspicion of ICH with features including intraparenchymal lesions, leukomalacia, schizencephaly, porencephaly, polymicrogyria, corpus callosum anomalies and/or lesions of the posterior fossa. In most cases, these anomalies were detected during the third trimester ultrasound examination, with a mean age at diagnosis of 28.3 +/- 5.2 weeks of gestation. As a result of the poor prognosis of those brain anomalies, 90% of pregnancies resulted in a termination of pregnancy and 10% in an IUFD. Based on Peddy analysis of geographical background, 72% of fetuses were of European origin, 12% were of African origin, 10% were of multiracial origin, 5% were classified as Admixed American, and 1% were of South Asian origin (Figure S2). Clinical charts indicated that six fetuses were consanguineous. Consanguinity was confirmed in four of them using exome data with F coefficients ranging from 0.0068 to 0.125 (Table S2) and homozygous by descent (HBD) segments number ranging from 1 to 26. The four inbred fetuses did not share any HBD segment with one another.

De novo variants identified in the 35 trios

Analysis of the 35 trios allowed the identification of 41 *de novo* qualifying variants (DNVs) occurring in 40 genes among 23 fetuses (Figure 1, Table S3).

There were five loss of function (LoF) variants (in *DCAF5, DSG4, PABPC1L, PLXNA2* and *TRA2A* genes) and 16 missense variants predicted to be damaging (Table 1). Among LoF variants, *DCAF5* was the only gene with a PLI score > 0.9 and a low number of heterozygous LoF variants in control databases (n=5 in gnomAD v.2.1, whereas n=23-166 for the 4 other genes). The *DCAF5* variant (p.D478As*17) was located in the last exon and thus predicted to generate a truncated protein. No Mendelian phenotype has been described so far for this gene; however an ICH phenotype was observed in a mouse model inactivated for DDB1, a direct partner of DCAF5 (Cang *et al.,* 2006). Among the 16 genes with predicted damaging missense DNVs only *F10* and *ITGA2B* genes, according to literature data, were associated with a human hemorrhagic phenotype but which was known to be autosomal recessive.

In the absence of recurrence of DNVs at gene level for *DCAF5* and for genes with the 16 predicted damaging missense DNVs, we investigated the entire cohort of fetuses (113 fetuses) to find out additional rare and potentially damaging variants in that list of genes. We identified a recurrence in nine genes, including three genes (*EXD3*, *TECTA* and *ZFYVE28*) for which there was a statistical enrichment (p-values < 0.05) compared to the gnomAD control database. *TECTA* encodes one of the major noncollagenous component of the tectorial membrane in the inner ear and pathogenic variations are responsible of an autosomal dominant non-syndromic hearing impairment and a recessive form of sensorineural pre-lingual non-syndromic deafness. In the literature, one patient with non-syndromic hearing loss carried a pathogenic variant at the same position (p.I1996N) that the DNV found in fetus F17 (p.I1996V) with no information about his neurological condition (Yasukawa *et al.*, 2019). Concerning *ZFYVE28* and *EXD3*, no human phenotype has been described, and Zfyve28 knocked out mice did not show any obvious abnormalities in major organs at macroscopic examination (Zambrano *et al.*, 2018).

Among the 41 DNVs, two distinct rare DNVs (p.I766T and p.P434S) in *ITGA2B* were identified in two unrelated fetuses (F06 and F106). This observation is a very unlikely event (Poisson-adjusted p-value = 0.047). *ITGA2B* p.(P434S) variant is absent of control databases and predicted damaging by several in silico softwares. The p.(I766T) variant had a frequency of $4x10^{-6}$ in gnomAD and was predicted benign. However, fetus F06 carrying this variant carried also a LoF variant in *ITGA2B* (p.Q250*) inherited from her mother. An allele-specific PCR confirmed that both variants were located on separate alleles (Figure S3-A). We queried multiple impact prediction algorithms and conducted a structural analysis of these two missense variations. The majority (>2/3) of predictive approaches for the P434S variant consider it as deleterious, and structural analysis showed a loss of all interactions leading to potential destabilization in the Beta-Propeller domain (Figure S3-B). Regarding the I766T variant, all predictive approaches consider it as benign and there is only a loss of one (weak) interaction, so its structural impact is very unlikely.

Inherited biallelic variants in trio and solo fetuses

A total of 73 homozygous qualifying variants located in 66 distinct genes were detected in 37 / 113 fetuses. In addition, we identified 17 fetuses from 35 trios and 3 families carrying qualifying compound heterozygous variants in 38 genes (Figure 2, Tables S4 and S5). To prioritize the strongest candidate genes, we used three additional criteria based on mutational

recurrency, absence of these qualifying variants at homozygous state in the gnomAD database, and absence of homozygous LoF variants in gnomAD for the studied gene.

Six genes showed biallelic (homozygous and compound heterozygous) qualifying variants (LoF and predicted damaging missense) in at least two unrelated fetuses (Table 2): *EPPK1, GOLGA6L2, HLA-DRB1, MUC19, MUC4* and *SCRIB*. *SCRIB* was the only gene without any homozygous LoF variant in gnomAD suggesting that it is essential for development. One fetus (F92) carried a homozygous missense variant in *SCRIB* (p.R221C) and another one (F112) carried compound heterozygous missense variants (p.H21Y/p.G1526R). The remaining five genes displayed several homozygous LoF variants in gnomAD and were not prioritized.
In the remaining 98 genes with either homozygous or compound heterozygous variants that did not show recurrence of qualifying variants at gene level, our prioritization strategy based on the absence of homozygous LoF in gnomAD resulted in selection of 16 genes with homozygous variants (*ATP5PO, ASIC4, CCDC102A, COQ2, CRELD1, ESAM, FGD5, MAFA, MYG1, NAXD, NOS3, NUAK2, PLEC, RET, RNPEP* and *ZNF251*) as well as seven genes with compound heterozygous variants (*ADRA2A, LRP8, MPL, PROC, TNC, ZNF687* and *ZNF865*) (Figure 2, Table 3).

As regard to homozygous LoF variants, four fetuses carried variants in *ATP5PO* (F58), *COQ2* (F32), *ESAM* (F92), *MAFA* (F92) and *RNPEP* (F35). Fetus F92 with variants in *ESAM* (p.L12*) and *MAFA* (p.A93Pfs*27) also carried one homozygous missense variant in *SCRIB*. It was a consanguineous fetus and homozygous qualifying variants were located in HBD segments. *ESAM* was concluded as the causing gene in F92 based on a collaborative effort reporting other patients with antenatal ICH and homozygous LoF variants in *ESAM* (Lecca et *al*., 2023). *MAFA* and *SCRIB* variants were deemed to be not associated with FICH in F92. One fetus (F32) carried a LoF variant in *COQ2* located in the last exon and expected to lead to a truncated protein. *COQ2* pathogenic variants are responsible of coenzyme Q10 deficiency, an autosomal recessive disorder for which some patients present an encephalopathy with signs resembling Leigh syndrome and stroke like episodes. Since fetus F32 presented a moderate ventriculomegaly, a lissencephaly and a leukomalacia, we considered the *COQ2* variant as probably disease causing. Fetus F35 carried a homozygous variant in *RNPEP*, which is a ubiquitous gene involved in inflammation and tumor development (Piesse *et al.*, 2022) and no human phenotype was reported. However, *RNPEP* remains a possible candidate gene for FICH based on the homozygous LoF nature of this variant. Finally, one consanguineous fetus (fetus F58) carried a homozygous variant in a splice region of *ATP5PO* (NM_001697.3:c.87+3A>G).

This variant has already been reported in two index children with brain MRI abnormalities consistent with Leigh syndrome. Studies on their muscle cell messenger RNA showed a majority transcript with exon 2 skipping and traces of the wild-type transcript. A complementary study of respiratory chain enzymes in fibroblasts showed a reduction in complex V ATPase activity in both families (Ganapathi *et al.*, 2022).

As regard to homozygous missense variants, one consanguineous fetus (F58) carried five homozygous variants in *ASIC4, CCDC102A, MYG1, RET* and *ZNF251*. Based on the literature, phenotypes associated with these genes were not consistent with FICH. A second consanguineous fetus (F01) carried two homozygous missense variants in *NAXD* and *NOS3*. This fetus belongs to a family in which a previous fetus had also been terminated in a context of suspected ICH. Unfortunately, the DNA quality of this fetus did not allow WES, but Sanger analysis revealed that he also carried the *NAXD* variant at homozygous state while he was heterozygous for the NOS3 variant. In addition, the *NAXD* qualifying variant (p.D232N) is predicted deleterious by almost all *in silico* softwares and is in the same domain harboring previously reported pathogenic variations (Van Bergen *et al.*, 219). *NAXD* pathogenic variations are associated with an early-onset progressive encephalopathy with brain edema and/or leukoencephalopathy. Clinically, F01 fetus presented microcephaly, hydranencephaly and a pontocerebellar hypoplasia.

Finally, regarding compound heterozygous variants, two affected fetuses from a family of six relatives carried bi-allelic variants in *MPL*. Each parent and the two unaffected siblings carried one heterozygous variant. These two variants (p.R102P and p.R537W) have been reported in a context of congenital amegakaryocytic thrombocytopenia (CAMT-MPL), a recessive condition for which FICH is as one of its manifestations. In addition, we identified biallelic variants in *PROC* in the fetus F12. One variant was a heterozygous LoF variant (p.G215Dfs*40), while the second variant was a missense variant (p.A388E) deemed to be likely pathogenic since a variant at the same position was reported pathogenic in heterozygous patients with low protein C levels (Downes *et al.*, 2019; Megy *et al.*, 2021). *MPL* and *PROC* genes were therefore considered as FICH causing genes. Available data for the other genes, as regard to gene function and the associated phenotype described in the literature, did not suggest any involvement in FICH.

X-linked variants in male trio and solo fetuses

A total of 42 hemizygous qualifying variants located in 41 distinct genes were identified in 28 / 58 male fetuses (Figure 3, Table S6). Two fetuses carried hemizygous missense variants in *ARHGAP4* gene (Table 2). This gene was not retained as a candidate gene since many hemizygous LoF variants (n= 11) are reported in gnomAD.

Six genes were considered as possible candidate genes despite the absence of recurrence at gene level, namely *BCOR, BTK, MOSPD2, NKAP, PGK1* and *TSPAN7* (Table 3). All qualifying variants were missense variants. *BCOR* was the sole gene for which an abnormal brain phenotype was reported in literature data. *BCOR* mutated patients present a syndromic microphthalmia (MIM# 300166) but without evidence of ICH. They carry almost exclusively truncated and frameshift variants (Ng *et al.*, 2004; Hilton *et al.*, 2009). Those observations were not consistent with the phenotype of fetus F77 who had triventricular dilatation with post-hemorrhagic remodeling in the germinal matrix.

Gene-Based collapsing test using gnomAD as a control database

WES analysis of large sets of unrelated probands using gene-based collapsing tests is a useful strategy in the absence of large families informative for linkage or to identify genes involved in diseases with incomplete penetrance. This approach compares proportions of individuals carrying rare protein-altering variants within a given gene between patients and appropriate controls. Gene-based collapsing test comparing ICH fetuses with the gnomAD v3 control database identified one gene (*PDHA1*) showing genome-wide significance (p= 2.11 x10-7) (Table S7). Three female fetuses were found to have a heterozygous LoF variant in *PDHA1*. The second ranked gene was *TIE1*, for which six fetuses carried a qualifying variant (Table S8) with a p-value near the genome wide significant threshold (p= 9.64 $x10^{-6}$). These six *TIE1* variants were all missense variants, not localized in a specific domain, and three of them were inherited from an asymptomatic parent. Other genes in the top 10 identified by the gene-based burden test included *ZNF775, ASH1L, PLB1, YPEL1, PPRC1, PDE8B, SULF1* and *TBL1X* whose p-values ranged from 1.06 $x10^{-4}$ to 6.62 $x10^{-4}$. Available data concerning gene functions, animal models and associated phenotypes in humans described in the literature for these genes were not suggestive of a potential involvement in FICH (Table S9).

DISCUSSION

We conducted a WES analysis using stringent criteria in 113 unrelated fetuses referred for a suspected ICH and in whom *COL4A1/COL4A2* gene screening was negative. Based on the in silico predicted pathogenicity of the variants and on information available in the literature (gene function, cellular and animal models, human data), we identified causing variants in seven genes (*ATP5PO, COQ2, ESAM, MPL, MECOM, PDHA1* and *PROC).* In addition, we identified rare and damaging variants in several strong candidate genes including *ITGA2B, DCAF5, NAXD* and *TIE1*, and several possible candidate genes as *EXD3, RNPEP* and *ZFYVE28*. For almost all causative and candidate genes, there was a very low recurrence, illustrating the extreme genetic heterogeneity of FICH (Figure 4). Our results also show that pathogenic variants in genes involved in mitochondrial metabolism can lead to phenotypes mimicking those observed in fetuses carrying pathogenic *COL4A1/COL4A2* variants, reinforcing the indication for whole exome sequencing in fetuses with ICH.

FICH associated with genes involved in coagulation, thrombophilia and platelet synthesis

Severe coagulation factor deficiencies, such as factor 5 and factor 7 deficiencies have previously been reported in fetuses with ICH. These deficiencies are very rare, with incidences ranging from 1/500,000 to 1/1,000,000 (Cavaliere *et al.,* 2021), which probably explains why we have not found any fetus carrying pathogenic variants in these coagulation genes.

However, we identified one fetus with biallelic pathogenic variants in *PROC*. Heterozygous protein C deficiency is relatively common, affecting as many as 1/200 individuals, and is a risk factor for venous thrombosis. Homozygous protein C deficiency is extremely rare, and classically presents with purpura fulminans, ophthalmic complications and thrombotic stroke in the newborn period (Lim *et al.,* 2016). However, homozygous deficiency could affect fetuses, as illustrated by Stutterd *et al.* who reported a fetus of 35 WG showing overall intracranial findings that suggest bilateral ischemic brain injury, with a resultant severe porencephaly (Stutterd *et al.,* 2014). While cerebral thrombosis is a well-known manifestation of congenital protein C deficiency, intracranial hemorrhage in the absence of cerebral thrombosis is a less recognized complication (Alsultan *et al.,* 2016). Some authors, hypothesized that a severe protein C deficiency could lead to ICH as a secondary consequence of microthrombi in blood capillaries (Martin et al., 2021). In addition, due to the relatively uncommon incidence, this severe deficit is underestimated in fetus (Fong *et al.,* 2010; Martin *et al.,* 2021).

In a family with two affected fetuses (F03) we identified biallelic *MPL* pathogenic variants. Those variants are known to be involved in classical congenital amegakaryocytic thrombocytopenia, a rare autosomal recessive bone marrow failure syndrome characterized by a depletion of hematopoietic progenitors described (Germeshausen *et al.,* 2020). Analysis of all genes associated with an inherited platelet disorder allowed the identification of a family carrying a pathogenic variant in *MECOM.* This gene was also associated with severe thrombocytopenia.

Considering these data, we concluded that genes involved inherited platelet disorders, as well as genes involved in coagulation and thrombophilia should be systemically investigated.

FICH associated with pathogenic variations in tight junction molecules

One homozygous LoF variant in *ESAM* was identified in a consanguineous fetus. This gene encodes an endothelial cell-selective adhesion molecule. No phenotype in human was reported when we identified this variant. GeneMatcher allowed the identification of eight unrelated families carrying homozygous LoF variants in *ESAM*; all these patients showed prenatal ICH (Lecca *et al.,* AJHG 2023). ESAM is a tight junction protein expressed exclusively in the vascular endothelium (Nasdala *et al.,* 2002). The novel phenotype associated with biallelic *ESAM* variants overlapped with other known conditions associated with pathogenic variation of genes encoding distinct tight junction molecules (*JAM2, JAM3*, and *OCLN*). Among those, homozygous LoF variants in *JAM3* have been identified in patients with multifocal antenatal intraparenchymal hemorrhage (MIM# 613730), and infancy death (Xu *et al.,* 2022). Phenotypes associated with *JAM2* and *OCLN* pathogenic variations, are respectively basal ganglia calcifications (MIM# 618824) and pseudo-TORCH syndrome (MIM# 251290). However, in these two conditions no cerebral hemorrhage is reported.

Potential candidate genes associated with FICH

We identified *TIE1* as a candidate gene in six fetuses carrying rare and predicted damaging heterozygous missense variants, based on the gene-based burden test. Those variants are not clustered in a specific domain and three of them were inherited from an asymptomatic parent. Heterozygous missense variants in *TIE1* located at other positions have been associated with an autosomal dominant lymphatic malformation syndrome (MIM# 619401). *TIE1* is highly expressed in vascular endothelial cells and plays a crucial role in angiogenesis, vascular permeability and homoeostasis via angiopoietin signaling (Eklund *et al.,* 2017; La Porta *et al.,*

2018; Korkhonen *et al.*, 2016; Griffin *et al.*, 2021). Altogether, these data support the potential involvement of *TIE1* in ICH. Additional in vitro experiments are however required to sort out the causality of these variants.

Two *de novo* variants in *ITGA2B* were identified in our FICH cohort. Biallelic pathogenic variants of *ITGA2B* lead to Glanzmann thrombasthenia, a rare autosomal recessive disease (Nurden *et al.*, 2011). A dominant form of the disease (Platelet-Type Bleeding Disorder 16; MIM number 187800) has been reported in some patients and is characterized by a macrothrombocytopenia and mild bleeding tendency without ICH (Kunishima *et al.*, 2011). Altered *ITGA2B* function could explain ICH, but we were unable to formally establish a causal link between *ITGA2B* and ICH. Indeed, in one case, the fetus was a compound heterozygous with a LoF variant and a *de novo* variant that is unlikely to affect ITGA2B function; the second fetus carried only one heterozygous predicted damaging *de novo* variant. There was no fetal tissue available for functional analysis. Our current working hypothesis is that those variants might have contributed to an immunization of the mother against the integrin IIb carrying the missense *de novo* variant, hypothesis supported by the existence of reported novel antigens responsible of FNAIT (Poles *et al.*, 2013). Another hypothesis would be a contribution of those heterozygous *ITGA2B* variants to FICH by acting as a risk factor in combination with another triggering event.

A *de novo* non NMD LoF variant in *DCAF5* (DDB1 and CUL4 Associated Factor 5) was detected in one fetus. This gene encodes a substrate receptor for the CUL4-DDB1 E3 ubiquitin-protein ligase complex. Inactivation of its DDB1 partner in mouse leads to intracerebral hemorrhage (Cang *et al.*, 2006). In the literature, three unrelated patients were reported with *de novo* deletions region encompassing *DCAF5* in the 14q24.1q24.3 region (sizes of 5.4, 2.8 and 2.3 Mb). None of these patients presented an ICH history (Oehl-Jaschkowitz *et al.*, 2014). In a recent report, a 4-month-old child with respiratory failure, hypotonia, and dysmorphia was found to carry a *de novo* non-NMD stop codon variant in *DCAF5*. Based on all these data, we consider *DCAF5* as a candidate gene in the pathogenesis of ICH, but we could not formally establish its causality so far.

Finally, we identified several genes with *de novo* missense variants showing a significant enrichment ($p<0.05$) for rare and damaging variants compared to gnomAD. First, *EXD3* encodes a protein possessing an exoribonuclease activity without phenotype data available, either in human or in animal models. Second, *ZFYVE28* encodes a FYVE domain-containing protein, a group of proteins involved in intracellular trafficking, signal transduction and

regulation of the actin cytoskeleton. Its function is still poorly understood and it has not yet been associated to a human disease so far.

Fetal phenotypes associated with genes involved in metabolic disorders may mimic phenotypes encountered in fetal intracerebral hemorrhage

Several pathogenic or likely pathogenic variants in genes involved in metabolism were identified in our cohort of fetuses referred for suspected ICH. We identified one pathogenic homozygous LoF variant in *COQ2*, encoding a mitochondrial 4-hydroxybenzoate polyprenyltransferase. Pathogenic variants of *COQ2* are associated with coenzyme Q10 deficiency (Multiple-System Atrophy Research Collaboration. NEJM 2013; Jakobs *et al.,* 2013). In this condition, prenatal brain lesions are highly variable and may include leukoencephalopathy, periventricular cysts, cerebellar atrophy, agenesis of the corpus callosum and ventriculomegaly (Hashemi *et al.,* 2020; Münch *et al.,* 2023). In addition, we identified, a pathogenic homozygous variant in *ATP5PO.* Pathogenic variants of *ATP5PO* could have serious consequences for cellular function and are associated with various mitochondrial diseases (Zech *et al.,* 2022). The variant identified in fetus F58 was the same than reported in two families showing atrophy, ventriculomegaly, periventricular cysts, white matter hypersignals, a thin corpus callosum and a cavum of the septum pellucidum on their brain MRI scans. One patient also presented with stigmata of old bilateral hemorrhages in the frontal and parietal lobes (Ganapathi *et al.,* 2022). In the fetus from our cohort, fetal MRI at 28 WG showed a midline anomaly, cerebellar hypoplasia, hypoplasia of the corpus callosum and the presence of periventricular cystic images suggestive of a recent intraventricular and intraparenchymal hemorrhagic process. Post-mortem examination was not accepted by parents. Another fetus carried a probably pathogenic homozygous missense variant in *NAXD,* a gene involved in metabolite repair. *NAXD* is associated with an infantile onset neurodegenerative condition including encephalopathy with brain edema and/or leukoencephalopathy (Van Bergen *et al.,* 2018; Majethia *et al.,* 2021). Additionally, using a gene-based collapsing test analysis of our cohort of 113 fetuses, we identified three fetuses with pathogenic variants in *PDHA1,* a gene involved in a mitochondrial metabolic pathway, with signs encountered in FICH, including ventriculomegaly, leukomalacia, dysgenesis of the corpus callosum, and abnormal posterior fossa (Coste *et al.,* 2022). In conclusion, associated signs of all these conditions are similar to those encountered in fetuses with COL4A1/A2 collagenopathy including ventriculomegaly and clastic lesions, which is why fetuses with pathogenic variants in metabolism genes were initially

suspected of FICH and referred for *COL4A1/COL4A2* gene screening. Overall, these results illustrate the complexity of phenotype interpretation in fetus, and the complexity of etiological orientation when the associated signs of different conditions are similar.

Strong genetic heterogeneity in FICH

Several disease-causing and candidate genes have been identified by our WES analysis of fetuses with suspected ICH and for each gene we found a low number of mutated fetuses. These findings are highly suggestive of an extreme genetic heterogeneity represented by a large number of potential genes involved in multiple pathophysiological processes (Figure 5). There are few reports of pangenomic analysis in fetus with suspected ICH. To our knowledge, in literature, only one study performed a WES analysis in a cohort of patients with antenatal ICH (Hausman-Kedem et *al.*, 2021). The authors reported results of 26 probands and causative / likely causative variants were found in 4 patients (*COL4A1*, *COL4A2* and *TREX1*). These results illustrate the interest of the WES but they also highlight that many fetuses remain negative after WES. There are still many genetic causes of FICH to be discovered, exome/whole genome analysis on large international fetal cohorts combined with data sharing is an indispensable strategy for identifying new genes responsible for ICH. In addition, other mechanisms should be investigated like pathogenic variations in non-coding regions, digenism or oligogenism inheritance or epigenetic alterations.

Clinical relevance

WES is a powerful tool for deciphering genetic causes of suspected FICH, for providing diagnostic tools, for better understanding genotype-phenotype correlations and for identifying possible therapeutic targets. Most importantly, identification of the underlying genetic causes is crucial for genetic counseling in order to prevent recurrence of ICH in future pregnancies by offering prenatal diagnosis. In addition, our findings point the complexity of diagnosis in deceased fetuses, partially due to phenotypic similarities between conditions associated with a cerebral hemorrhage and those reported in metabolic encephalopathy; WES is of major interest for these fetuses.

Altogether, we strongly suggest that an exome-wide analysis should be part of the routine evaluation of fetuses with suspected ICH, prioritizing recessive and *de novo* variants, especially when parents are both asymptomatic.

DATA AVAILABILITY

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

ACKNOWLEDGEMENTS

The authors thank families for their participation in this study. We also acknowledge clinicians, geneticists and pathologists who referred fetuses enrolled in this study.

CONTRIBUTORS

T.C., A.L.L and E.T.L. were in charge of the conception, design of the study and drafting of the article. C.A., J.M., T.A-B., F.P., D.H., A.G.B., R.R and R.P contributed to the acquisition, analysis of data and contributed to the manuscript.

FUNDING

This work was supported by the National Institutes of Health (R01NS096173 grant) and by the Assistance Publique des Hôpitaux de Paris (AP-HP).

POTENTIAL CONFLICTS OF INTEREST

Nothing to report.

ETHICS APPROVAL

This study was approved by the INSERM Ethics Review Committee (IRB00003888) and an informed and signed consent was obtained from parents of fetus probands included in the study.

FIGURES AND TABLES

SUPPLEMENTAL DATA

REFERENCES

Adiego B, Martínez-Ten P, Bermejo C, Estévez M, Recio Rodriguez M, Illescas T. Fetal intracranial hemorrhage. Prenatal diagnosis and postnatal outcomes. The Journal of Maternal-Fetal & Neonatal Medicine. 2 janv 2019;32(1):21 30.

Alsultan A, Gale AJ, Kurban K, Khalifah M, Albadr FB, Griffin JH. Activation-resistant homozygous protein C R229W mutation causing familial perinatal intracranial hemorrhage and delayed onset of thrombosis. Thrombosis Research. 2016;143:17 21.

Bussel JB, Vander Haar EL, Berkowitz RL. New developments in fetal and neonatal alloimmune thrombocytopenia. American Journal of Obstetrics and Gynecology. 2021;225(2):120 7.

Bouchghoul H, Quelin C, Loget P, Encha-Razavi F, Senat MV, Maheut L, et al. Fetal cerebral hemorrhage due to X-linked GATA1 gene mutation. Prenatal Diagnosis. 2018;38(10):772 8.

Cavaliere AF, Turrini I, Pallottini M, Vidiri A, Marchi L, Perelli F, et al. Genetic Profiling of Idiopathic Antenatal Intracranial Haemorrhage: What We Know? Genes. avr 2021;12(4):573.

Cang Y, Zhang J, Nicholas SA, Bastien J, Li B, Zhou P, et al. Deletion of DDB1 in Mouse Brain and Lens Leads to p53-Dependent Elimination of Proliferating Cells. Cell. déc 2006;127(5):929 40.

Coste T, Vincent-Delorme C, Stichelbout M, Devisme L, Gelot A, Deryabin I, et al. COL4A1/COL4A2 and inherited platelet disorder gene variants in fetuses showing intracranial hemorrhage. Prenatal Diagnosis. 2022;42(5):601-610.

Coste T, Aloui C, Petit F, Moutton S, Devisme L, Wells CF, et al. Rare metabolic disease mimicking COL4A1/COL4A2 fetal brain phenotype. Ultrasound in Obstetrics & Gynecology. 2022;60(6):805 11.

Downes K, Megy K, Duarte D, Vries M, Gebhart J, Hofer S, et al. Diagnostic high-throughput sequencing of 2396 patients with bleeding, thrombotic, and platelet disorders. Blood. 5 déc 2019;134(23):2082 91.

Eklund L, Kangas J, Saharinen P. Angiopoietin–Tie signalling in the cardiovascular and lymphatic systems. Clin Sci (Lond). 1 janv 2017;131(1):87 103.

Fong CY, Mumford AD, Likeman MJ, Jardine PE. Cerebral palsy in siblings caused by compound heterozygous mutations in the gene encoding protein C. Developmental Medicine & Child Neurology. 2010;52(5):489 93.

Ganapathi M, Friocourt G, Gueguen N, et al. A homozygous splice variant in ATP5PO, disrupts mitochondrial complex V function and causes Leigh syndrome in two unrelated families. Journal of Inherited Metabolic Disease. 2022;45(5):996-1012.

Ghi T, Simonazzi G, Perolo A, Savelli L, Sandri F, Bernardi B, et al. Outcome of antenatally diagnosed intracranial hemorrhage: case series and review of the literature. Ultrasound in Obstetrics & Gynecology. 2003;22(2):121 30.

Germeshausen M, Ballmaier M. CAMT-MPL: congenital amegakaryocytic thrombocytopenia caused by MPL mutations - heterogeneity of a monogenic disorder - a comprehensive analysis of 56 patients. Haematologica. 2021;106(9):2439 48.

Griffin ME, Sorum AW, Miller GM, Goddard WA, Hsieh-Wilson LC. Sulfated glycans engage the Ang/Tie pathway to regulate vascular development. Nat Chem Biol. févr 2021;17(2):178 86.

Hashemi SS, Zare-Abdollahi D, Bakhshandeh MK, Vafaee A, Abolhasani S, Inanloo Rahatloo K, et al. Clinical spectrum in multiple families with primary COQ10 deficiency. American Journal of Medical Genetics Part A. 2021;185(2):440-52.

Hilton E, Johnston J, Whalen S, Okamoto N, Hatsukawa Y, Nishio J, et al. BCOR analysis in patients with OFCD and Lenz microphthalmia syndromes, mental retardation with ocular anomalies, and cardiac laterality defects. Eur J Hum Genet. 2009;17(10):1325 35.

Itai T, Miyatake S, Taguri M, Nozaki F, Ohta M, Osaka H, et al. Prenatal clinical manifestations in individuals with COL4A1/2 variants. J Med Genet. 2021;58(8):505 13.

Jakobs BS, van den Heuvel LP, Smeets RJP, de Vries MC, Hien S, Schaible T, et al. A novel mutation in COQ2 leading to fatal infantile multisystem disease. Journal of the Neurological Sciences. 2013;326(1 2):24 8.

Korhonen EA, Lampinen A, Giri H, Anisimov A, Kim M, Allen B, et al. Tie1 controls angiopoietin function in vascular remodeling and inflammation. J Clin Invest. 126(9):3495 510.

Kunishima S, Kashiwagi H, Otsu M, Takayama N, Eto K, Onodera M, et al. Heterozygous ITGA2B R995W mutation inducing constitutive activation of the αIIbβ3 receptor affects proplatelet formation and causes congenital macrothrombocytopenia. Blood. 2011;117(20):5479 84.

La Porta S, Roth L, Singhal M, Mogler C, Spegg C, Schieb B, et al. Endothelial Tie1–mediated angiogenesis and vascular abnormalization promote tumor progression and metastasis. J Clin Invest. 2018;128(2):834 45.

Lecca M, Pehlivan D, Suñer DH, Weiss K, Coste T, Zweier M, et al. Bi-allelic variants in the ESAM tight-junction gene cause a neurodevelopmental disorder associated with fetal intracranial hemorrhage. Am J Hum Genet. 2023 Apr 6;110(4):681-690.

Leutenegger AL, Prum B, Génin E, Verny C, Lemainque A, Clerget-Darpoux F, et al. Estimation of the Inbreeding Coefficient through Use of Genomic Data. The American Journal of Human Genetics. 2003;73(3):516 23.

Leutenegger AL, Sahbatou M, Gazal S, Cann H, Génin E. Consanguinity around the world: what do the genomic data of the HGDP-CEPH diversity panel tell us? Eur J Hum Genet. 2011 May;19(5):583-7.

Lim MS, Shin JE, Lee SM, Eun HS, Park MS, Park KI, et al. Diagnosis of Severe Protein C Deficiency Confirmed by Presence of Rare PROC Gene Mutation. Neonatal Med. 2016;23(4):233.

MacArthur DG, Balasubramanian S, Frankish A, Huang N, Morris J, Walter K, et al. A Systematic Survey of Loss-of-Function Variants in Human Protein-Coding Genes. Science. 2012;335(6070):823 8.

Majethia P. NAD(P)HX dehydratase (NAXD) deficiency due to a novel biallelic missense variant and review of literature. European Journal of Medical Genetics. 2021;6.

Martin G, Thomas MA, Wei XC, Le D. Diffuse Intracerebral Hemorrhage in an Infant With a Novel Homozygous Variant Leading to Severe Protein C Deficiency. J Pediatr Hematol Oncol. 2021;43(6):e763 5.

Megy K, Downes K, Morel-Kopp MC, Bastida JM, Brooks S, Bury L, et al. GoldVariants, a resource for sharing rare genetic variants detected in bleeding, thrombotic, and platelet disorders: Communication from the ISTH SSC Subcommittee on Genomics in Thrombosis and Hemostasis. Journal of Thrombosis and Haemostasis. 2021;19(10):2612 7.

Meuwissen MEC, Halley DJJ, Smit LS, Lequin MH, Cobben JM, de Coo R, et al. The expanding phenotype of COL4A1 and COL4A2 mutations: clinical data on 13 newly identified families and a review of the literature. Genet Med. nov 2015;17(11):843 53.

Monteagudo A. Intracranial Hemorrhage. Am J Obstet Gynecol. 2020 Dec;223(6):B34-B37.

Münch J, Prasuhn J, Laugwitz L, Fung CW, Chung BHY, Bellusci M, et al. Neuroimaging in Primary Coenzyme-Q10-Deficiency Disorders. Antioxidants. 2023;12(3):718.

Nasdala I, Wolburg-Buchholz K, Wolburg H, Kuhn A, Ebnet K, Brachtendorf G, et al. A Transmembrane Tight Junction Protein Selectively Expressed on Endothelial Cells and Platelets. Journal of Biological Chemistry. 2002;277(18):16294 303.

Ng D, Thakker N, Corcoran CM, Donnai D, Perveen R, Schneider A, et al. Oculofaciocardiodental and Lenz microphthalmia syndromes result from distinct classes of mutations in BCOR. Nat Genet. 2004;36(4):411 6.

Nurden P, Debili N, Coupry I, Bryckaert M, Youlyouz-Marfak I, Solé G, et al. Thrombocytopenia resulting from mutations in filamin A can be expressed as an isolated syndrome. Blood. 2011;118(22):5928 37.

Oehl-Jaschkowitz B, Vanakker OM, Paepe AD, Menten B, Martin T, Weber G, et al. Deletions in 14q24.1q24.3 are associated with congenital heart defects, brachydactyly, and mild intellectual disability. American Journal of Medical Genetics Part A. 2014;164(3):620 6.

Pedersen BS, Quinlan AR. Who's Who? Detecting and Resolving Sample Anomalies in Human DNA Sequencing Studies with Peddy. The American Journal of Human Genetics. 2017;100(3):406 13.

Piesse C, Tymms M, Garrafa E, Gouzy C, Lacasa M, Cadel S, et al. Human aminopeptidase B (rnpep) on chromosome 1q32.2: complementary DNA, genomic structure and expression. Gene. 2002;292(1):129 40.

Poles A, Woźniak MJ, Walser P, Ridgwell K, Fitzgerald J, Green A, et al. A V740L mutation in glycoprotein IIb defines a novel epitope (War) associated with fetomaternal alloimmune thrombocytopenia. Transfusion. 2013;53(9):1965 73.

Richards S. 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. 2015;(17):405 24.

Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: A Matching Tool for Connecting Investigators with an Interest in the Same Gene. Human Mutation. 2015;36(10):928 30.

Stutterd C, Savoia H, Fink AM, Stark Z. Severe fetal ischaemic brain injury caused by homozygous protein C deficiency. Prenatal Diagnosis. 2014;34(2):192 4.

The Multiple-System Atrophy Research Collaboration. Mutations in COQ2 in Familial and Sporadic Multiple-System Atrophy. N Engl J Med. 2013;369(3):233 44.

Van Bergen NJ, Guo Y, Rankin J, Paczia N, Becker-Kettern J, Kremer LS, et al. NAD(P)HX dehydratase (NAXD) deficiency: a novel neurodegenerative disorder exacerbated by febrile illnesses. Brain. 2019;142(1):50 8.

Xu M, Jin P, Huang Y, Qian Y, Lin M, Zuo J, et al. Case report: Prenatal diagnosis of fetal intracranial hemorrhage due to compound mutations in the JAM3 gene. Front Genet. 2022;13:1036231.

Yasukawa R, Moteki H, Nishio S ya, Ishikawa K, Abe S, Honkura Y, et al. The Prevalence and Clinical Characteristics of TECTA-Associated Autosomal Dominant Hearing Loss. Genes. 2019;10(10):744.

Yoneda Y, Haginoya K, Kato M, Osaka H, Yokochi K, Arai H, et al. Phenotypic Spectrum of COL4A1 Mutations: Porencephaly to Schizencephaly. Annals of Neurology. 2013;73(1):48 57.

Zambrano S, Rodriguez PQ, Guo J, Möller-Hackbarth K, Schwarz A, Patrakka J. FYVE domain-containing protein ZFYVE28 regulates EGFR-signaling in podocytes but is not critical for the function of filtration barrier in mice. Sci Rep. 2018;8(1):4712.

Zech M, Kopajtich R, Steinbrücker K, et al. Variants in Mitochondrial ATP Synthase Cause Variable Neurologic Phenotypes. Annals of Neurology. 2022;91(2):225-237.

Figure 1

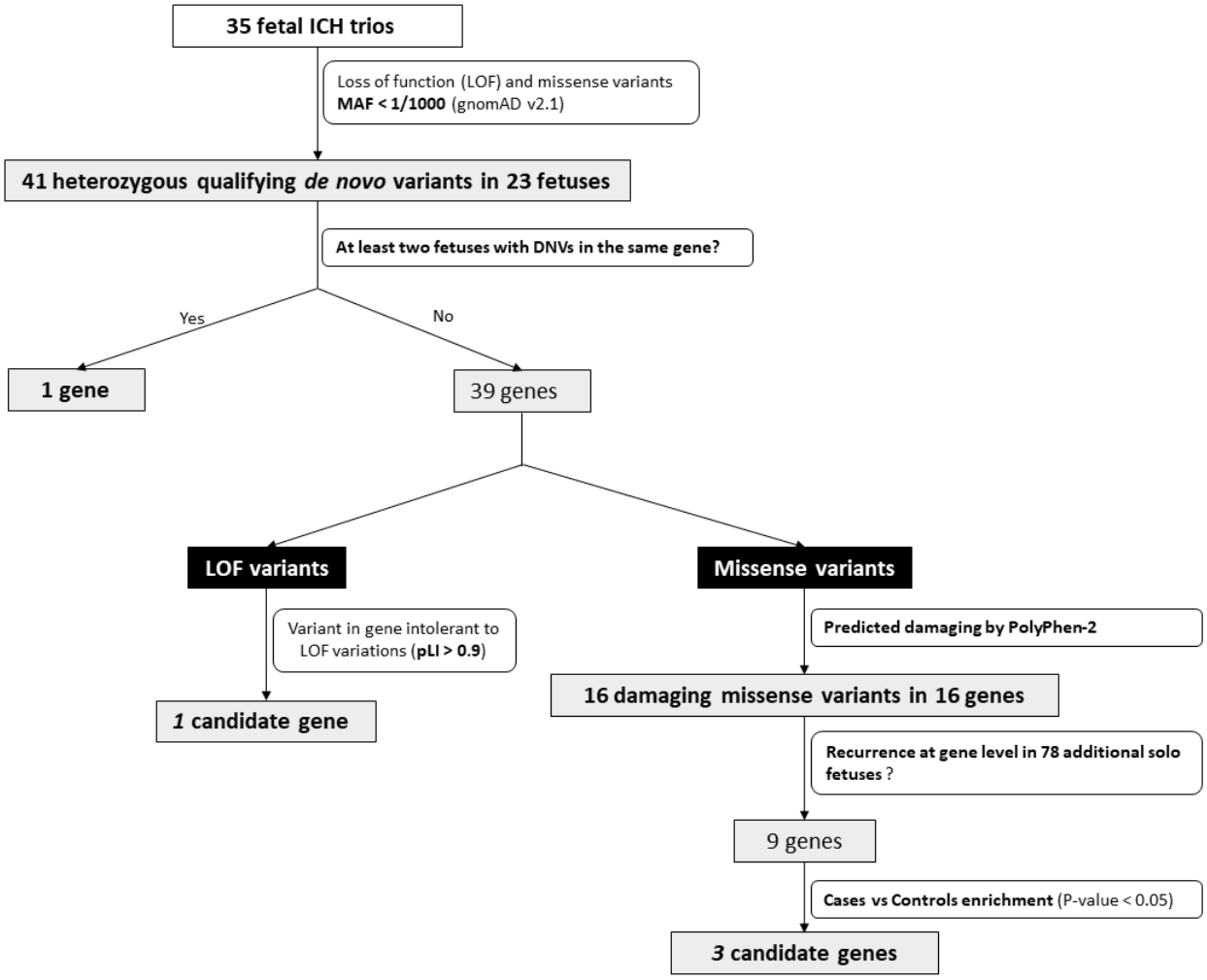

35 fetal ICH trios
Loss of function (LOF) and missense variants MAF < 1/1000 (gnomAD v2.1)
41 heterozygous qualifying de novo variants in 23 fetuses
At least two fetuses with DNVs in the same gene?
Yes
No
1 gene
39 genes
LOF variants
Missense variants
Variant in gene intolerant to LOF variations (pLI > 0.9)
Predicted damaging by PolyPhen-2
1 candidate gene
16 damaging missense variants in 16 genes
Recurrence at gene level in 78 additional solo fetuses ?
9 genes
Cases vs Controls enrichment (P-value < 0.05)
3 candidate genes

Figure 2

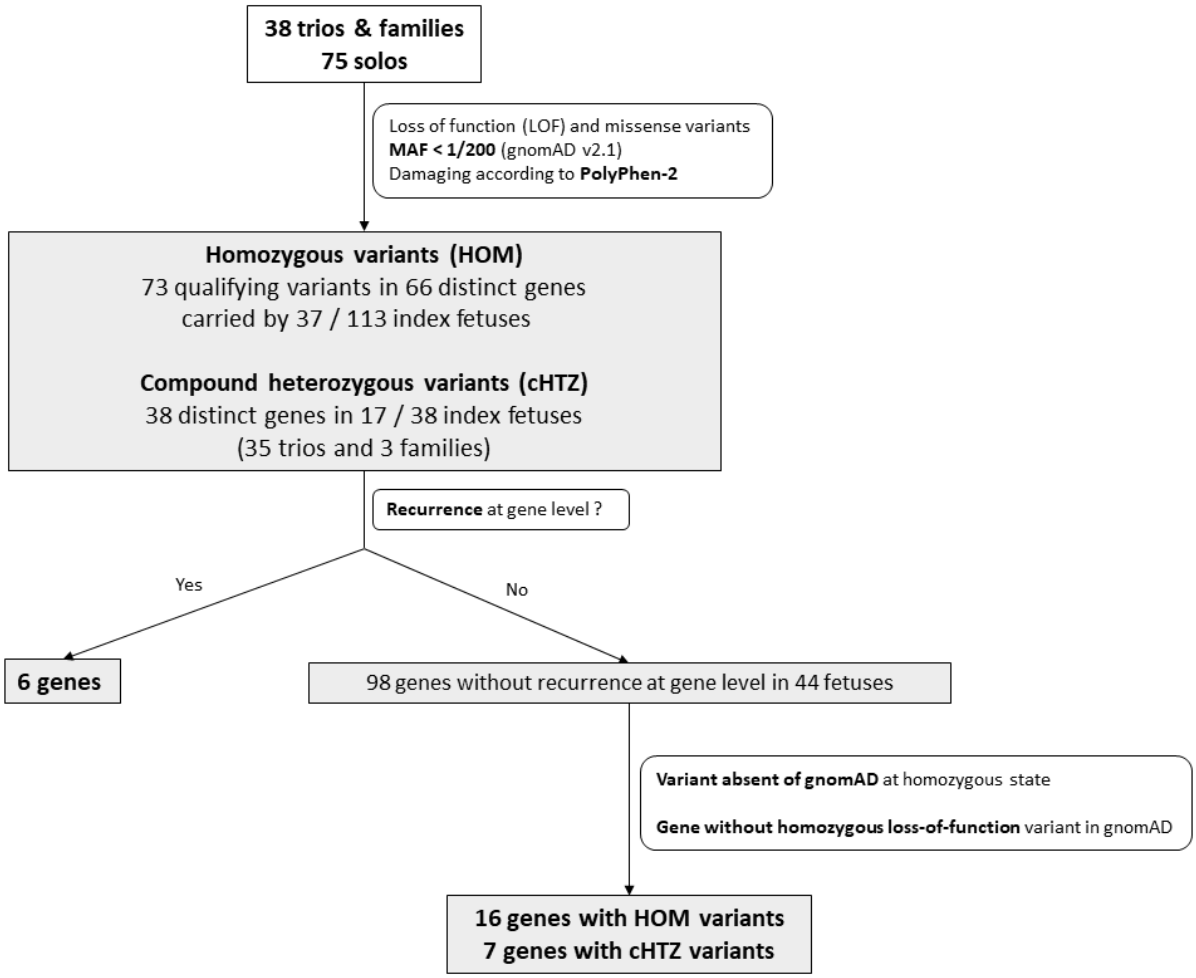

38 trios & families
75 solos
Loss of function (LOF) and missense variants
MAF < 1/200 (gnomAD v2.1)
Damaging according to PolyPhen-2
Homozygous variants (HOM)
73 qualifying variants in 66 distinct genes
carried by 37 / 113 index fetuses
Compound heterozygous variants (cHTZ)
38 distinct genes in 17 / 38 index fetuses
(35 trios and 3 families)
Recurrence at gene level ?
Yes
No
6 genes
98 genes without recurrence at gene level in 44 fetuses
Variant absent of gnomAD at homozygous state
Gene without homozygous loss-of-function variant in gnomAD
16 genes with HOM variants
7 genes with cHTZ variants

Figure 3

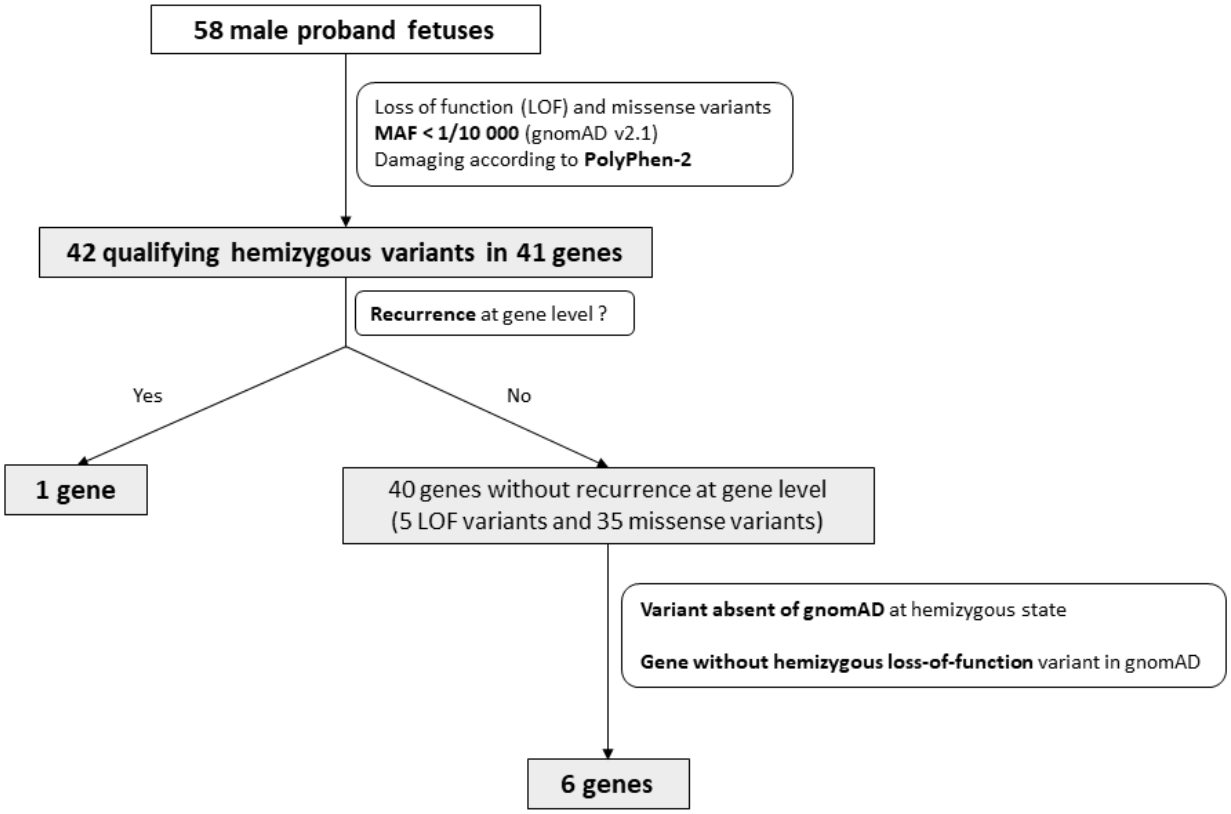

58 male proband fetuses
Loss of function (LOF) and missense variants
MAF < 1/10 000 (gnomAD v2.1)
Damaging according to PolyPhen-2
42 qualifying hemizygous variants in 41 genes
Recurrence at gene level ?
Yes
No
1 gene
40 genes without recurrence at gene level
(5 LOF variants and 35 missense variants)
Variant absent of gnomAD at hemizygous state
Gene without hemizygous loss-of-function variant in gnomAD
6 genes

Figure 4

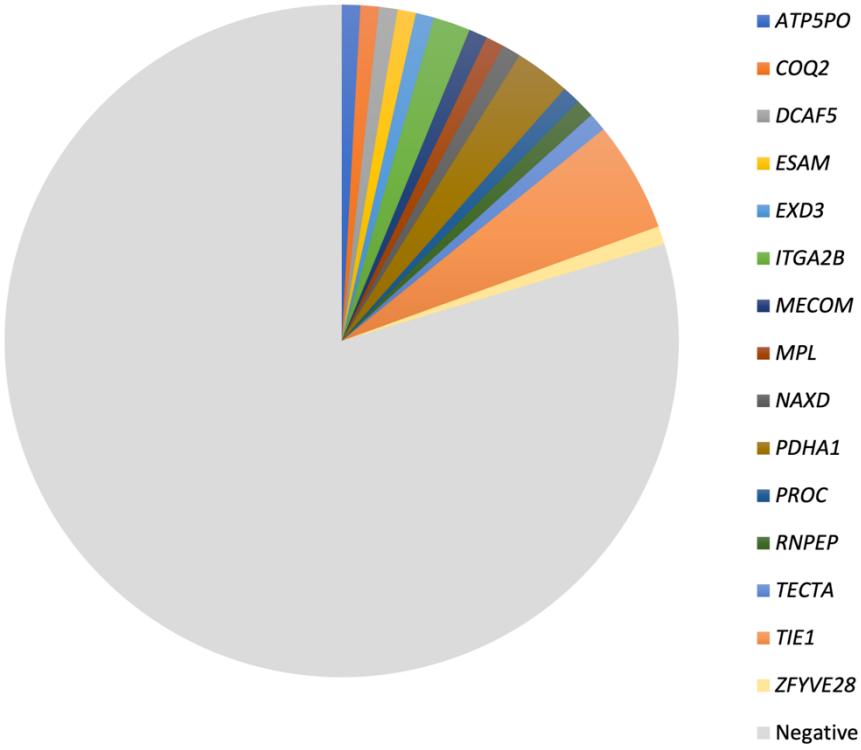

ATP5PO
COQ2
DCAF5
ESAM
EXD3
ITGA2B
MECOM
MPL
NAXD
PDHA1
PROC
RNPEP
TECTA
TIE1
ZFYVE28
Negative

Figure 5

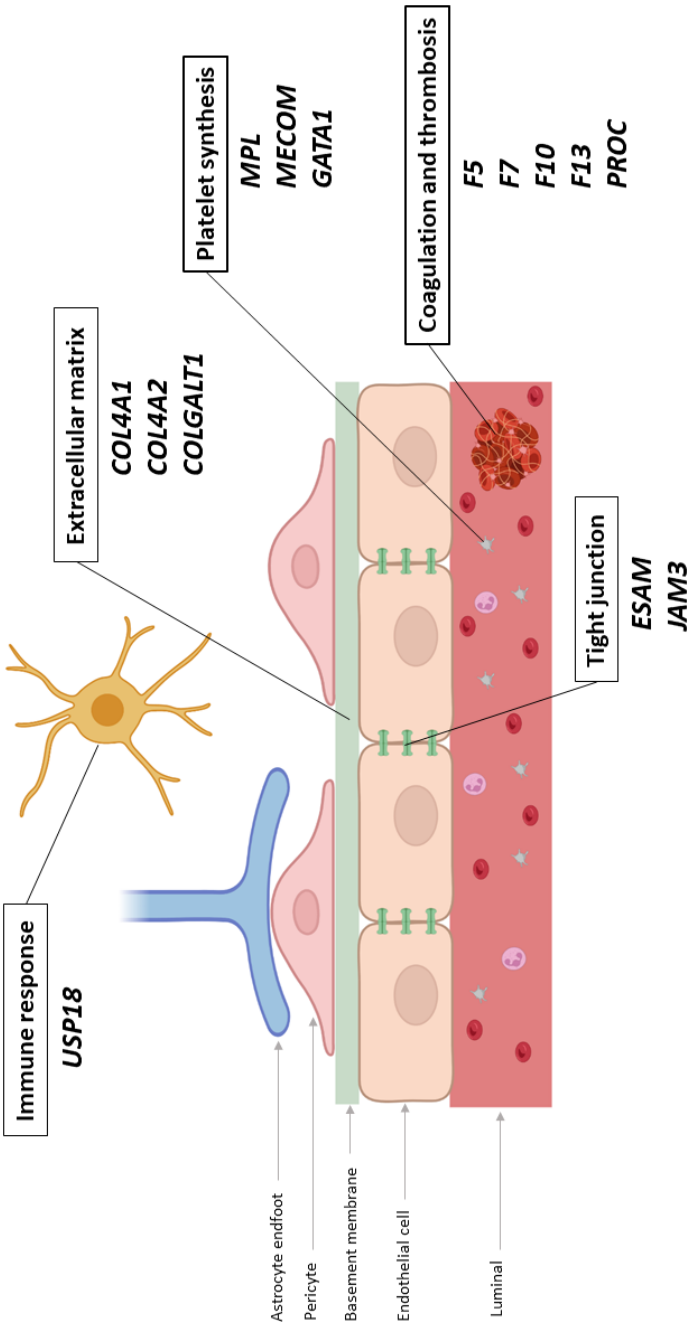

Immune response
USP18
Extracellular matrix
COL4A1
COL4A2
COLGALT1
Platelet synthesis
MPL
MECOM
GATA1
Coagulation and thrombosis
F5
F7
F10
F13
PROC
Tight junction
ESAM
JAM3
Astrocyte endfoot
Pericyte
Basement membrane
Endothelial cell
Luminal

Table 1. Rare LoF and predicted damaging *de novo* variants identified in the 35 trios.

Fetus ID	Gene	Amino acid change	gnomAD (v2.1) allele frequency	PolyPhen-2	Recurrence in non-trio fetuses	Cases/Controls p-value*
F34	*DCAF5*	p.D478Afs*17	Abs	NA	No	NA
F94	*DSG4*	p.S285*	Abs	NA	No	NA
F108	*EXD3*	p.G672E	0.0000098	Possibly damaging	Yes	0.0039
F12	*F10*	p.R68C	Abs	Probably damaging	No	NA
F106	*ITGA2B*	p.P434S	Abs	Probably damaging	Yes	0.0070
F94	*KIRREL3*	p.I581V	Abs	Possibly damaging	Yes	0.19
F94	*LRTM1*	p.P22L	0.00002	Possibly damaging	No	NA
F30	*METTL13*	p.E426K	Abs	Probably damaging	Yes	0.086
F92	*MYO9B*	p.L384V	Abs	Probably damaging	Yes	0.66
F108	*NKX3-1*	p.C225R	Abs	Probably damaging	No	NA
F01	*PABPC1L*	p.Q402Hfs*25	Abs	NA	No	NA
F105	*PLTP*	p.R372H	0.00036	Probably damaging	Yes	0.19
F35	*PLXNA2*	p.Y95Cfs*76	Abs	NA	No	NA
F113	*RCOR2*	p.F369L	Abs	Probably damaging	No	NA
F111	*SLC23A1*	p.D291N	0.000016	Probably damaging	No	NA
F109	*SMC3*	p.E457K	Abs	Possibly damaging	No	NA
F103	*TBC1D13*	p.D124N	0.000060	Probably damaging	No	NA
F17	*TECTA*	p.I1996V	Abs	Probably damaging	Yes	0.040
F17	*TRA2A*	p.R95Gfs*51	Abs	NA	No	NA
F109	*ZFYVE28*	p.A515V	Abs	Possibly damaging	Yes	0.00067
F35	*ZNF560*	p.R694Q	0.00013	Possibly damaging	Yes	0.08

Abs: absent; NA: not applicable; * Analysis performed for *de novo* variants showing recurrence in non-trios fetuses.

Table 2. Identified genes according to a recessive inheritance hypothesis and mutated in at least two fetuses.

Fetus ID	Consanguinity	Gene	Amino acid change	Status	Allele frequency in gnomAD v2.1 / Number of homozygous variants in gnomAD v2.1	Total number of homozygous LoF variants at gene level in gnomAD v2.1
F108	No	*EPPK1*	p.D2018N	cHTZ	0.00095 / 1	17
			p.V1020M	cHTZ	Abs	
F34	No	*EPPK1*	p.R2179*	cHTZ	0.00009 / 0	17
			p.V1902X	cHTZ	0.0005 / 0	
			p.R274C	cHTZ	0.00093 / 2	
F97	No	*GOLGA6L2*	p.R736*	HOM	0.00001093 / 0	36
F08	No	*GOLGA6L2*	p.S63L	HOM	0.00003962 / 0	36
F90	No	*HLA-DRB1*	p.E98X	HOM	0.005478 / 76	809
F94	No	*HLA-DRB1*	p.E98X	HOM	0.005478 / 76	809
F01	Yes	*MUC4*	p.T94M	HOM	0.001393 / 0	9497
F112	No	*MUC4*	p.A3982T	HOM	0.004509 / 135	9497
F81	Yes	*MUC4*	p.R394H	cHTZ	0.0046 / 21	9497
			p.A925E	cHTZ	0.00094 / 1	
			p.A459T	cHTZ	0.00096 / 1	
F111	No	*MUC19*	p.T2299K	cHTZ	0.0008 / 0	42702
			p.G5952V	cHTZ	0.0008 / 0	
F104	No	*MUC19*	c.4018-2A>G	cHTZ	Abs	42702
			p.G5683A	cHTZ	0.001 / 0	
F34	No	*MUC19*	p.G247A	cHTZ	0.0014 / 3	42702
			p.S1828A	cHTZ	0.0012 / 1	
			p.P7722S	cHTZ	0.0016 / 0	
F92	Yes	*SCRIB*	p.R221C	HOM	0.00001615 / 0	0
F112	No	*SCRIB*	p.G1526R	cHTZ	0.000009 / 0	0
			p.H21Y	cHTZ	Abs	
F92	Yes	*ARHGAP4*	p.R268W	HEM	Abs	11
F11	No	*ARHGAP4*	p.R80C	HEM	Abs	11

HOM: homozygous; HTZc: compound heterozygous; HEM: hemizygous; LoF: loss-of-function;

Table 3. Genes without recurrence at gene level identified according to a recessive inheritance hypothesis.

Fetus ID	Consanguinity*	Gene	Amino acid change	Status	Gene function and associated Mendelian disease
F58	Yes	*ATP5PO*	p.?	HOM	Encodes a V-complex assembly protein, and plays an essential role in the formation and assembly of complex V of the mitochondrial respiratory chain. Associated disease : Mitochondrial complex V (ATP synthase) deficiency, nuclear type 7 (MIM# 620359)
F32	No	*COQ2*	p.N351Ifs*15	HOM	Catalyzes the prenylation of para-hydroxybenzoate (PHB) with an all-trans polyprenyl group. Associated disease: Coenzyme Q10 deficiency (MIM# 607426)
F92	Yes	*ESAM*	p.L12*	HOM	Endothelial cell-selective adhesion molecule
F92	Yes	*MAFA*	p.A93Pfs*27	HOM	Involved either as an oncogene or as a tumor suppressor, depending on the cell context. Associated disease: Insulinomatosis and diabetes mellitus (MIM# 147630)
F35	No	*RNPEP*	p.E269*	HOM	Exopeptidase which selectively removes arginine and/or lysine residues from the N-terminus of several peptide substrates
F58	Yes	*ASIC4*	p.P240L	HOM	Channel implicated in synaptic transmission, pain perception as well as mechanoperception
F58	Yes	*CCDC102A*	p.E496K	HOM	Predicted to be part of myosin complex
F108	No	*CRELD1*	p.R107H	HOM	Protein disulfide isomerase. Associated disease: Atrioventricular septal defect with heterotaxy syndrome (MIM# 606217)
F34	No	*FGD5*	p.R1080C	HOM	Enable guanyl-nucleotide exchange factor activity and small GTPase binding activity. May regulate proangiogenic action of VEGF in vascular endothelial cells, including network formation, directional movement and proliferation
F58	Yes	*MYG1*	p.Y235H	HOM	3'-5' RNA exonuclease which cleaves in situ on specific transcripts in both nucleus and mitochondrion
F01	Yes	*NAXD*	p.D232N	HOM	Catalyzes the dehydration of the S-form of NAD(P)HX at the expense of ATP, which is converted to ADP. Involved in metabolite repair. Associated disease: Early-onset progressive encephalopathy with brain edema and/or leukoencephalopathy-2 (MIM# 618321)
F01	Yes	*NOS3*	p.T854M	HOM	Produces nitric oxide (NO) which is implicated in vascular smooth muscle relaxation. NO mediates vascular endothelial growth factor (VEGF)-induced angiogenesis in coronary vessels and promotes blood clotting through the activation of platelets. Associated disease: Susceptibility to late-onset Alzheimer disease; hypertension; coronary artery spasm; ischemic stroke and placental abruption (MIM+ 163729)
F27	No	*NUAK2*	p.T10S	HOM	Stress-activated kinase involved in tolerance to glucose starvation. Induces cell-cell detachment by increasing F-actin conversion to G- Associated disease: Anencephaly (MIM# 619452)
F38	No	*PLEC*	p.R2206Q	HOM	Interlinks intermediate filaments with microtubules and microfilaments and anchors intermediate filaments to desmosomes or hemidesmosomes. Associated disease: Epidermolysis bullosa (MIM* 601282)
F58	Yes	*RET*	p.A766V	HOM	Receptor tyrosine-protein kinase involved in numerous cellular mechanisms including cell proliferation, neuronal navigation, cell migration, and cell differentiation. Associated diseases: Hirschsprung disease; medullary thyroid carcinoma; Multiple endocrine neoplasia and pheochromocytoma (MIM* 164761)
F58	Yes	*ZNF251*	p.S529C	HOM	May be involved in transcriptional regulation
F104	No	*ADRA2A*	p.A126V	cHTZ	Alpha-2 adrenergic receptors mediate the catecholamine-induced inhibition of adenylate cyclase through the action of G proteins
			p.C457Gfs*20	cHTZ	

F81	Yes	*LRP8*	p.P798R	cHTZ	Encodes a member of the low-density lipoprotein receptor (LDLR) family
			p.R472H	cHTZ	
F03	No	*MPL*	p.R102P	cHTZ	Receptor for thrombopoietin that acts as a primary regulator of megakaryopoiesis and platelet production. Associated disease: Congenital amegakaryocytic thrombocytopenia (CAMT-MPL) (MIM# 604498), thrombocythemia and myelofibrosis with myeloid metaplasia
			p.R537W	cHTZ	
F12	No	*PROC*	p.G215Dfs*40	cHTZ	Vitamin K-dependent serine protease that regulates blood coagulation by inactivating factors Va and VIIIa in the presence of calcium ions and phospholipids. Associated disease: Thrombophilia due to protein C deficiency (MIM# 612304)
			p.A388E	cHTZ	
F101	No	*TNC*	p.G1146E	cHTZ	Extracellular matrix protein implicated in guidance of migrating neurons as well as axons during development, synaptic plasticity as well as neuronal regeneration. Associated disease: Deafness (MIM# 615629)
			p.G1784D	cHTZ	
F12	No	*ZNF687*	p.L1142S	cHTZ	May be involved in transcriptional regulation and play a role in bone differentiation and development. Associated disease: Paget disease of bone-6 (MIM# 616833)
			p.R1154C	cHTZ	
F36	No	*ZNF865*	p.G780C	cHTZ	May be involved in transcriptional regulation
			p.P190S	cHTZ	
F77	No	*BCOR*	p.V1605L	HEM	Transcriptional corepressor. Associate disease: Syndromic microphthalmia-2 (MIM # 300166)
F92	Yes	*BTK*	p.D504N	HEM	Non-receptor tyrosine kinase indispensable for B lymphocyte development, differentiation and signaling. Associated disease: X-linked agammaglobulinemia (MIM# 300755)
F52	No	*MOSPD2*	p.S386L	HEM	Endoplasmic reticulum-anchored receptor that modulates interorganelle contacts by interacting with other organelle-bound proteins via their FFAT motif. Important role in endoplasmic reticulum and endosomes contacts
F51	No	*NKAP*	p.K256N	HEM	Involved in the activation of the ubiquitous transcription factor NF-kappaB. Associated disease: Hackmann-Di Donato-type X-linked syndromic intellectual developmental disorder (MIM# 301039)
F100	No	*PGK1*	p.I265T	HEM	Catalyzes one of the two ATP producing reactions in the glycolytic pathway via the reversible conversion of 1,3-diphosphoglycerate to 3-phosphoglycerate. Associated disease: Phosphoglycerate kinase-1 deficiency (MIM# 300653)
F82	No	*TSPAN7*	p.L88V	HEM	Mediates signal transduction events that play a role in the regulation of cell development, activation, growth and motility. Associated disease: Non-syndromic X-linked intellectual developmental disorder-58 (MIM# 300210)

*Consanguinity status was estimated on WES data using Fantasio R package. HOM: homozygous; cHTZ: compound heterozygous; HEM: hemizygous.

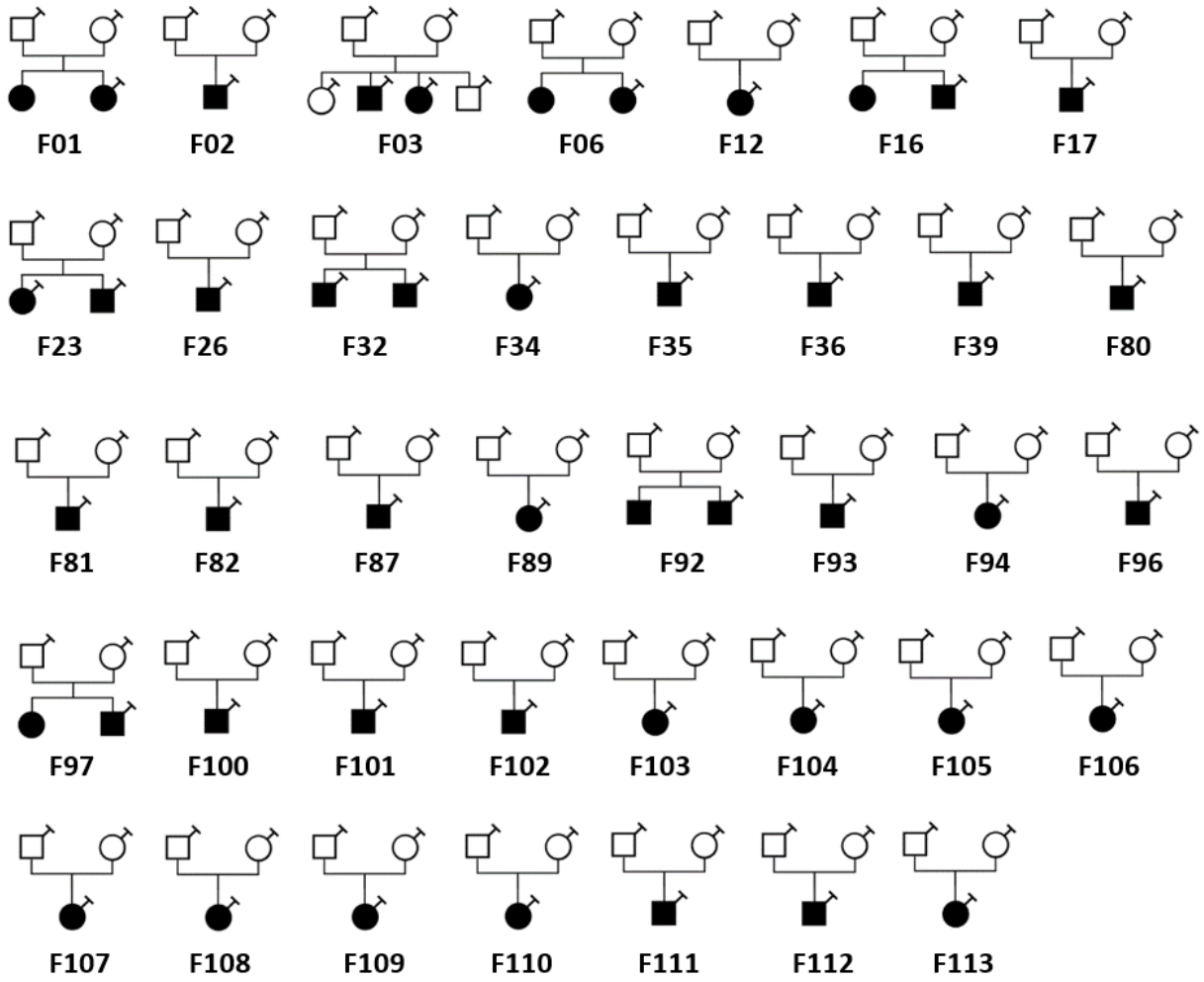
SUPPORTING INFORMATION

This Supporting Information includes:

Figure S1. Genealogical trees of the 113 fetuses

A/ Fetuses with trio exome sequencing (and relatives if available)

Black filled symbols indicate proven affected individuals. Syringe symbols indicate blood sampled individuals.



B/ Fetuses whose exome have only been performed in the index case (solo fetuses)

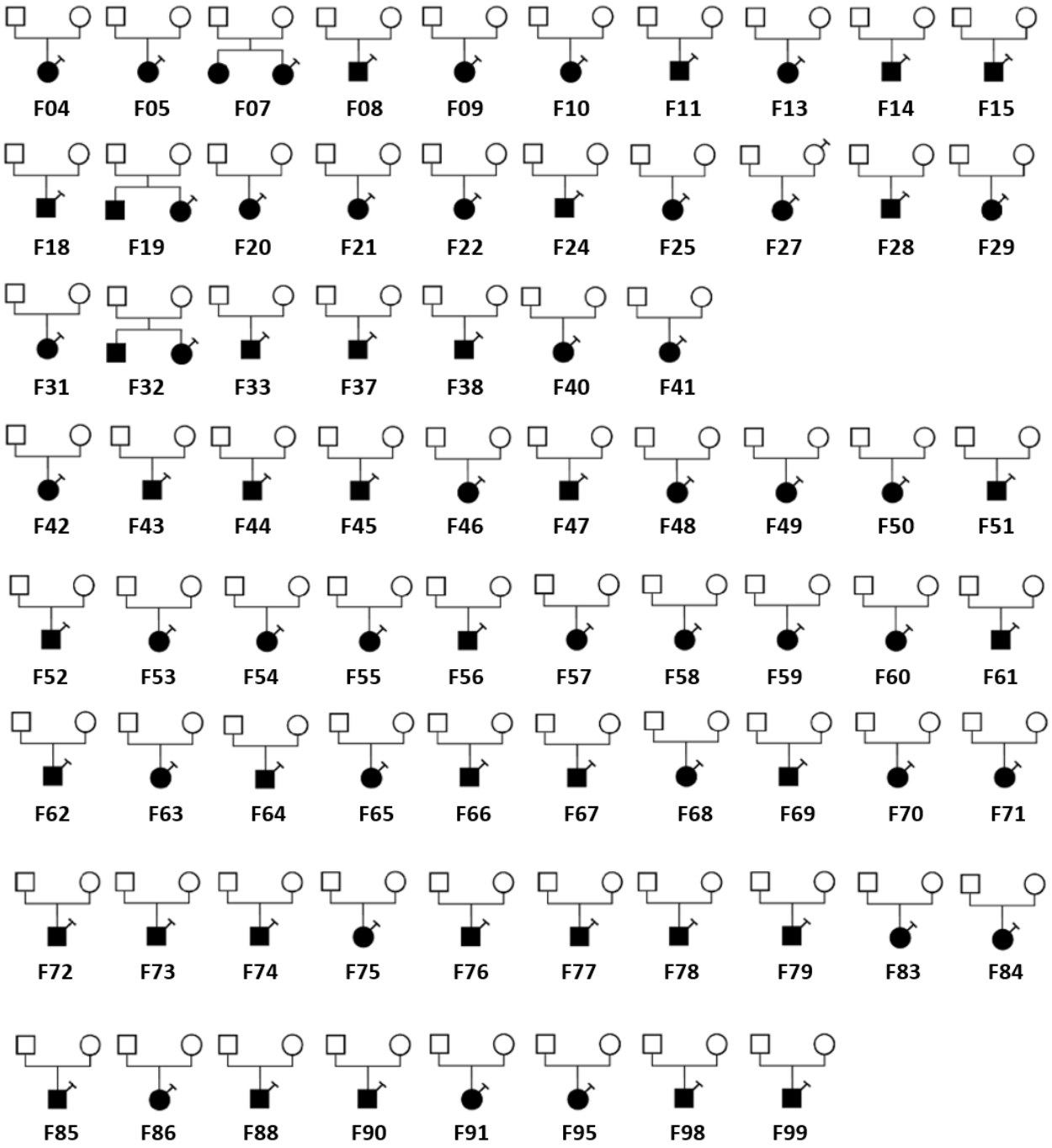
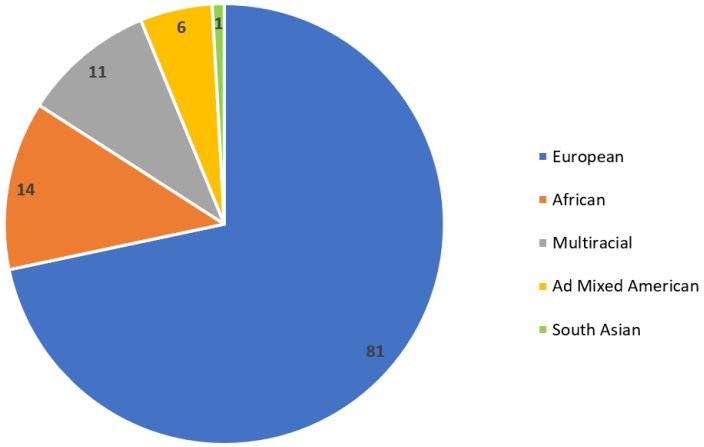


Figure S2. Geographical origin distribution of the 113 unrelated fetuses

A/ Pie chart of the number of fetuses in each geographical category



B / Results of principal component analysis using Peddy

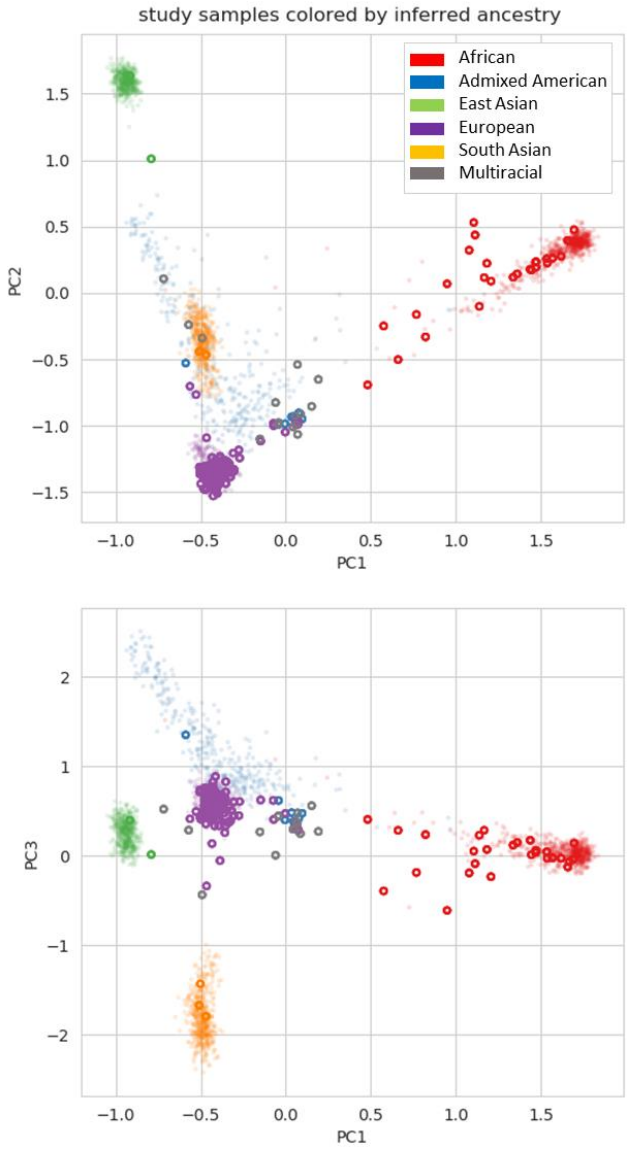


Figure S3. Characteristics of *ITGA2B* qualifying variants identified in fetuses

Panel A illustrate allelic positions of variants in the two families (F06 and F106). Panel B illustrates structural changes of ITGA2B de novo variants. The I766T variant leads to a loss of one hydrophobic interaction (L768-I766) and the P434S variant leads to a loss of hydrophobic interactions with L421, F432 and V437 residues.

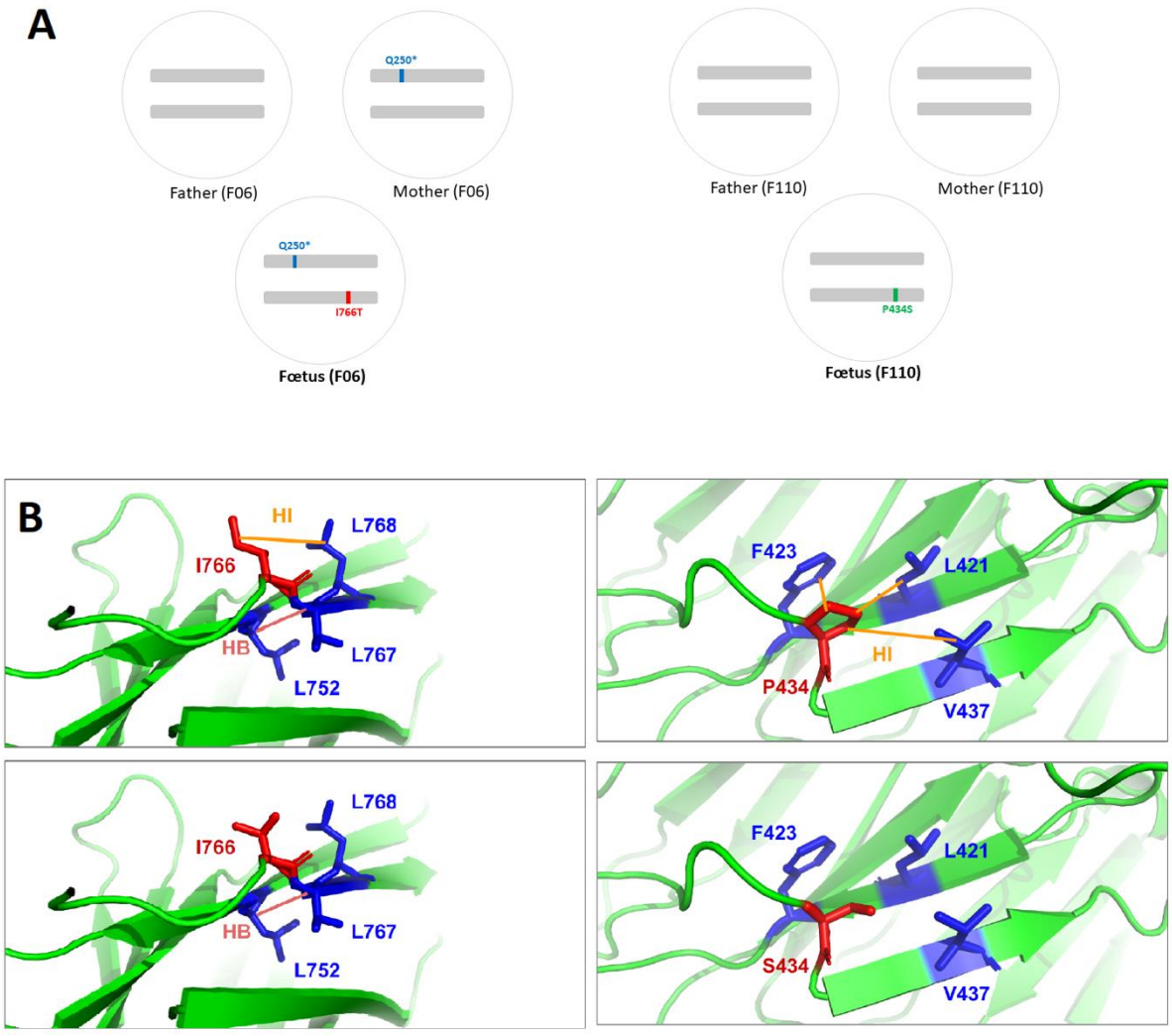


Table S1. Molecular data of the six previously published fetuses

Fetus	Gene	HGVS nomenclature	Protein change	Status	ACMG class	Reference
F03	*MPL*	NM_005373.2:c.305G>C	p.R102P	cHtz	5	Coste et al., 2022a
F03	*MPL*	NM_005373.2:c.1609C>T	p.R537W	cHtz	4	Coste et al., 2022a
F19	*MECOM*	NM_001105077.3:c.1268del	p.S423Lfs*6	Htz	5	Coste et al., 2022a
F92	*ESAM*	NM_138961.2:c.35T>A	p.L12*	Hom	5	Lecca et al., 2023
F54	*PDHA1*	NM_000284.4:c.705_706del	p.R235Sfs*6	Htz	5	Coste et al., 2022b
F49	*PDHA1*	NM_000284.4:c.355C>T	p.R119W	Htz	5	Coste et al., 2022b
F95	*PDHA1*	NM_000284.4:c.910C>T	p.R304*	Htz	5	Coste et al., 2022b

cHtz: compound heterozygous; Htz: heterozygous; Hom: homozygous

Table S2. Results of the consanguinity analysis using exome-sequencing data of the 113 fetuses

A/ Consanguineous fetuses according to clinical chart and results of the analysis on exome data

Fetus ID	Ancestry prediction	Pedigree consanguinity coefficient	Consanguinity analysis on exome data		
			Quality score (%)	Estimated coefficient Fmedian (SE)	Consanguinity inference
F01	African	NA	100	0.066 (0.047-0.088)	True
F81	African	NA	89	0.0068 (0-0.038)	True
F58	European	0.0625	100	0.125 (0.100-0.143)	True
F92	European	0.0078	100	0.062 (0.043-0.079)	True
F94	European	0.031	100	0	False
F101	South Asian	NA	82	0	False

NA: consanguinity reported in clinical chart but pedigree not available to calculate the coefficient. SE: standard error.

B/ Homozygous by descent (HBD) regions in the consanguineous fetuses

Fetus	Chromosome	Start position	End position	Size (kb)
F01	1	10949646	13369564	2419.918
F01	2	88526447	115161724	26635.277
F01	3	140459643	151881605	11421.962
F01	4	374138	426899	52.761
F01	4	786498	959462	172.964
F01	4	991584	1002902	11.318
F01	4	1170590	1380907	210.317
F01	4	1730877	4415824	2684.947
F01	4	177353411	183683913	6330.502
F01	6	161716115	170584376	8868.261
F01	7	151076720	152344249	1267.529
F01	9	4586919	6329888	1742.969
F01	9	124415457	129322342	4906.885
F01	11	193112	7651868	7458.756
F01	11	134194645	134224505	29.86
F01	14	33578235	45195518	11617.283
F01	20	44750623	46118303	1367.68
F01	20	61718449	62656375	937.926
F81	9	110090189	112220183	2129.994
F58	2	111766970	156549784	44782.814
F58	2	235852770	241871857	6019.087
F58	5	173085	840227	667.142
F58	5	1076246	1216660	140.414

F58	5	1244310	1593211	348.901
F58	5	1801575	1882674	81.099
F58	5	9629417	16783328	7153.911
F58	6	12749523	20212238	7462.715
F58	6	157084796	165280135	8195.339
F58	7	143936064	151806933	7870.869
F58	7	154880952	155070911	189.959
F58	8	138877729	145003654	6125.925
F58	9	32408498	39132977	6724.479
F58	10	30700259	54664208	23963.949
F58	10	119161076	128122931	8961.855
F58	11	99819753	108593482	8773.729
F58	11	117505698	126521402	9015.704
F58	12	24584127	53945268	29361.141
F58	13	26050797	33835466	7784.669
F58	13	94382495	102735870	8353.375
F58	14	54987401	60457065	5469.664
F58	16	19030273	23701127	4670.854
F58	16	54933184	64948568	10015.384
F58	16	68252524	68353438	100.914
F58	18	56691831	57147794	455.963
F58	21	25592788	41167642	15574.854
F92	2	6865861	10210473	3344.612
F92	3	8767498	16409365	7641.867
F92	8	69659078	88069502	18410.424
F92	8	127416393	145003654	17587.261
F92	9	135770427	138175883	2405.456
F92	11	102713876	102771985	58.109
F92	11	123610273	128781998	5171.725
F92	20	36244679	50614796	14370.117
F92	22	16592176	20319733	3727.557

Table S3. Rare LoF and missense ***de novo*** **variants** (MAF<1/1000) identified in the 35 trios

Fetus ID	Gene	Variant class	RefSeq	Nucleotide change	Amino acid change	Allele frequency (gnomAD total)	PolyPhen-2	Recurrence in non-trio fetuses	Cases/Controls p-value*
F16	*CHL1*	Missense	NM_006614.3	c.3260C>G	A1087G	Abs	Benign	N/D	N/A
F89	*CSNK2B*	Missense	NM_001320.6	c.610G>A	A204T	0.000012	Benign	N/D	N/A
F34	*DCAF5*	Frameshift	NM_003861.2	c.1433del	D478Afs*17	Abs	N/A	No	N/A
F97	*DDX10*	Missense	NM_004398.3	c.263A>C	Q88P	Abs	Benign	N/D	N/A
F94	*DSG4*	Stop gained	NM_177986.4	c.854C>A	S285*	Abs	N/A	N/D	N/A
F108	*EXD3*	Missense	NM_017820.4	c.2015G>A	G672E	0.0000098	Possibly damaging	Yes	0.0039
F12	*F10*	Missense	NM_000504.3	c.202C>T	R68C	Abs	Probably damaging	No	N/A
F81	*FAM240A*	Missense	NM_001195442.1	c.28T>C	F10L	Abs	Benign	N/D	N/A
F16	*GABRD*	Missense	NM_000815.4	c.1120C>G	L374V	Abs	Benign	N/D	N/A
F96	*GLIS2*	Missense	NM_001318918.1	c.560A>C	D187A	Abs	Benign	N/D	N/A
F80	*HMCN2*	Missense	NM_001291815.1	c.4036A>C	I1346L	Abs	Benign	N/D	N/A
F06	*ITGA2B*	Missense	NM_000419.4	c.2297T>C	I766T	0.000004	Benign	Yes	0.0070
F106	*ITGA2B*	Missense	NM_000419.4	c.1300C>T	P434S	Abs	Probably damaging	Yes	0.0070
F94	*KIRREL3*	Missense	NM_032531.3	c.1741A>G	I581V	Abs	Possibly damaging	Yes	0.19
F94	*LRTM1*	Missense	NM_020678.3	c.65C>T	P22L	0.00002	Possibly damaging	No	N/A
F30	*METTL13*	Missense	NM_015935.4	c.1276G>A	E426K	Abs	Probably damaging	Yes	0.086
F80	*MGME1*	Missense	NM_001310339.1	c.514T>C	S172P	Abs	Benign	N/D	N/A
F92	*MYO9B*	Missense	NM_004145.3	c.1150C>G	L384V	Abs	Probably damaging	Yes	0.66
F108	*NKX3-1*	Missense	NM_006167.3	c.673T>C	C225R	Abs	Probably damaging	No	N/A
F103	*OPA1*	Missense	NM_130837.2	c.254G>A	R85H	0.00023	Benign	N/D	N/A
F01	*PABPC1L*	Frameshift	NM_001372179.1	c.1206del	Q402Hfs*25	Abs	N/A	N/D	N/A
F105	*PGM1*	Missense	NM_002633.2	c.253C>T	R85C	0.00012	Benign	N/D	N/A
F105	*PLTP*	Missense	NM_006227.3	c.1115G>A	R372H	0.00036	Probably damaging	Yes	0.19
F35	*PLXNA2*	Frameshift	NM_025179.3	c.284_287del	Y95Cfs*76	Abs	N/A	N/D	N/A

F113	*RCOR2*	Missense	NM_173587.3	c.1105T>C	F369L	Abs	Probably damaging	No	N/A
F26	*RNF213*	Missense	NM_001256071.2	c.5869G>A	E1957K	0.000012	Benign	N/D	N/A
F106	*RUFY1*	Missense	NM_025158.4	c.298A>C	T100P	Abs	Benign	N/D	N/A
F80	*SCN1A*	Missense	NM_001165963.2	c.1499G>A	R500Q	0.00012	Benign	N/D	N/A
F111	*SLC23A1*	Missense	NM_005847.4	c.871G>A	D291N	0.000016	Probably damaging	No	N/A
F109	*SMC3*	Missense	NM_005445.3	c.1369G>A	E457K	Abs	Possibly damaging	No	N/A
F103	*TBC1D13*	Missense	NM_018201.4	c.370G>A	D124N	0.000060	Probably damaging	No	N/A
F17	*TECTA*	Missense	NM_005422.2	c.5986A>G	I1996V	Abs	Probably damaging	Yes	0.040
F103	*TP53BP2*	Missense	NM_001031685.2	c.2297C>T	A766V	Abs	Benign	N/D	N/A
F108	*TPM2*	Missense	NM_001301227.1	c.566G>A	R189Q	0.000016	Benign	N/D	N/A
F17	*TRA2A*	Frameshift	NM_013293.4	c.283del	R95Gfs*51	Abs	N/A	N/D	N/A
F113	*TRIO*	Missense	NM_007118.3	c.6886G>A	G2296S	Abs	Benign	N/D	N/A
F80	*TUBD1*	Missense	NM_016261.4	c.678T>A	D226E	Abs	Benign	N/D	N/A
F89	*USP20*	Missense	NM_001110303.3	c.1559T>G	I520S	Abs	Benign	N/D	N/A
F109	*ZFYVE28*	Missense	NM_020972.2	c.1544C>T	A515V	Abs	Possibly damaging	Yes	0.00067
F89	*ZNF541*	Missense	NM_001277075.2	c.3079A>C	I1027L	Abs	Benign	N/D	N/A
F35	*ZNF560*	Missense	NM_152476.2	c.2081G>A	R694Q	0.00013	Possibly damaging	Yes	0.08

Abs: absent; N/A: data not available; N/D: data not determined; * Analysis performed for missense *de novo* variants showing recurrence in non-trios fetuses.

Table S4. Rare and predicted damaging **homozygous variants** identified in the 113 proband fetuses

Fetus ID	Inbred	Gene	Variant class	RefSeq	Nucleotide change	Amino acid change	Allele frequency in gnomAD v2.1 / Number of homozygous variants in gnomAD	Total number of homozygous Lof variants at gene level in gnomAD v2.1
F90	No	*AHNAK2*	Missense	NM_001350929.1	c.4156T>A	L1386M	0.00002942 / 1	50
F55	No	*ASCL3*	Missense	NM_020646.2	c.374T>C	L125S	0 0.00002783 / 0	1
F58	Yes	*ASIC4*	Missense	NM_018674.5	c.719C>T	P240L	00.00001433 / 0	0
F16	No	*ATF5*	Missense	NM_001193646.1	c.785A>G	Q262R	0.001363 / 4	0
F58	Yes	*ATP5PO*	Splice region	NM_001697.2	c.87+3A>G	p.?	Abs	0
F63	No	*CALHM6*	Missense	NM_001010919.2	c.839T>C	V280A	0.001969 / 3	414
F70	No	*CAPG*	Frameshift	NM_001747.4	c.332del	P111X	0.0008883 / 0	2
F58	Yes	*CCDC102A*	Missense	NM_033212.3	c.1486G>A	E496K	0.00001234 / 0	0
F63	No	*CCDC168*	Missense	NM_001146197.2	c.10210G>A	E3404K	0.0002449 / 0	1
F14	No	*CCDC27*	Missense	NM_152492.2	c.1208T>C	L403P	0.001044 / 1	1
F109	No	*CDK11A*	Missense	NM_001313896.1	c.286A>C	K96Q	0.0002244 / 0	10
F16	No	*COL11A2*	Missense	NM_080679.2	c.1865G>A	R622Q	0.001028 / 1	0
F32	No	*COQ2*	Frameshift	NM_001358921.2	c.1047del	N351X	0.00001217 / 0	0
F108	No	*CRELD1*	Missense	NM_001031717.3	c.320G>A	R107H	0.0004443 / 0	0
F101	Yes	*EBNA1BP2*	Missense	NM_001159936.1	c.889C>T	R297W	0.002384 / 7	1
F92	Yes	*ESAM*	Stop gained	NM_138961.3	c.35T>A	L12*	Abs	0
F34	No	*FGD5*	Missense	NM_001320276.1	c.3238C>T	R1080C	0.0002139 / 0	0
F59	No	*FGD6*	Missense	NM_018351.3	c.3004C>T	R1002C	0.004302 / 4	0
F63	No	*FRK*	Missense	NM_002031.2	c.296A>T	Y99F	0.0009722 / 1	1
F01	Yes	*FTO*	Missense	NM_001080432.2	c.545G>C	G182A	0.001347 / 3	0
F52	No	*GBP3*	Stop gained	NM_018284.3	c.1360C>T	Q454*	0.001432 / 25	36
F97	No	*GOLGA6L2*	Stop gained	NM_001304388.2	c.2206C>T	R736*	0.00001093 / 0	51
F08	No	*GOLGA6L2*	Missense	NM_001304388.1	c.188C>T	S63L	0.00003962 / 0	51
F92	Yes	*GP1BB*	Missense	NM_000407.4	c.119G>A	G40E	0.001439 / 2	0

F51	No	*GPR42*	Missense	NM_001348195.1	c.140C>G	P47R	Abs	2
F25	No	*HAUS5*	Missense	NM_015302.1	c.1694A>G	E565G	0.001763 / 2	0
F90	No	*HLA-DRB1*	Frameshift	NM_002124.3	c.294_295ins	98-99X	Abs	35974
F94	Yes	*HLA-DRB1*	Frameshift	NM_002124.3	c.294_295ins	98-99X	Abs	35974
F90	No	*HLA-DRB1*	Frameshift	NM_002124.3	c.293dup	E98X	0.005478 / 76	35974
F94	Yes	*HLA-DRB1*	Frameshift	NM_002124.3	c.293dup	E98X	0.005478 / 76	35974
F58	Yes	*KMT2D*	Missense	NM_003482.3	c.10256A>G	D3419G	0.001575 / 3	0
F68	No	*KRTAP9-6*	Missense	NM_001277331.1	c.401A>T	Q134L	0.009025 / 394	4
F58	Yes	*LBP*	Missense	NM_004139.4	c.641A>G	Q214R	0.002051 / 1	1
F25	No	*LIN37*	Missense	NM_019104.2	c.389G>A	R130H	0.001813 / 2	0
F58	Yes	*LPA*	Missense	NM_005577.2	c.3764C>T	T1255M	0.00008192 / 0	363
F46	No	*MADCAM1*	Frameshift	NM_130760.2	c.699_700ins15	D234X	0.00001213 / 0	794
F92	Yes	*MAFA*	Frameshift	NM_201589.4	c.277del	A93PfsX27	Abs	0
F69	No	*MRRF*	Missense	NM_001346345.1	c.583A>T	S195C	0.001305 / 2	0
F01	Yes	*MUC4*	Missense	NM_018406.6	c.281C>T	T94M	0.001393 / 0	9497
F112	No	*MUC4*	Missense	NM_018406.6	c.11944G>A	A3982T	0.004509 / 135	9497
F58	Yes	*MYG1*	Missense	NM_021640.3	c.703T>C	Y235H	0.00005303 / 0	0
F01	Yes	*NAXD*	Missense	NM_001242882.2	c.694G>A	D232N	0.00001061 / 0	6
F01	Yes	*NOS3*	Missense	NM_000603.4	c.2561C>T	T854M	0.0002278 / 0	0
F44	No	*NRAP*	Splice site	NM_198060.4	c.994-1G>T	p.?	0.0003399 / 0	3
F27	No	*NUAK2*	Missense	NM_030952.2	c.29C>G	T10S	0.009111 / 0	0
F01	Yes	*NUP188*	Missense	NM_015354.2	c.3839A>G	D1280G	0.001084 / 1	0
F92	Yes	*NUP210*	Missense	NM_024923.3	c.3154G>A	G1052S	0.002722 / 1	1
F44	No	*OR10A2*	Frameshift	NM_001004460.1	c.224del	Q75X	0.0001308 / 0	1
F01	Yes	*OR10A7*	Frameshift	NM_001005280.1	c.500del	P167X	0.001951 / 3	3
F01	Yes	*OR10A7*	Missense	NM_001005280.1	c.497T>G	L166R	0.001951 / 3	3
F01	Yes	*OR10A7*	Missense	NM_001005280.1	c.196C>G	L66V	0.001950 / 3	3
F58	Yes	*OR2F2*	Missense	NM_001004685.1	c.698G>A	G233E	0.00007424 / 0	0
F01	Yes	*OR2T10*	Missense	NM_001004693.1	c.175A>G	M59V	0.001090 / 57	10

F01	Yes	*OR51B4*	Missense	NM_033179.2	c.503C>T	S168F	0.001057 / 1	0
F81	Yes	*OR5P2*	Frameshift	NM_153444.1	c.851del	L284X	0.001706 / 81	104
F92	Yes	*OR6X1*	Missense	NM_001005188.1	c.7A>C	N3H	0.0008495 / 1	0
F21	No	*PABPC1L*	Missense	NM_001124756.2	c.200G>A	R67Q	0.001259 / 1	0
F38	No	*PLEC*	Missense	NM_000445.4	c.6617G>A	R2206Q	0.0004425 / 0	0
F70	No	*PLPPR3*	Missense	NM_001270366.1	c.208C>T	L70F	0.0009996 / 1	2
F105	No	*PRAMEF4*	Stop gained	NM_001009611.4	c.653G>A	W218*	0.0006206 / 11	2054
F84	No	*RAX2*	Splice site	NM_001319074.1	c.-130-3_-130-2dup	p.?	0.04092 / 110	0
F58	Yes	*RET*	Missense	NM_001355216.1	c.2297C>T	A766V	0.00002386 / 0	0
F70	No	*RETSAT*	Stop gained	NM_017750.3	c.1458_1459del	Y486*	0.001181 / 1	56
F20	No	*RHBG*	Missense	NM_001256395.1	c.542C>T	T181M	0.002459 / 2	15739
F35	No	*RNPEP*	stop_gained	NM_020216.4	c.805G>T	E269*	0.0001096 / 0	0
F41	No	*SAMSN1*	Missense	NM_001256370.1	c.123G>T	K41N	0.001853 / 2	0
F92	Yes	*SCRIB*	Missense	NM_015356.4	c.661C>T	R221C	0.00001615 / 0	0
F34	No	*SFRP5*	Frameshift	NM_003015.3	c.20_21insG	V10X	0.0001809 / 0	2
F03	No	*SIGLEC14*	Missense	NM_001098612.1	c.1091C>T	A364V	0.001641 / 65	6
F52	No	*SLC5A11*	Missense	NM_001258411.2	c.766G>A	V256M	0.0007192 / 2	0
F73	No	*SPATA31C1*	Missense	NM_001145124.1	c.3069T>G	F1023L	0.001440 / 76	0
F01	No	*USP45*	Stop gained	NM_001080481.1	c.658G>T	E220*	0.0006362 / 1	27
F58	Yes	*ZNF251*	Missense	NM_138367.1	c.1586C>G	S529C	0.0001422 / 0	0

Lof: loss of function; Abs: absent

Table S5. Rare and predicted damaging **compound heterozygous variants** identified in the trios and families

Fetus ID	Gene	Variant class	RefSeq	Nucleotide change	Amino acid change	Inheritance	Allele frequency in gnomAD v2.1 / Number of homozygous variants in gnomAD	Total number of homozygous Lof variants at gene level in gnomAD v2.1
F81	*ABCC10*	Missense	NM_001198934.1	c.967G>A	G323S	Father	0.00036 / 0	1
F81	*ABCC10*	Missense	NM_001198934.1	c.2776A>C	T926P	Mother	0.00012 / 0	1
F104	*ADRA2A*	Frameshift	NM_000681.3	c.1369_1372del	L456X	Father	Abs	0
F104	*ADRA2A*	Missense	NM_000681.3	c.377C>T	A126V	Mother	0.000018 / 0	0
F36	*ATP2C2*	Missense	NM_001286527.2	c.493C>A	L165M	Father	Abs	23
F36	*ATP2C2*	Missense	NM_001286527.2	c.796G>A	E266K	Mother	0.000024 / 0	23
F36	*ATP2C2*	Missense	NM_001286527.2	c.839A>T	E280V	Mother	Abs	23
F93	*CMYA5*	Stop gained	NM_153610.4	c.10789A>T	K3597*	Father	0.00081 / 0	2
F93	*CMYA5*	Missense	NM_153610.4	c.5412G>C	Q1804H	Mother	0.00014 / 0	2
F93	*DNAH7*	Missense	NM_018897.2	c.9236T>G	F3079C	Father	0.00026 / 0	14
F93	*DNAH7*	Missense	NM_018897.2	c.3179A>G	Y1060C	Mother	0.000025 / 0	14
F108	*ELN*	Missense	NM_000501.3	c.493G>T	V165L	Father	0.0016 / 4	0
F108	*ELN*	Missense	NM_000501.3	c.1388A>G	K463R	Mother	0.0016 / 2	0
F108	*EPPK1*	Missense	NM_031308.3	c.6052G>A	D2018N	Mother	0.00095 / 1	17
F108	*EPPK1*	Missense	NM_031308.3	c.3058G>A	V1020M	Father	Abs	17
F34	*EPPK1*	Stop gained	NM_031308.3	c.6535C>T	R2179*	Father	0.00009 / 0	17
F34	*EPPK1*	Frameshift	NM_031308.3	c.5704del	V1902X	Mother	0.0005 / 0	17
F34	*EPPK1*	Missense	NM_031308.3	c.820C>T	R274C	Father	0.00093 / 2	17
F97	*HEATR5A*	Missense	NM_015473.3	c.4544C>T	T1515M	Mother	0.00035 / 1	0
F97	*HEATR5A*	Missense	NM_015473.3	c.3590A>G	E1197G	Father	0.00077 / 0	0
F97	*HEATR5A*	Missense	NM_015473.3	c.1495A>C	K499Q	Father	0.00031 / 0	0
F34	*HELZ2*	Missense	NM_001037335.2	c.5288A>T	E1763V	Father	0.00024 / 0	1
F34	*HELZ2*	Missense	NM_001037335.2	c.3382C>T	R1128W	Mother	0.00018 / 0	1
F34	*HMCN1*	Missense	NM_031935.2	c.1480G>A	E494K	Father	0.00034 / 1	0

F34	*HMCN1*	Missense	NM_031935.2	c.11042A>G	N3681S	Mother	0.00013 / 0	0
F101	*IGSF10*	Missense	NM_001178145.1	c.845C>A	S282Y	Mother	Abs	2
F101	*IGSF10*	Missense	NM_178822.4	c.3025C>T	R1009C	Father	0.0027 / 17	2
F81	*LRP8*	Missense	NM_001018054.2	c.2393C>G	P798R	Father	0.000016 / 0	0
F81	*LRP8*	Missense	NM_001018054.2	c.1415G>A	R472H	Mother	0.000042 / 0	0
F03	*MPL*	Missense	NM_005373.2	c.305G>C	R102P	Father	0.00038 / 0	0
F03	*MPL*	Missense	NM_005373.2	c.1609C>T	R537W	Mother	0.000025 / 0	0
F108	*MUC16*	Missense	NM_024690.2	c.35663C>G	T11888R	Mother	0.0016 / 3	3
F108	*MUC16*	Missense	NM_024690.2	c.4709C>T	S1570L	Father	Abs	3
F111	*MUC19*	Missense	NM_173600.2	c.6896C>A	T2299K	Mother	0.0008 / 0	60237
F111	*MUC19*	Missense	NM_173600.2	c.17855G>T	G5952V	Father	0.0008 / 0	60237
F104	*MUC19*	Splice site	NM_173600.2	c.4018-2A>G	p.?	Mother	Abs	60237
F104	*MUC19*	Missense	NM_173600.2	c.17048G>C	G5683A	Father	0.001 / 0	60237
F34	*MUC19*	Missense	NM_173600.2	c.740G>C	G247A	Father	0.0014 / 3	60237
F34	*MUC19*	Missense	NM_173600.2	c.5482T>G	S1828A	Mother	0.0012 / 1	60237
F34	*MUC19*	Missense	NM_173600.2	c.23164C>T	P7722S	Mother	0.0016 / 0	60237
F81	*MUC4*	Missense	NM_004532.5	c.1181G>A	R394H	Father	0.0046 / 21	9497
F81	*MUC4*	Missense	NM_018406.6	c.2774C>A	A925E	Mother	0.00094 / 1	9497
F81	*MUC4*	Missense	NM_018406.6	c.1375G>A	A459T	Mother	0.00096 / 1	9497
F108	*MUC5AC*	Missense	NM_001304359.1	c.3298G>A	G1100S	Mother	Abs	0
F108	*MUC5AC*	Missense	NM_001304359.1	c.3523G>A	E1175K	Father	Abs	0
F108	*MUC5AC*	Missense	NM_001304359.1	c.4348C>G	P1450A	Mother	Abs	0
F108	*MUC5AC*	Missense	NM_001304359.1	c.6971T>C	I2324T	Mother	Abs	0
F108	*MUC5AC*	Missense	NM_001304359.1	c.6986C>T	T2329I	Mother	Abs	0
F108	*MUC5AC*	Missense	NM_001304359.1	c.7358G>T	S2453I	Mother	Abs	0
F108	*MUC5AC*	Missense	NM_001304359.1	c.8618C>A	T2873K	Mother	Abs	0
F108	*MUC5AC*	Missense	NM_001304359.1	c.9674G>T	R3225L	Father	Abs	0
F107	*MYO3B*	Missense	NM_001083615.3	c.1603T>C	Y535H	Mother	0.000026 / 0	0
F107	*MYO3B*	Missense	NM_001083615.3	c.2968C>T	R990C	Father	0.0018 / 1	0

F102	*NEB*	Missense	NM_001164507.1	c.18928A>G	I6310V	Father	0.00001 / 0	2
F102	*NEB*	Missense	NM_001164507.1	c.2132A>G	E711G	Mother	Abs	2
F36	*NPHP4*	Missense	NM_001291593.1	c.1426G>A	E476K	Mother	0.00045 / 0	88272
F36	*NPHP4*	Missense	NM_001291593.1	c.1412C>T	T471M	Mother	0.000046 / 0	88272
F36	*NPHP4*	Missense	NM_001291593.1	c.1309C>T	R437W	Father	0.000012 / 0	88272
F36	*NPHP4*	Missense	NM_001291593.1	c.664C>T	R222W	Mother	0.00043 / 0	88272
F36	*OBSCN*	Missense	NM_001098623.2	c.8137G>A	A2713T	Father	0.001 / 0	960
F36	*OBSCN*	Missense	NM_001098623.2	c.13453G>A	D4485N	Father	0.00017 / 0	960
F36	*OBSCN*	Missense	NM_001098623.2	c.19859A>C	Q6620P	Mother	0.000039 / 0	960
F81	*PCLO*	Missense	NM_014510.2	c.5632A>G	I1878V	Mother	0.00077 / 2	0
F81	*PCLO*	Missense	NM_014510.2	c.3862G>T	G1288W	Father	0.00026 / 0	0
F97	*PKD1L1*	Missense	NM_138295.4	c.7876T>C	C2626R	Father	0.00034 / 0	2
F97	*PKD1L1*	Missense	NM_138295.4	c.2008G>A	V670M	Mother	0.00027 / 0	2
F12	*PLOD3*	Missense	NM_001084.4	c.1466C>T	P489L	Father	0.00051 / 3	0
F12	*PLOD3*	Missense	NM_001084.4	c.589C>T	R197W	Mother	0.000012 / 0	0
F12	*PROC*	Frameshift	NM_000312.3	c.644_651del	G215X	Mother	Abs	0
F12	*PROC*	Missense	NM_000312.3	c.1163C>A	A388E	Father	Abs	0
F34	*RAB11FIP5*	Missense	NM_015470.2	c.1076C>T	S359L	Father	0.00023 / 0	0
F34	*RAB11FIP5*	Missense	NM_015470.2	c.235C>G	P79A	Mother	0.0014 / 2	0
F112	*SCRIB*	Missense	NM_015356.4	c.4576G>C	G1526R	Father	0.000009 / 0	0
F112	*SCRIB*	Missense	NM_015356.4	c.61C>T	H21Y	Mother	Abs	0
F112	*SHKBP1*	Missense	NM_138392.3	c.416G>A	R139Q	Father	0.00085 / 2	0
F112	*SHKBP1*	Missense	NM_138392.3	c.1883C>A	S628Y	Mother	0.0015 / 1	0
F108	*SLFN12L*	Frameshift	NM_001195790.1	c.1716_1719del	L572X	Mother	0.00016 / 0	335
F108	*SLFN12L*	Frameshift	NM_001195790.1	c.508dup	T170X	Father	Abs	335
F108	*SLFN12L*	Missense	NM_001195790.1	c.118A>C	T40P	Mother	0.000039 / 0	335
F108	*SSPO*	Missense	NM_198455.2	c.2789A>G	Y930C	Mother	0.000022 / 0	25072
F108	*SSPO*	Missense	NM_198455.2	c.11264G>C	C3755S	Father	0.00023 / 0	25072
F105	*TBC1D3J*	Stop gained	NM_001355571.1	c.19C>T	R7*	Mother	0.000026 / 0	0

F105	*TBC1D3J*	Missense	NM_001355571.1	c.1310C>T	S437L	Father	0.00096 / 0	0
F34	*TJP2*	Missense	NM_001170414.2	c.2567A>G	Q856R	Father	0.00086 / 1	0
F34	*TJP2*	Missense	NM_001170414.2	c.2657C>T	A886V	Mother	0.00013 / 0	0
F16	*TNC*	Missense	NM_002160.3	c.5351G>A	G1784D	Mother	0.0001 / 0	0
F16	*TNC*	Missense	NM_002160.3	c.3437G>A	G1146E	Father	Abs	0
F108	*TNXB*	Missense	NM_019105.6	c.5746A>G	T1916A	Mother	0.0021 / 3	1
F108	*TNXB*	Missense	NM_019105.6	c.113G>A	R38Q	Father	0.0026 / 9	1
F104	*TTC28*	Missense	NM_001145418.1	c.5437C>T	R1813W	Mother	0.0027 / 3	0
F104	*TTC28*	Frameshift	NM_001145418.1	c.4004dup	N1335X	Father	Abs	0
F101	*TTN*	Missense	NM_001256850.1	c.98224G>C	E32742Q	Mother	0.0038 / 12	5
F101	*TTN*	Missense	NM_001256850.1	c.71999G>A	R24000H	Father	0.00063 / 4	5
F101	*TTN*	Missense	NM_001256850.1	c.68521A>G	T22841A	Father	0.000004 / 0	5
F12	*ZNF687*	Missense	NM_001304763.1	c.3425T>C	L1142S	Mother	Abs	0
F12	*ZNF687*	Missense	NM_001304763.1	c.3460C>T	R1154C	Father	0.000018 / 0	0
F36	*ZNF865*	Missense	NM_001195605.1	c.568C>T	P190S	Mother	0.00035 / 0	0
F36	*ZNF865*	Missense	NM_001195605.1	c.2338G>T	G780C	Father	0.000014 / 0	0

Lof: loss of function; Abs: absent

Table S6. Rare and predicted damaging **hemizygous variants** identified in the 58 male probands

Fetus	Gene	Variant class	RefSeq	Nucleotide change	Amino acid change	Allele frequency in gnomAD v2.1 / Number of hemizygous variants in gnomAD	Total number of hemizygous Lof variants at gene level in gnomAD v2.1
13	*ACTRT1*	Frameshift	NM_138289.3	c.724del	R242X	Abs	140
100	*ADGRG4*	Missense	NM_153834.3	c.1866G>C	L622F	Abs	96
57	*ARHGAP4*	Missense	NM_001164741.1	c.802C>T	R268W	Abs	11
F11	*ARHGAP4*	Missense	NM_001164741.1	c.238C>T	R80C	Abs	11
103	*ARSH*	Missense	NM_001011719.1	c.1444C>G	L482V	Abs	29
115	*ATP11C*	Missense	NM_001010986.2	c.71C>T	T24I	Abs	5
100	*ATP6AP1*	Missense	NM_001183.5	c.344G>A	S115N	0.00007800 / 4	0
41	*BCOR*	Missense	NM_001123385.2	c.4813G>C	V1605L	Abs	0
57	*BTK*	Missense	NM_000061.2	c.1510G>A	D504N	Abs	0
99	*CHST7*	Missense	NM_019886.3	c.52T>A	L18M	Abs	2
115	*CT47B1*	Missense	NM_001145718.1	c.562G>A	G188S	Abs	7
F15	*DMD*	Missense	NM_000109.3	c.3695T>G	L1232R	0.00002446 / 3	7
F42	*DOCK11*	Missense	NM_144658.3	c.4286A>G	H1429R	Abs	17
F15	*DRP2*	Missense	NM_001171184.1	c.2506G>A	E836K	0.00001096 / 1	17
50	*FAAH2*	Missense	NM_001353840.1	c.533A>T	E178V	0.00005457 / 1	138
F02	*FAM47B*	Stop gained	NM_152631.2	c.331C>T	Q111*	0.000005457 / 1	30
23	*FLNA*	Missense	NM_001110556.1	c.6728C>G	A2243G	0.00002801 / 1	1
17	*FRMD7*	Missense	NM_001306193.1	c.1438G>A	V480I	Abs	6
41	*FRMPD3*	Missense	NM_032428.2	c.680A>G	Y227C	Abs	21
116	*GATA1*	Missense	NM_002049.3	c.343G>T	D115Y	0.00001091 / 1	0
F39	*GNL3L*	Missense	NM_001184819.1	c.760C>T	R254C	0.0001088 / 5	1
40	*GUCY2F*	Missense	NM_001522.2	c.3196T>C	F1066L	0.00001095 / 1	63
100	*IL13RA1*	Missense	NM_001560.2	c.1207G>A	D403N	0.00009916 / 7	8
12	*KDM6A*	Missense	NM_001291415.1	c.2614T>G	S872A	0.000005474 / 1	0

23	*MAGEC1*	Frameshift	NM_005462.4	c.802_803ins	L268X	Abs	548
F39	*MAGEC3*	Missense	NM_177456.2	c.92C>G	S31C	0.00001464 / 1	1064
43	*MAMLD1*	Missense	NM_001177465.2	c.2573C>T	P858L	0.00007845 / 0	4
F41	*MED14OS*	Frameshift	NM_001289773.1	c.177dup	R60X	Abs	1
13	*MOSPD2*	Missense	NM_001177475.1	c.968C>T	S323L	Abs	0
105	*MTMR8*	Missense	NM_017677.3	c.725G>C	R242T	Abs	31
12	*NKAP*	Missense	NM_024528.3	c.768G>T	K256N	0.00001701 / 0	0
103	*PGK1*	Missense	NM_000291.3	c.794T>C	I265T	Abs	0
F16	*PLAC1*	Missense	NM_001316887.1	c.314C>T	T105M	0.00002181 / 2	0
26	*POF1B*	Missense	NM_001307940.1	c.797G>A	R266H	0.00004528 / 4	9
4	*SH3KBP1*	Missense	NM_001024666.2	c.838G>A	V280M	Abs	1
29	*TBL1X*	Missense	NM_001139466.1	c.1726C>T	R576W	0.000009919 / 1	2
4	*TENM1*	Missense	NM_001163278.1	c.593C>T	P198L	0.00004957 / 3	4
17	*TMEM255A*	Splice site	NM_001104544.1	c.423+1G>C	p.?	0.00001090 / 0	4
46	*TSPAN7*	Missense	NM_004615.3	c.262C>G	L88V	Abs	0
F25	*VCX2*	Missense	NM_016378.3	c.380G>C	W127S	Abs	32
22	*ZIC3*	Missense	NM_001330661.1	c.1367T>A	I456N	Abs	1
43	*ZNF185*	Missense	NM_001178106.1	c.260T>C	I87T	0.00003445 / 5	38

Lof: loss of function; Abs: absent

Table S7. Top 10 genes identified using a gene-based burden test comparing qualifying variants between fetal ICH and gnomAD v3.0 cohorts

Rank	Gene	Case count in fetuses (N=113)	Case count in gnomAD v3 (N=71702)	p-value
1	*PDHA1*	3 (2.65%)	5 (0.007%)	2.11E-07
2	*TIE1*	6 (5.31%)	299 (0.417%)	9.64E-06
3	*ZNF775*	4 (3.54%)	149 (0.208%)	0.00010646
4	*ASH1L*	5 (4.42%)	352 (0.491%)	0.00026699
5	*PLB1*	5 (4.42%)	402 (0.561%)	0.00048492
6	*YPEL1*	2 (1.77%)	19 (0.026%)	0.00050535
7	*PPRC1*	4 (3.54%)	240 (0.335%)	0.00062628
8	*PDE8B*	3 (2.65%)	102 (0.142%)	0.00063263
9	*SULF1*	3 (2.65%)	103 (0.144%)	0.00065032
10	*TBL1X*	1 (0.88%)	22 (0.031%)	0.00066213

Table S8. Qualifying variants identified in *TIE1* gene

Fetus ID	Gene	Coding variation (NM_005424.4)	Protein change (NP_005415.1)	gnomAD v2.1 global AF	Polyphen2 (score)	Status
F38	*TIE1*	c.1382G>A	p.Arg461His	0.000028	Probably damaging (0.961)	Heterozygous
F47	*TIE1*	c.207C>G	p.Ile69Met	0.000037	Probably damaging (0.997)	Heterozygous
F55	*TIE1*	c.1297G>A	p.Gly433Ser	0.000030	Probably damaging (1)	Heterozygous
F101	*TIE1*	c.3053G>A	p.Arg1018His	0.0000040	Probably damaging (1)	Heterozygous
F103	*TIE1*	c.2462G>A	p.Arg821Gln	0.000012	Probably damaging (0.994)	Heterozygous
F112	*TIE1*	c.3052C>T	p.Arg1018Cys	0.0000040	Probably damaging (1)	Heterozygous

Table S9. Characteristics of the 10 genes identified using a gene-based burden test

Gene	Function	Expression	Human phenotype	Animal models
PDHA1	Subunit of the pyruvate dehydrogenase complex, which is a nuclear-encoded mitochondrial multienzyme complex that catalyzes the overall conversion of pyruvate to acetyl-CoA	Ubiquitous	OMIM : Pyruvate dehydrogenase E1-alpha deficiency (XLD inheritance)	Embryonic lethality in knocked out mouse
TIE1	Tyrosine kinase with a role in angiogenesis and blood vessel stability by inhibiting angiopoietin 1 signaling through the endothelial receptor tyrosine kinase Tie2	Broad	OMIM : Lymphatic malformation 11 (AD inheritance)	Hemorrhagic phenotype in knocked out mice (PMID: 20223757)
ZNF775	Predicted to be Involved in regulation of transcription by RNA polymerase II	Ubiquitous	Unknown	Unknown
ASH1L	Histone methyltransferase member of the trithorax group of transcriptional activators	Ubiquitous	OMIM : Mental retardation (AD inheritance)	No obvious gross morphological abnormalities of individual organs, except that the majority of Ash1L-KO newborns displayed aberrant rib numbers (PMID: 34145365)
PLB1	Phospholipase that catalyzes complete diacylation of phospholipids. May contribute to digestion of dietary phospholipids, glycerolipids and retinoids, facilitating lipid absorption at the brush border	Intestine ++ Skin Whole blood	Gene proposed to contribute to virulence of C. albicans and C. neoformans	Hyperactivity in knocked out mouse
YPEL1	The encoded protein localizes to the centrosome and nucleolus and may play a role in the regulation of cell division	Broad Testis++ Brain+	Unknown	Unknown
PPRC1	Activate mitochondrial biogenesis in part through a direct interaction with nuclear respiratory factor 1	Ubiquitous	Unknown	Homozygous deficient PPRC1 mice failed to form egg cylinders and died after implantation Mice heterozygous for PPRC1 were viable, fertile and indistinguishable from their wild-type littermates (PMID: 22411706)

PDE8B	Cyclic nucleotide phosphodiesterase that catalyzes the hydrolysis of the second messenger cAMP	Thyroid++	OMIM: Pigmented nodular adrenocortical disease Striatal degeneration (AD inheritance)	PDE8B knocked out mice demonstrate an enhancement in contextual fear, spatial memory, performance in an appetitive instrumental conditioning task, motor-coordination, and have an attenuation of age-induced motor coordination decline (PMID: 22925203) Pde8b-/- male mice were infertile and modest adrenal defect (PMID: 33338547)
SULF1	Extracellular heparan sulfate endosulfatase that removes 6-O-sulfate groups from heparan sulfate chains of heparan sulfate proteoglycans	Broad Arteries++ Fibroblasts+	Mesomelia-Synostoses Syndrome Results from deletion of SULF1 and SLCO5A1 Genes at 8q13 (PMID: 20602915)	No abnormal phenotype detected in Sulf1 knocked out mouse (PMID: 17720696) Targeted knockdown of sulf1 by antisense morpholinos in zebrafish resulted in severe vascular patterning and maturation defects (PMID: 23959107)
TBL1X	F-box-like protein involved in the recruitment of the ubiquitin/19S proteasome complex to nuclear receptor-regulated transcription units. Plays an essential role in transcription activation	Ubiquitous	OMIM: Hypothyroidism (XLR inheritance)	Unknown

DISCUSSION

Plusieurs gènes causaux porteurs de variants pathogènes biallélіques ont été identifiés (*COQ2, ESAM, PROC* et *ATP5PO)* ainsi que des gènes considérés comme candidats forts *(DCAF5, EXD3, ITGA2B, NAXD, RNPEP, TIE1 et ZFYVE28*) mais nécessitant des explorations complémentaires. Les autres gènes causaux étaient les gènes *MPL, MECOM* et *PDHA1* qui ont fait l'objet d'articles distincts. Pour presque tous les gènes causaux et candidats identifiés, la récurrence est faible. Ces résultats illustrent l'extrême hétérogénéité génétique des HIC fœtales et l'intérêt des analyses pangénomiques chez les fœtus référés pour HIC. Enfin, ces résultats soulignent également que des variants pathogènes dans des gènes impliqués dans le métabolisme mitochondrial sont responsables de phénotypes pouvant mimer celui rencontré chez les fœtus porteurs de variants pathogènes de *COL4A1/COL4A2.*

Pour 70% des fœtus, l'information sur l'existence ou non d'une consanguinité était disponible sur le compte rendu ou l'arbre généalogique. Sur ces données provenant de l'interrogatoire des parents, six fœtus étaient considérés comme issus d'une union entre apparentés avec, pour 4 d'entre eux, des taux de consanguinité attendus allant de 0,78% à 6,25%. Afin de préciser cette information pour l'ensemble des fœtus, nous avons effectué une étude de la consanguinité sur les données de WES avec un package R (Fantasio). Cette analyse a révélé que quatre fœtus étaient consanguins avec des coefficients de consanguinité F allant de 0,7% à 12,5% (ce dernier présentant plusieurs boucles de consanguinité). Il y avait donc environ 4% de fœtus consanguins dans notre cohorte. Ce chiffre est plus élevé que le pourcentage estimé d'individus consanguins (0.15% ; $p=3.12x10^{-5}$) dans la population générale Française (Serre *et al.*, 2014), et qui conforte l'hypothèse de l'identification de variants dans des gènes causaux transmis selon un mode autosomique récessif. Après calcul du coefficient de consanguinité, l'outil Fantasio nous a fourni la liste des régions HBD qui sont potentiellement impliquées dans le phénotype d'HIC avec un nombre allant de 1 à 26 pour celui avec un plus fort coefficient de consanguinité. Les quatre fœtus consanguins n'avaient aucun segment homozygote en commun. Le fœtus F58 est celui qui présente le plus de régions HBD (Figure 25).

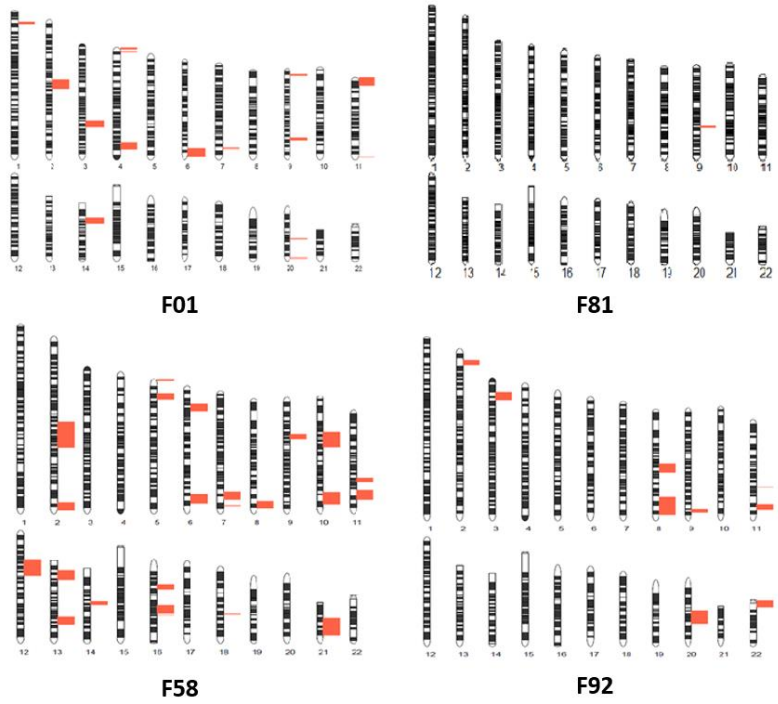


Figure 25. Illustration de la position chromosomique des segments HBD chez les quatre fœtus consanguins selon le package R Fantasio.

Il y avait une corrélation positive entre le nombre de variants homozygotes candidats (variants rares avec MAF < 0,5% et de nature disruptive ou faux-sens) et le taux de consanguinité (Figure 26).

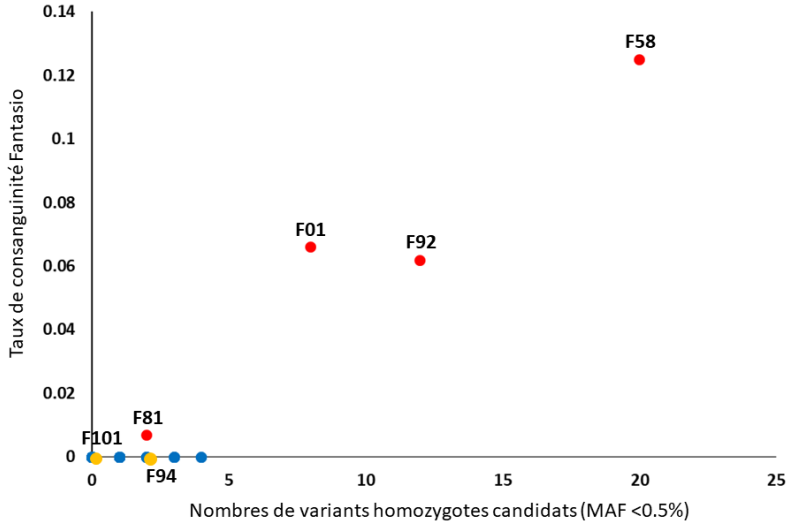


Figure 26. Nombre de variants homozygotes candidats selon le taux de consanguinité estimé avec l'outil Fantasio. En rouge : individus consanguins selon Fantasio ; orange : fœtus avec consanguinité a priori mais non confirmée par Fantasio ; Bleu : fœtus non consanguins.

Le fœtus F58 (le plus consanguin) présente effectivement un nombre important de variants homozygotes candidats (26). Il s'agit du fœtus pour lequel nous avons identifié un variant causal perte de fonction homozygote dans le gène *ESAM* (dans une région HBD).

Par ailleurs, deux fœtus (F94 et F101) étaient dits consanguins selon l'interrogatoire des parents, mais cette consanguinité n'était pas retrouvée sur les données de WES. Ils présentaient un nombre faible de variants candidats homozygotes rares (2 et 0 respectivement) (Tableau 2).

Fœtus	Origine	Nombre de variants homozygotes candidats (MAF <0.5% et nature disruptive ou faux-sens)	Taux de consanguinité attendu	Taux de consanguinité Fantasio	Nombre de sous-cartes	Nombre de segments HBD	Nombre de variants homozygotes candidats dans segments HBD
F01	AFR	8	Nd*	0.066	100	18	3
F58	EUR	20	0.0625	0.125	100	26	15
F81	AFR	2	<0.0625	0.007	89	1	0
F92	EUR	12	0.0078	0.062	100	9	11
F94	EUR	2	0.03125	0	100	Nd	Nd
F101	SAS	0	Nd*	0	82	Nd	Nd

Tableau 2. Résultats de l'analyse de consanguinité avec Fantasio et nombre de variants homozygotes chez les six fœtus à priori consanguins. AFR : Africaine ; EUR : Européenne, SAS ; Asie du Sud. ; Nd : donnée non disponible ; * : notion de consanguinité sans arbre généalogique précis.

Pour ces fœtus F94 et F101, plusieurs hypothèses sont possibles pour expliquer la discordance avec l'analyse sur les données de WES. Soit ces fœtus ne sont pas consanguins soit la qualité des données de WES ne permet pas d'identifier une consanguinité. Cette dernière hypothèse est probable pour le fœtus F94 car il possède un taux de variants hétérozygotes plus élevé que les autres individus (Figure 27), gage d'une moins bonne qualité de l'échantillon d'ADN (Høy Hansen *et al.*, 2022).

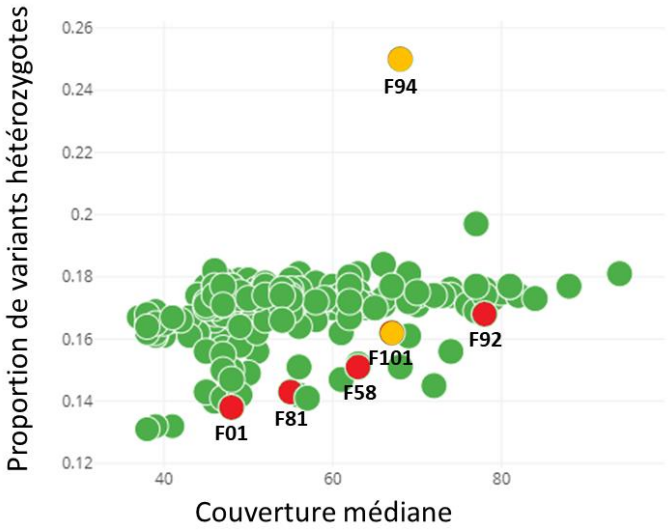


Figure 27. Proportion de variants hétérozygotes chez les fœtus consanguins selon les données de l'interrogatoire. Rouge : fœtus consanguin selon Fantasio ; Orange : consanguinité selon interrogatoire mais non confirmée par Fantasio. On peut remarquer une distribution plus basse de la proportion en variants hétérozygotes chez ces sujets.

Des filtres moins stringents ont été utilisés pour l'analyse des variants homozygotes dans les régions HBD des quatre fœtus consanguins. Nous avons analysé tous les variants homozygotes (incluant les synonymes) avec une fréquence dans les bases de données de polymorphismes inférieure à 1%. Pour les variants synonymes nous avons regardé s'ils pouvaient avoir un effet sur l'épissage à l'aide de l'outil MaxEntScan (Yeo et Burge, 2004). Nous avons aussi recherché la présence de CNV homozygotes dans les segments HBD à l'aide du logiciel HMZDelFinder (résultats de cette analyse en Annexe 3). Pour 3 des 4 fœtus consanguins (selon les données de WES) nous avons identifié un variant homozygote causal ou probablement causal (gènes *ATP5PO*, *ESAM* et *NAXD*). Pour le fœtus restant (F81), aucun variant homozygote candidat ni CNV n'a été détecté, ce fœtus ayant le taux de consanguinité le plus faible (0,007).

Dans la littérature, il y a peu de données sur la génétique des HIC chez le fœtus et celles-ci sont en majorité portées sur *COL4A1/COL4A2* (Cavaliere *et al.*, 2021). De plus, à notre connaissance, une seule étude de cohorte a été rapportée avec analyse de WES dans un contexte d'HIC périnatales (Hausman-Kedem *et al.*, 2021). Dans cette étude, les auteurs avaient utilisé des critères d'inclusion différents des nôtres : pas de criblage préalable des gènes *COL4A1* et *COL4A2* et HIC identifiée en période anténatale ou néonatale. Leur cohorte comprenait 26 cas issus de 25 familles pour lesquels une HIC avait été diagnostiquée en période anténatale pour 9 cas et en période postnatale pour 17 cas. Les auteurs identifièrent des variants pathogènes ou

probablement pathogènes chez 4 patients issus de 3 familles dans les gènes *COL4A1, COL4A2* et *TREX1* (rendement diagnostic de 12%). La causalité de *TREX1* dans les HIC fœtales est toutefois discutable (Wilms *et al.,* 2022). Les auteurs suggèrent également plusieurs autres gènes candidats : *DNAH5, GP1BA, FGA, F7, F11* et *VWF*, gènes pour lesquels ils ont identifié des variants hétérozygotes de signification inconnue (interprétation basée sur la fréquence dans les bases de données de polymorphismes, statut hérité des parents et le mode de transmission de la pathologie associée). Curieusement, parmi ces 6 gènes, 5 sont impliqués dans des troubles de l'hémostase et de la coagulation. Cette voie biologique semble être une voie importante impliquée dans les HIC fœtales puisque nous avons également identifié des variants dans des gènes de l'hémostase et de la coagulation (gènes *MPL, MECOM, PROC* et *ITGA2B).* Par ailleurs, cette étude et la nôtre montrent qu'après analyse de WES il n'y a pas de variants identifiés dans des gènes causaux ou candidats pour de nombreux fœtus avec HIC, ce qui soulève plusieurs hypothèses :

-Un manque de puissance statistique du fait d'une très grande hétérogénéité génétique qui impose la mise en place de collaborations à l'échelle internationale afin d'agrandir la taille des cohortes étudiées et le partage de données pour permettre l'identification de variations récurrentes

-Des limites techniques inhérentes au WES : régions mal couvertes, absence de détection des variants introniques, des variants dans des régions régulatrices ou des variants structuraux

-Un mode de transmission non mendélien et complexe

Très récemment, il a été publié une étude portant sur 85 patients incluant 39 enfants prématurés avec hémorragie périventriculaire, six enfants avec hémorragie périventriculaire anténatale et 40 enfants avec infarctus veineux périventriculaire présumé anténatal (Ilves *et al.,* 2023). Sur des données de WES, les auteurs ont analysé une liste de 160 gènes candidats associés à un AVC (tout âge confondu), une anomalie vasculaire ou une anomalie de l'hémostase / coagulation. Ils conclurent que 11/85 patients (13%) étaient porteurs de variants pathogènes (cinq patients porteurs de variants pathogènes de *COL4A1,* un patient *COL4A2,* un patient *COL5A1,* un patient *NOTCH3* et un patient *MT-TL1* et deux patients avec des variants de prédisposition à la thrombophilie). L'association de variants pathogènes de *NOTCH3* (responsable de la maladie CADASIL), *MT-TL1* (responsable du syndrome de MELAS) ou de *COL5A1* (responsable du syndrome d'Ehlers-Danlos) avec une HIC anténatale ou néonatale n'a jamais été établie à ce jour. La liste des gènes candidats incluait certains gènes que nous avons

identifié dans notre cohorte : *MPL, MECOM* et *PROC* ; et il n'a pas été retrouvé par les auteurs de variants pathogènes ou de signification inconnue dans ces gènes.

Ces résultats illustrent la prépondérance des variations pathogènes de *COL4A1*, la possibilité que des enfants prématurés ayant fait une HIC puissent avoir des variations pathogènes de *COL4A1/COL4A2* et ils illustrent aussi la grande proportion de patient négatifs après criblage moléculaire.

Ce qui est également intéressant dans cet article, c'est que les auteurs ont identifié un patient porteur d'un variant hétérozygote de signification inconnue (p.R1028Q) dans le gène *ITGA2B*, dont le statut de novo/hérité n'était pas connu. Dans notre cohorte, nous avons mis en évidence des variants *de novo* du gène *ITGA2B* chez 2 fœtus index. La probabilité d'observer la survenue au hasard de 2 évènements dans le même gène au sein des 35 trios est très faible (valeur de p ajustée = 0,047). Par ailleurs un de ces fœtus était également porteur d'un codon stop hérité de sa mère présent sur l'autre allèle du gène *ITGA2B* (hétérozygote composite). Ces éléments nous font ainsi penser que ce gène est probablement causal. *ITGA2B* code pour la sous-unité alpha 2b de l'intégrine αIIbβ3, une protéine exprimée sur la plaquette et qui est essentielle à l'agrégation plaquettaire. Les variations pathogènes d'*ITGA2B* sont associées à une maladie récessive rare appelée thrombasthénie de Glanzmann (prévalence < 1/1 000 000), qui se caractérise par des saignements essentiellement cutanéo-muqueux. Les HIC chez ces patients sont exceptionnelles bien que des cas d'HIC fœtales soient connus chez les femmes atteintes de Glanzmann par le développement d'anticorps contre l'intégrine αIIbβ3 des plaquettes du fœtus (Wijemanne *et al.*, 2016). Plusieurs hypothèses ont été émises et décrites dans l'article de synthèse pour ces 2 variants *de novo* (immunisation de la mère contre l'intégrine IIb portant le variant faux de novo ou facteur de risque en combinaison avec un autre événement déclencheur) mais il ne nous a pas été possible de conclure, en l'absence de matériel biologique disponible (plaquettes du fœtus, sérum de la mère), sur l'implication de ces variants.

Enfin, de façon à priori surprenante, nous avons identifié au sein de notre cohorte de fœtus référés pour une suspicion d'HIC, des variants pathogènes ou probablement pathogènes dans des gènes impliqués dans le métabolisme mitochondrial. Nous avons ainsi mis en évidence un fœtus (F32) avec variant disruptif homozygote dans le gène *COQ2*, codant pour une enzyme mitochondriale. Un fœtus consanguin (F58) était porteur d'un variant homozygote pathogène (c.87+3A>G) dans *ATP5PO*. Le variant (c.87+3A>G) a été rapporté précédemment par Ganapathi *et al.* (2022) chez deux enfants index qui présentaient à l'IRM cérébrale une atrophie,

une ventriculomégalie, des kystes périventriculaires, des hypersignaux de la substance blanche, un corps calleux fin et un cavum du septum pellucidum. Un de ces deux enfants présentait également des stigmates d'hémorragies anciennes bilatérales dans les lobes frontal et pariétal. Ces similitudes phénotypiques conduisent actuellement à une demande de criblage des gènes *COL4A1* et *COL4A2*. D'autre part, un fœtus consanguin (F01) était porteur d'un variant homozygote faux-sens probablement pathogène dans le gène *NAXD*. Ce fœtus appartient à une famille pour laquelle un fœtus précédent avait également été interrompu dans un contexte de suspicion d'HIC et qui porte aussi le variant à l'état homozygote (identifié en Sanger car le WES n'était pas possible en raison d'une mauvaise qualité de l'ADN). Et nous avons identifié trois fœtus de sexe féminin (F50, F54, F95) avec des variants pathogènes dans *PDHA1*, responsables d'un déficit en pyruvate déshydrogénase dominant lié à l'X.

Pour comparaison, les différentes anomalies neurologiques chez les fœtus avec un variant causal dans les gènes responsables d'une thrombopénie sévère, d'un déficit en protéine C et d'une maladie mitochondriale sont décrites dans le tableau suivant (Tableau 3).

Dans toutes ces conditions, de nombreux signes associés sont semblables à ceux rencontrés lors de collagénopathies *COL4A1/A2*, ce qui illustre la complexité de l'orientation diagnostique lorsque les signes associés de différentes affections sont similaires.

En conclusion, cet article présente une caractérisation génétique approfondie de la cohorte de fœtus référés pour HIC. Il souligne l'hétérogénéité génétique extrême des HIC fœtales, avec une faible récurrence en variants pour les gènes causaux et candidats identifiés, et met en avant la complexité de l'interprétation phénotypique chez le fœtus et l'importance de l'analyse de l'exome entier pour identifier les causes génétiques sous-jacentes.

Voie biologique	Gène causal	N° Fœtus	Anomalies cérébrales à l'IRM fœtale	Anomalies cérébrales à l'examen neuropathologique
Mégacaryopoïèse	***MPL***	F03	*Non disponible*	Dilatation quadriventriculaire /Nécrose parenchymateuse pariétale droite / Pseudorosettes épendymaires / Foyers hémorragiques intra-parenchymateux péri-ventriculaires (hématies fraiches + pigments hémosidériniques) / Microcalcifications périventriculaires / Lésions de polymicrogyrie pariétales droites (témoins de lésions ischémiques surajoutés) / Lésions hémorragiques anciennes et récentes au niveau des méninges
	MECOM	F19	*IRM non faite (MFIU découverte à l'échographie)*	Hématome sous-arachnoïdien englobant le pôle temporal / Dilatation ventriculaire / Pétéchies intraparenchymateuses diffuses centrées par des capillaires / Signes ischémiques organisés (proliférations capillaires et macrophagiques) au niveau thalamus et noyaux gris centraux
Jonctions serrées	***ESAM***	F92	Importante dilatation ventriculaire / Hémorragies intéressant 3ème ventricule, ventricules latéraux et probablement le corps calleux (mal identifié) / Parois ventriculaires des ventricules latéraux irrégulières / Atteinte hémorragique frontale gauche / Kyste et hémosidérose dans fosse postérieure	*Examen non souhaité par les parents*
Inhibition de la coagulation	***PROC***	F12	Dilatation asymétrique des ventricules latéraux avec lésions parenchymateuses associées ischémo-hémorragiques hémisphère gauche / Lésions hémorragiques pures au niveau hémisphère droit / Kyste arachnoïdien dans fosse postérieure	Ventriculomégalie latérale asymétrique / Hémorragies intraparenchymateuses prédominantes à gauche / Foyers d'ischémie corticale et sous-corticale associés / Nombreuses thromboses organisées des vaisseaux dans la substance blanche / Face interne des ventricules avec infiltration de sidérophages évoquant une origine hémorragique de la ventriculomégalie
Métabolisme mitochondriale	***PDHA1***	F49	Dilatation ventriculaire / Microkystes de germinolyse / Anomalie de signal au niveau de la substance blanche profonde	Microcéphalie/ Ventriculomégalie majeure / Kystes périventriculaires / Foyers de porencéphalie et remaniement diffus des ventricules
		F54	Ventriculomégalie bilatérale / Retard de gyration / Agénésie de la partie postérieure du corps calleux / Formations kystiques sous-épendymaires frontales et temporales	Microcéphalie/ Retard de gyration / Agénésie partiel du corps calleux / Colpocéphalie / Pseudokystes sous épendymaires en pariétal / Spongiose périventriculaire
		F95	Agénésie corps calleux / Ventriculomégalie supra-tentorielles à berges irrégulières / Retard de gyration / Lésions microkystiques / Hypoplasie cérébelleuse globale / Séquences T2 écho de gradient sans anomalie en faveur de stigmates hémorragiques	*Examen non souhaité par les parents*
	COQ2	F32	Lissencéphalie avec anomalie de la substance blanche	Ventriculomégalie modérée / Fonte de la substance blanche à prédominance sous corticale et frontale / Calcifications et kystes sous épendymaires / Importante pathologie de la substance blanche avec foyers de leucomalacie d'âges variables parfois calcifiés, parfois cavitaires / Retard de gyration / Petits kystes péri-ventriculaires / Epuisement prématuré de la zone germinative

	ATP5PO	F58	Formations kystiques intra-parenchymateuse en périphérie des cornes frontales / Formations kystiques en chapelet le long de la corne frontale (conclut comme séquelles de saignement intra et périventriculaire sous forme de cavités porencéphaliques) / Hyposignaux périventriculaires résiduels d'un saignement intra et périventriculaire (dépôt d'hémosidérine) / Hypoplasie cérébelleuse	*Examen non souhaité par les parents*

Tableau 3. Détail des lésions cérébrales identifiées par IRM fœtale et lors de l'examen fœtopathologique pour les fœtus avec un variant causal.

C- Des variants bialléliques du gène *ESAM* (molécule des jonctions serrées) sont à l'origine de troubles du développement neurologique et d'hémorragies intracrâniennes fœtales

INTRODUCTION

Le séquençage de l'exome entier permet d'identifier des gènes candidats mais leur causalité doit souvent être étayée sur des arguments supplémentaires. Des études fonctionnelles complémentaires et/ou la récurrence de l'implication d'un même gène chez des cas non apparentés sont nécessaires. Comme montré précédemment, l'hétérogénéité génétique et la faible récurrence rencontrées dans les HIC fœtales rendent difficile l'établissement d'une causalité. Le partage international des données permet de rassembler d'autres patients porteurs de variants du même gène afin de tirer des conclusions sur leur implication dans la pathologie étudiée. Pour tous les gènes candidats sans récurrence nous avons utilisé GeneMatcher (Sobreira *et al.,* 2015), un outil de mise en relation des chercheurs qui s'intéressent au même gène, et nous avons pu identifier d'autres cas présentant des similitudes génotypiques et phénotypiques avec l'un de nos fœtus, comme cela a été le cas pour le gène *ESAM*. Pour celui-ci nous avions un cas atteint d'une HIC fœtale détectée au 3ème trimestre de grossesse avec hydranencéphalie et un variant perte de fonction (p.L12*) homozygote dans *ESAM*. Les parents étaient consanguins et une précédente grossesse avait été interrompue à 32 SA devant une hémorragie cérébrale intraparenchymateuse et des calcifications intracrâniennes.

Par collaboration, il a ainsi été possible de regrouper 13 individus, dont quatre fœtus, appartenant à huit familles non apparentées présentant des variants bialléliques de type perte de fonction du gène *ESAM*. Ce gène code pour une protéine des jonctions serrées des cellules endothéliales et n'était associé à aucune condition pathologique humaine.

Les résultats sont présentés dans l'article suivant. La famille issue de notre cohorte de fœtus est la famille D comprenant les individus 6 et 11.

Please cite this article in press as: Lecca et al., Bi-allelic variants in the *ESAM* tight-junction gene cause a neurodevelopmental disorder associated with fetal intracranial hemorrhage, The American Journal of Human Genetics (2023), https://doi.org/10.1016/j.ajhg.2023.03.005

REPORT

Bi-allelic variants in the *ESAM* tight-junction gene cause a neurodevelopmental disorder associated with fetal intracranial hemorrhage

Mauro Lecca,[1] Davut Pehlivan,[2,3,4] Damià Heine Suñer,[5,6] Karin Weiss,[7,8] Thibault Coste,[9,10] Markus Zweier,[11] Yavuz Oktay,[12,13,14] Nada Danial-Farran,[15] Vittorio Rosti,[16] Maria Paola Bonasoni,[17] Alessandro Malara,[1,18] Gianluca Contrò,[19] Roberta Zuntini,[19] Marzia Pollazzon,[19] Rosario Pascarella,[20] Alberto Neri,[21] Carlo Fusco,[22] Dana Marafi,[3,23] Tadahiro Mitani,[3] Jennifer Ellen Posey,[3] Sadik Etka Bayramoglu,[24] Alper Gezdirici,[25] Jessica Hernandez-Rodriguez,[6] Emilia Amengual Cladera,[6] Elena Miravet,[26] Jorge Roldan-Busto,[27] María Angeles Ruiz,[26] Cristofol Vives Bauzá,[28] Liat Ben-Sira,[29,30] Sabine Sigaudy,[31] Anaïs Begemann,[11] Sheila Unger,[32] Serdal Güngör,[33] Semra Hiz,[13,34] Ece Sonmezler,[12,13] Yoav Zehavi,[8,35] Michael Jerdev,[36] Alessandra Balduini,[1,37] Orsetta Zuffardi,[1] Rita Horvath,[38,39] Hanns Lochmüller,[40,41,42] Anita Rauch,[11,43] Livia Garavelli,[19] Elisabeth Tournier-Lasserve,[9,10] Ronen Spiegel,[8,35] James R. Lupski,[3,4,44,45] and Edoardo Errichiello[1,46,*]

Summary

The blood-brain barrier (BBB) is an essential gatekeeper for the central nervous system and incidence of neurodevelopmental disorders (NDDs) is higher in infants with a history of intracerebral hemorrhage (ICH). We discovered a rare disease trait in thirteen individuals, including four fetuses, from eight unrelated families associated with homozygous loss-of-function variant alleles of *ESAM* which encodes an endothelial cell adhesion molecule. The c.115del (p.Arg39Glyfs*33) variant, identified in six individuals from four independent families of Southeastern Anatolia, severely impaired the *in vitro* tubulogenic process of endothelial colony-forming cells, recapitulating previous evidence in null mice, and caused lack of *ESAM* expression in the capillary endothelial cells of damaged brain. Affected individuals with bi-allelic *ESAM* variants showed profound global developmental delay/unspecified intellectual disability, epilepsy, absent or severely delayed speech, varying degrees of spasticity, ventriculomegaly, and ICH/cerebral calcifications, the latter being also observed in the fetuses. Phenotypic traits observed in individuals with bi-allelic *ESAM* variants overlap very closely with other known conditions characterized by endothelial dysfunction due to mutation of genes encoding tight junction molecules. Our findings emphasize the role of brain endothelial dysfunction in NDDs and contribute to the expansion of an emerging group of diseases that we propose to rename as "tightjunctionopathies."

[1]Department of Molecular Medicine, University of Pavia, Pavia, Italy; [2]Section of Pediatric Neurology and Developmental Neuroscience, Department of Pediatrics, Baylor College of Medicine, Houston, TX, USA; [3]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA; [4]Texas Children's Hospital, Houston, TX, USA; [5]Molecular Diagnostics and Clinical Genetics Unit, Hospital Universitari Son Espases, Palma, Illes Balears, Spain; [6]Genomics of Health, Institute of Health Research of the Balearic Islands, Palma, Illes Balears, Spain; [7]Genetics Institute, Rambam Health Care Campus, Haifa, Israel; [8]The Ruth and Bruce Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, Haifa, Israel; [9]AP-HP, Service de Génétique Moléculaire Neurovasculaire, Hôpital Saint-Louis, Paris, France; [10]Université de Paris, INSERM UMR-1141 Neurodiderot, Paris, France; [11]Institute of Medical Genetics, University of Zurich, Schlieren-Zurich, Switzerland; [12]Izmir Biomedicine and Genome Center, Dokuz Eylul University Health Campus, Izmir 35340, Turkey; [13]Izmir International Biomedicine and Genome Institute, Dokuz Eylul University, Izmir 35340, Turkey; [14]Department of Medical Biology, School of Medicine, Dokuz Eylul University, Izmir 35340, Turkey; [15]The Genetic Institute, Emek Medical Center, Afula, Israel; [16]Center for the Study of Myelofibrosis, Laboratory of Biochemistry, Biotechnology and Advanced Diagnosis, IRCCS Policlinico San Matteo Foundation, Pavia, Italy; [17]Pathology Unit, Azienda USL-IRCCS di Reggio Emilia, Reggio Emilia, Italy; [18]Laboratory of Biochemistry-Biotechnology and Advanced Diagnostics, IRCCS Policlinico San Matteo Foundation, Pavia, Italy; [19]Medical Genetics Unit, Azienda USL-IRCCS di Reggio Emilia, Reggio Emilia, Italy; [20]Neuroradiology Unit, Azienda USL-IRCCS di Reggio Emilia, Reggio Emilia, Italy; [21]Ophthalmology Unit, Azienda USL-IRCCS di Reggio Emilia, Reggio Emilia, Italy; [22]Child Neurology and Psychiatry Unit, Azienda USL-IRCCS di Reggio Emilia, Reggio Emilia, Italy; [23]Department of Pediatrics, Faculty of Medicine, Kuwait University, P.O. Box 24923, Safat 13110, Kuwait; [24]Tertiary ROP Center, Health Science University Kanuni Sultan Suleyman Training and Research Hospital, Istanbul 34303, Turkey; [25]Department of Medical Genetics, Basaksehir Cam and Sakura City Hospital, Istanbul 34480, Turkey; [26]Metabolic Pathologies and Pediatric Neurology Unit, Pediatric Service, Hospital Universitari Son Espases, Palma, Illes Balears, Spain; [27]Pediatric Radiology Unit, Radiology Service, Hospital Universitari Son Espases, Palma, Illes Balears, Spain; [28]Neurobiology, Institute of Health Research of the Balearic Islands, Palma, Illes Balears, Spain; [29]Department of Radiology, Division of Pediatric Radiology, Dana Children's Hospital, Tel Aviv Sourasky Medical Center, Tel Aviv University, Tel Aviv, Israel; [30]Sackler School of Medicine, Tel Aviv University, Tel-Aviv, Israel; [31]AP-HM, Service de Génétique, Hôpital de la Timone, Marseille, France; [32]Medical Genetics Service, CHUV, University of Lausanne, Lausanne, Switzerland; [33]Inonu University, Faculty of Medicine, Turgut Ozal Research Center, Department of Pediatric Neurology, Malatya, Turkey; [34]Department of Pediatric Neurology, School of Medicine, Dokuz Eylul University, Izmir 35340, Turkey; [35]Department of Pediatrics B, Emek Medical Center, Afula, Israel; [36]Poriya Medical Center and the Azrieli Faculty of Medicine, Bar-Ilan University, Ramat-Gan, Israel; [37]Department of Biomedical Engineering, Tufts University, Medford, MA, USA; [38]Department of Clinical Neurosciences, School of Clinical Medicine, University of Cambridge, Cambridge Biomedical Campus, Cambridge CB2 0PY, UK; [39]Department of Clinical Neurosciences, John Van Geest Centre for Brain Repair, School of Clinical Medicine, University of Cambridge, Cambridge CB2 0PY, UK; [40]Children's Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, ON K1H 8L1, Canada; [41]Brain and Mind Research Institute, University of Ottawa, Ottawa ON K1H 8L1, Canada; [42]Division of Neurology, Department of Medicine, The Ottawa Hospital, Ottawa, ON K1H 8L1, Canada; [43]University Children's Hospital Zurich, University of Zurich, Zurich, Switzerland; [44]Department of Pediatrics, Baylor College of Medicine, Houston, TX, USA; [45]Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA; [46]Neurogenetics Research Center, IRCCS Mondino Foundation, Pavia, Italy
*Correspondence: edoardo.errichiello@unipv.it
https://doi.org/10.1016/j.ajhg.2023.03.005.

Please cite this article in press as: Lecca et al., Bi-allelic variants in the *ESAM* tight-junction gene cause a neurodevelopmental disorder associated with fetal intracranial hemorrhage, The American Journal of Human Genetics (2023), https://doi.org/10.1016/j.ajhg.2023.03.005

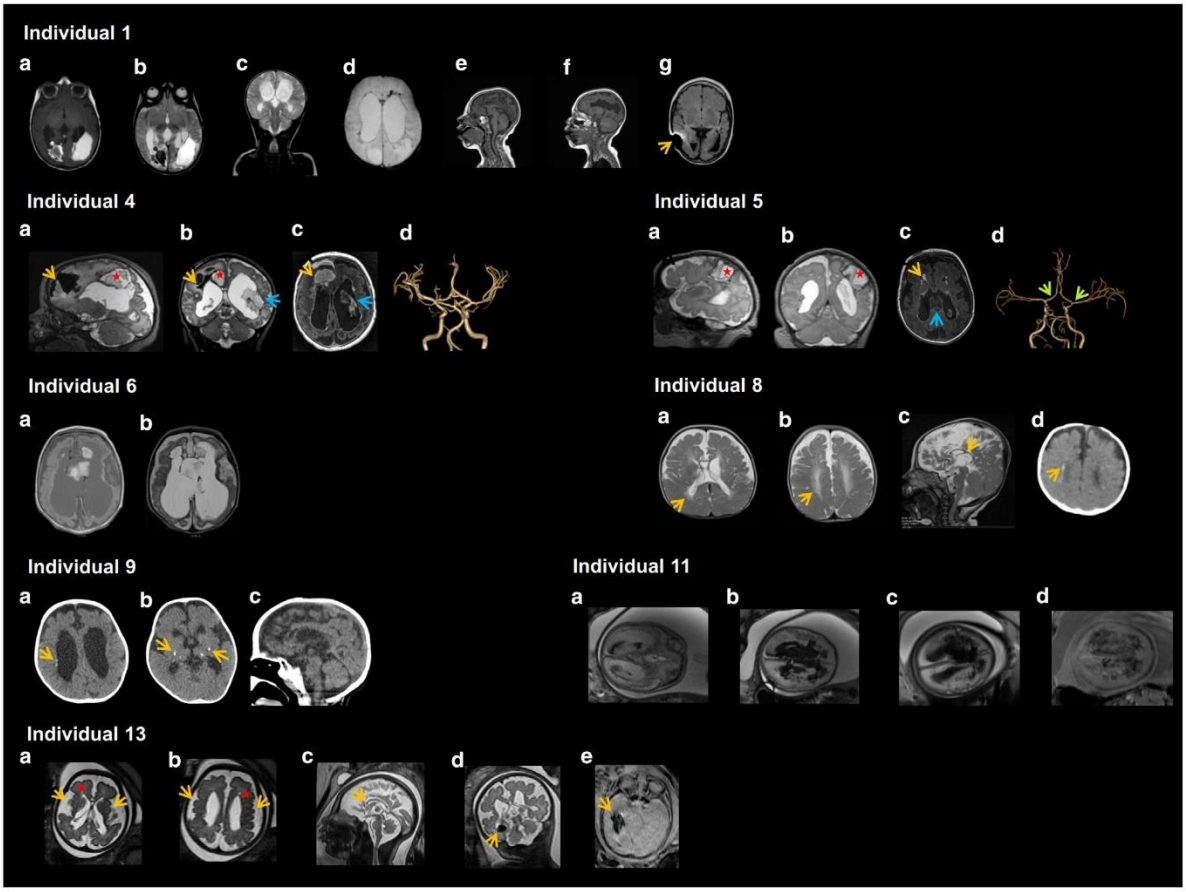


Figure 1. Neuroimaging abnormalities in individuals with homozygous *ESAM* variants

Individual 1: T1-weighted (a and b) images showing various parenchymal hemorrhagic area in different evolution phases, located at the parietal-occipital area. T2-weighted and multi-echo gradient-recalled echo (GRE) T2-weighted images (c and d) showing blood occupying the ventricular spaces (hemoventricle). T1-weighted MRI (e and f) and FLAIR sequences (g) highlight a dysmorphic corpus callosum (e) and poromalacic evolution of the hemorrhagic areas with ventricular dilatation (f and g) and diffuse subependymal microhemorrhage foci. The orange arrow in (g) indicates ventricular derivation.

Individual 4: MRI images at 2 days after birth: 3DT2 sagittal (a), 3DT2 coronal MPR (b), 3DT1 axial (c), 3D TOF (d). Bilateral grade III germinal matrix - intraventricular hemorrhage (blue arrows) with supratentorial posthemorrhagic hydrocephalus. Bilateral parenchymal hematomas with associated subpial hemorrhages (orange arrows) at different stages and parietal encephalomalacia secondary to prior bleeding with residual clot inside (red stars). No evident vascular anomalies in MRI angiography (3D TOF sequence) were found.

Individual 5 (sibling of individual 4): MRI images at 7 days after birth: 3DT2 sagittal (a), 3DT2 coronal MPR (b), 3DT1 axial (c), 3D TOF (d). Microcephaly with simplified gyral pattern and severe hypoplastic corpus callosum (blue arrow). Left parietal subcortical white matter hematoma with subpial hemorrhage (red stars) and encephalomalacia. Right frontal (orange arrow) and multiple bilateral periventricular hemorrhages. MRI angiography (3D TOF sequence) shows slight narrowing of the proximal anterior and middle cerebral arteries (green arrows).

Individual 6: MRI images at 5 days old. T1- (a) and T2- (b) weighted images showing massive dilatation of lateral ventricles with global cerebral parenchymal destruction. Diffuse meningeal and intraventricular hemorrhage with intraventricular clotting. Focal destruction of the septum pellucidum.

Individual 8: axial T2-weighted images showed hyperintensity and volume loss in the bilateral periventricular white matter, frontotemporal atrophy, dilatation in the lateral ventricles (a and b) and hypoplasia of the corpus callosum (c) (orange arrows). Bilateral periventricular calcifications were observed on cranial CT (d), as indicated by the arrow.

Individual 9: Head CT (axial) showing bilateral subependymal (a) and basal ganglia/thalamus calcifications (b) (orange arrows), suggestive of a previous bleeding. Dilation of the lateral ventricles and their straight shape suggest agenesis of the corpus callosum, as confirmed in the sagittal multiplanar reformation (c).

Individual 11: Axial T2-weighted true fast sequence (a and b), axial T2 HASTE sequence (c), and axial T1-weighted sequence (d). The fetal brain MRI shows hydrocephalus, intraventricular hemorrhage, diffuse intraparenchymal, and periventricular hemorrhages with calcifications (hypointense signals on T1-weighted images).

Individual 13: Brain MRI at 37 weeks of gestation. Axial T2 (a and b) showing severe irregular ventriculomegaly with increased extra-axial spaces and significant decrease in gray and white matter volume. Orange arrows indicate an abnormal Sylvian fissure development in

(legend continued on next page)

Please cite this article in press as: Lecca et al., Bi-allelic variants in the *ESAM* tight-junction gene cause a neurodevelopmental disorder associated with fetal intracranial hemorrhage, The American Journal of Human Genetics (2023), https://doi.org/10.1016/j.ajhg.2023.03.005

Neurodevelopmental disorders (NDDs) are a large group of disabilities involving impairment of the brain and neurocognitive development affecting >3% of children worldwide.[1] NDDs include a broad spectrum of phenotypes such as intellectual disability/developmental delay (ID/DD), autism spectrum disorders (ASD), attention-deficit/hyperactivity disorder (ADHD), epilepsy, cerebral palsy, and language impairment. Comorbidity of two or more of these conditions is frequently observed.[2] Furthermore, neuroimaging findings, such as corpus callosum and white matter anomalies, may suggest a specific etiology for NDD.[3] At least 30% of NDDs are thought to have a genetic basis, and a plethora of genes and molecular alterations have been identified, underlying wide clinical and genetic heterogeneity, overlapping phenotypes, and gene pleiotropy, which together make molecular diagnosis a challenging task.[4,5] The implementation of next-generation sequencing (NGS) technologies, especially exome sequencing (ES), which has been proposed as a first-tier test in the diagnostic algorithm of all NDDs followed by array-CGH when necessary, has dramatically increased the percentage of NDD-affected individuals who receive a molecular diagnosis.[6] Furthermore, diagnostic rates reported in recent studies and meta-analyses are 30%–40% or higher when trio analysis is performed and when NDD-affected individuals eligible for diagnostic ES are prioritized before testing based on their phenotypic presentation.[7–13]

Prenatal/perinatal brain damage due to intracranial/intraventricular hemorrhage (ICH/IVH) is one of the leading causes of lifelong disability, including cerebral palsy, epilepsy, sensory impairment, and cognitive deficit.[14–18] The blood-brain barrier (BBB), whose integrity is maintained by junctional adhesion molecules (JAMs), is considered the core structure of the neurovascular unit (NVU) and plays a critical role in maintaining central nervous system homeostasis. Tight junctions (TJs) are an essential BBB component and bi-allelic variants in the tight-junction genes *JAM2/3* (junctional adhesion molecule 2/3 [MIM: 606870 and 606871]) and *OCLN* (occludin [MIM: 602876]) have been implicated in individuals with compromised BBB permeability. These subjects typically show a combination of brain hemorrhage and calcification together with movement disorders and cognitive and neurobehavioral manifestations.[19–24]

In this study, we identified bi-allelic variants in *ESAM* (endothelial cell adhesion molecule [MIM: 614281]), a gene not previously associated with a rare disease trait in humans, which encodes a TJ protein related to JAM proteins involved in the formation and maintenance of the BBB. Genotypic (ES) and clinical phenotypic data from individuals with *ESAM* variants were recruited through GeneMatcher[25] and our international collaborative network. All individuals were carefully evaluated by a multidisciplinary team of pediatric neurologists/neuroophthalmologists/neuroradiologists and clinical geneticists of their respective referral center. We collected clinical information related to neurodevelopment, growth parameters, dysmorphology, neurological, ocular, and vascular manifestations, plus behavior and neuroimaging. No formal intelligence testing was possible in individuals over the age of 5 years because of the severe global developmental delay, including complete lack of language and motor skills. Written human subject research informed consent for genetic analysis and publication of the clinical information, including relevant clinical pictures, was obtained from the parents or legal guardians of each research subject according to the Declaration of Helsinki and Institutional Review Boards of participating research centers.

The age of affected individuals included in the study, 8 males and 4 females (sex was unknown in a 32-week-old fetus), ranged from 31 weeks of gestation to 13 years. Onset of symptoms occurred in the antenatal/neonatal period. Global developmental delay/unspecified intellectual disability (GDD/UID), absent or severely delayed speech, epilepsy, spasticity (mainly consisting of spastic tetraparesis), hypotonia (which frequently occurred neonatally), and dilation of lateral ventricles were observed in all nine live-born children, the latter being also noted in the four fetuses examined in the study (Table S1), while variable microcephaly was reported in four out of nine children (mean OFC *Z* score: −3.35; range: −2.1 to −5.7). Following dilation of lateral ventricles, thinning of the corpus callosum, hydrocephalus, and focal white matter lesions were the most frequently observed neuroimaging abnormalities (eight, seven, and five individuals, respectively) (Figure 1). Notably, hydrocephalus was a feature also observed in all fetuses as well as in a previously reported individual carrying an *ESAM* homozygous nonsense variant (Table S1).[26] Intracranial hemorrhages (ICH) or suspected cerebral microangiopathy, frequently associated with cerebral calcifications (Figure 1), were reported in all individuals, including the four fetuses. Other vascular anomalies were represented by retinal hemorrhage (four individuals)—due to increased tortuosity of retinal vasculature (Figures S1 and S2)—and renal medullary hemorrhage (in a 31-week stillbirth). Dysmorphic facial features were noted in most individuals, mainly consisting of bitemporal narrowing (6/9), highly arched eyebrow and bulbous nasal tip, which were detected with the same frequency (5/9), followed by presence of long eyelashes (4/9), abnormal vermilion (4/9), high narrow palate (3/9), wide nasal bridge (3/9),

(a) and an asymmetric gyral pattern suggestive of dysgyria/polymicrogyria in (b). Small periventricular cysts are indicated by red stars. Sagittal T2 (c) highlighting corpus callosum dysgenesis with absence of rostrum (arrow). Coronal T2 (d): hypointense lesions in the temporal parenchyma (indicated by the arrow) protruding into the temporal horn of the right ventricle. GRE T2 sequences (e) demonstrate susceptibility artifacts consistent with brain calcification (arrow).

Please cite this article in press as: Lecca et al., Bi-allelic variants in the *ESAM* tight-junction gene cause a neurodevelopmental disorder associated with fetal intracranial hemorrhage, The American Journal of Human Genetics (2023), https://doi.org/10.1016/j.ajhg.2023.03.005

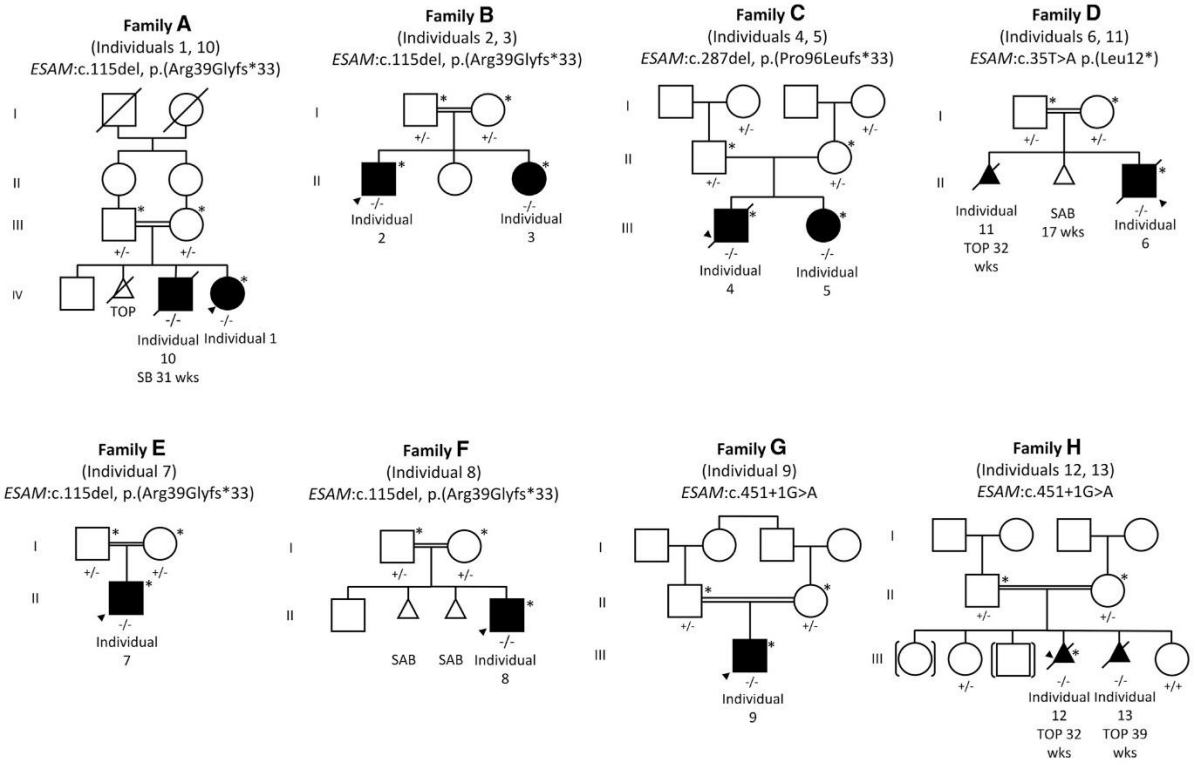


Figure 2. Family pedigrees and genetic findings

Bi-allelic *ESAM* variants are shown in thirteen affected individuals, including four fetuses, from eight unrelated families. The c.115del (p.Arg39Glyfs*33) frameshift variant was identified in six individuals from families A, B, E, and F, originating from the same geographic area in Turkey (southeastern Anatolia). The c.287del (p.Pro96Leufs*33) frameshift variant was detected in two affected siblings from Spain (family C). The c.35T>A (p.Leu12*) nonsense variant was identified in a family of Algerian descent (family D). The c.451+1G>A splice site variant was identified in two independent families of Arab Bedouin descent (families G and H). +/+, +/−, and −/− represents homozygosity for the wild-type allele, heterozygosity, and homozygosity for the mutant allele, respectively. An asterisk beneath an individual indicates that ES was performed. Arrowheads indicate the probands. TOP, termination of pregnancy; SB, stillbirth; SAB, spontaneous abortion; wks, weeks of gestation.

upslanted palpebral fissures (3/9), and microretrognathia (2/9), the latter being also present in a fetus (Table S1 and supplemental note).

ES identified homozygous *ESAM* loss-of-function (LoF) variant alleles in all affected individuals belonging to eight independent families (from A to H) (Figure 2 and supplemental note). The same frameshift variant in exon 2, c.115del (p.Arg39Glyfs*33), was detected in six individuals from four unrelated families, including a 31-week stillbirth (individuals 1, 2, 3, 7, 8, and 10). Individuals 4 and 5 are affected siblings carrying the c.287del (p.Pro96Leufs*33) frameshift variant in exon 3. Individual 6 harbored a nonsense variant within the first gene exon, c.35T>A (p.Leu12*). Although archival autopsy material was not available for his affected sibling (individual 11), a fetus of 32 weeks, it is highly likely that the same variant was present based on strict phenotypic comparison. Individuals 9, 12, and 13 carried the same homozygous splice variant in intron 3 (c.451+1G>A) predicted to cause loss of donor splice site by multiple *in silico* tools (e.g., HSF, MaxEntScan, SpliceAI, CADD-Splice). In all probands no other potentially relevant variants were identified either in known genes associated with neurological and vascular conditions/phenotypes, according to PanelApp, HPO, and GTR, or by filtering the variants for coherent inheritance patterns.

Interestingly, all individuals harboring the c.115del variant originated from the same geographic region in Turkey (southeastern Anatolia), suggesting a founder effect. The variant was not reported in the SHGP (Saudi Human Genome Program) database, the largest genome repository in the Middle East,[27,28] nor in the Iranome database,[29] the GME (Greater Middle East) Variome Project,[30] or the Turkish Variome (TRV) database.[31] Similarly, the c.451+1G>A variant was detected in three individuals from two independent families with the same ethnic origin (Arab Bedouin) and was also absent in the aforementioned databases.

All the variants were classified as likely pathogenic according to the ACMG/AMP criteria (PVS1+PM2) and were predicted to trigger nonsense-mediated mRNA decay (NMD). In all cases, *ESAM* was embedded in ROH/AOH

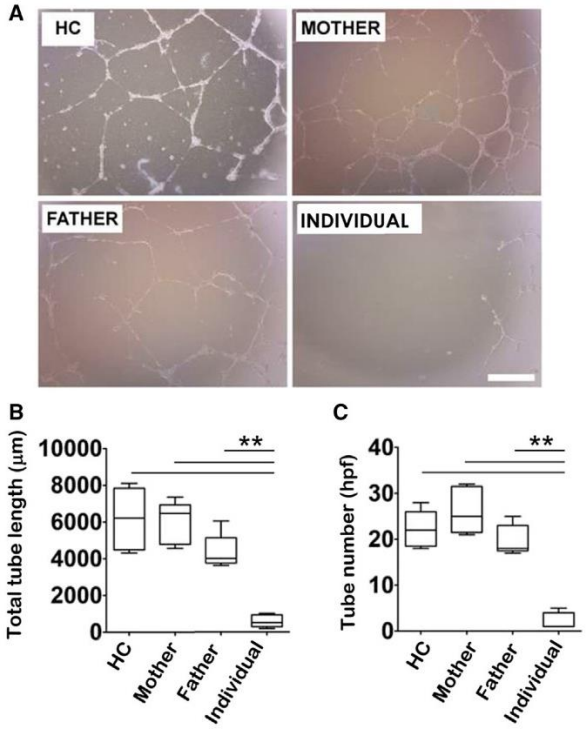


Figure 3. *In vitro* matrigel tubulogenesis assay
(A) Endothelial colony-forming cells (ECFCs) isolated from individual 1 showed significantly decreased tubulogenesis compared to the ECFCs obtained from a healthy control (HC) and parents of individual 1.
(B and C) Quantification of tubulogenesis was performed considering the number of branch points (B) and the tube length (C). Data shown (±SD) are from 3 independent experiments. **$p < 0.01$. Scale bar: 200 mm. More technical details are provided in the supplemental note.

(runs of homozygosity/absence of heterozygosity) stretches spanning from 1.8 to 13.85 Mb in size, consistent with shared ancestry and/or recent consanguinity. Details of the variants are publicly available in the ClinVar database (https://www.ncbi.nlm.nih.gov/clinvar/) with the following accession numbers: SCV002818296 (c.115del), SCV002818297 (c.287del), SCV002818299 (c.35T>A), and SCV002818300 (c.451+1G>A). The predicted 3D structures of the mutant protein products are shown in Figure S3. The crucial immunoglobulin-like (Ig) domains (aa 30–150 and 157–239) and the helical transmembrane domain (aa 247–279) were abolished, whereas the putative motifs predicted by AlphaFold (https://alphafold.ebi.ac.uk/), which are absent in the wild-type protein, did not match with any known protein signature explored through different databases and bioinformatic tools (e.g., Motif Scan, ScanProsite, BlastP).

Considering the crucial role of ESAM in angiogenesis, endothelial permeability, and leukocyte transmigration[32,33] as well as its enriched expression in the endothelium (GTEx, BioGPS, HPA) and the peculiar vascular alterations observed in the identified individuals harboring *ESAM* variants (especially of the brain), we assessed the capability of endothelial progenitors (endothelial colony-forming cells [ECFCs]) to form capillary-like structures *in vitro*. Compared to an age- and sex-matched control sample, ECFCs of individual 1 showed dramatic phenotypical changes of proliferation, migration, and tubulogenesis (Figure 3A) in terms of both number of branch points and tube length (Figures 3B and 3C), which were almost completely abrogated. On the other hand, the ECFCs isolated from her healthy heterozygous parents showed comparable values to the control sample.

The fetal brain of individual 10, carrying the c.115del (p.Arg39Glyfs*33) homozygous frameshift variant, showed lack of ESAM staining in the capillary endothelial cells (Figure 4A), which is consistent with the LoF effect of the c.115del bi-allelic variant. On the other hand, intense expression of ESAM was observed in the cerebral endothelium of an age- and sex-matched control sample, where it was abundantly localized in the plasma membrane of endotheliocytes (Figure 4B). The immunohistochemical findings are in agreement with the results of reverse transcription polymerase chain reaction (RT-PCR) performed on ECFCs of individual 1, showing no detectable *ESAM* expression (Figure S4), as well as with the *in silico* NMD predictions and the 3D-modeling, which overall indicate a lack of functional protein. Furthermore, microscopic and histochemical analysis revealed periventricular leukomalacia (Figure 4C) and brain calcification with abundant presence of so-called "ferruginated" neurons, implying mineralization of neurons as a consequence of hypoxic-ischemic damage (Figure 4D).

In this study, we describe a severe neurodevelopmental rare disease trait caused by homozygous LoF variants of *ESAM* in thirteen individuals, including four fetuses, from eight unrelated families. Global developmental delay/unspecified intellectual disability (GDD/UID) with absent or severely delayed speech, epilepsy, varying degrees of spasticity, ventriculomegaly, and intracranial hemorrhage constitute the cardinal clinical features of the *ESAM*-related phenotype, being observed in all live-born individuals. Other notable clinical signs were thin corpus callosum and variable microcephaly. Among the clinical features, intracranial hemorrhage, together with hydrocephalus and cerebral calcifications, represented the first observable prenatal anomalies that prompted further genetic investigations and, thus, is hypothesized as the etiological trigger of downstream neurodevelopmental defects. In addition to the frank neurological manifestations, *ESAM* alterations also appear to cause a well-defined spectrum of ophthalmological signs, particularly retinal ischemia and abnormal retinal vascular morphology.

ESAM encodes an endothelial cell-selective adhesion molecule, a member of the immunoglobulin receptor family, which mediates homophilic interactions between endothelial cells. Previous studies demonstrated that

Please cite this article in press as: Lecca et al., Bi-allelic variants in the *ESAM* tight-junction gene cause a neurodevelopmental disorder associated with fetal intracranial hemorrhage, The American Journal of Human Genetics (2023), https://doi.org/10.1016/j.ajhg.2023.03.005

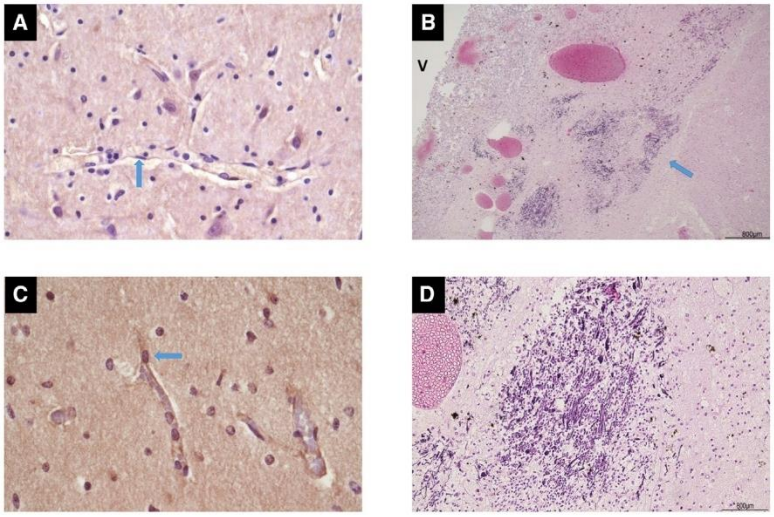


Figure 4. ESAM immunohistochemistry and hematoxylin and eosin staining in the damaged brain tissue (white matter) of individual 10

(A and B) A capillary with ESAM-negative endothelial cells in the proband (A) compared to a control sample (B), as indicated by the blue arrows. Magnification: 40 HPF.

(C) Periventricular leukomalacia in individual 10: the white matter close to the lateral ventricle (V) shows multiple areas of calcification (blue arrow). Hematoxylin and eosin staining, magnification 4 HPF.

(D) Calcified areas showing elongated sticks and round bodies as a result of encrusted ("ferruginated") neurons and their axons with mineral salts. Hematoxylin and eosin staining, magnification 20 HPF. More technical details are provided in the supplemental note.

ESAM expression is primarily restricted to embryonic and adult vasculature, where it regulates endothelial permeability and neutrophil extravasation.[33–35] Mouse model experiments showed that ESAM regulates albumin extravasation at the glomeruli and plays a role in the initiation of diabetic nephropathy.[36] A recent proteomic study of individuals affected by diabetic kidney disease confirmed ESAM as a candidate circulating biomarker to predict the risk of progression to kidney failure.[37] In this regard, we did not observe renal dysfunction in our pediatric subjects, although later manifestations cannot be completely ruled out and careful follow-up in the next years would be indicated. Interestingly, a proteomic analysis revealed that ESAM interacts with endoglin,[38] encoded by *ENG*, whose mutation causes hereditary hemorrhagic telangiectasia (HHT [MIM: 187300]), a condition that includes ICH as a clinical feature.

In a previous study, a *de novo* heterozygous frameshift variant in *ESAM* exon 4, c.543_546dup (p.Val183Argfs*48), was reported in a 23-year-old male individual affected by schizophrenia and developmental delay.[39] Therefore, it could be hypothesized that monoallelic LoF variants in *ESAM* may at most contribute (possibly together with other factors) to mild neurological phenotypes, in contrast to the severe phenotype we observed in all our homozygous probands. Nevertheless, we did not observe overt neuropsychiatric features in any of our probands' heterozygous parents nor in other carrier family members. This finding is in line with the gene constraint metrics (pLI: 0.20; LOEUF: 0.59; sHet: 0.05; %HI: 63.38) and the presence of an ~34 kb deletion encompassing *ESAM* in the DGV (Database of Genomic Variants) repository (nsv556487), indicating that *ESAM* is quite tolerant to haploinsufficiency, as it is for triplosensitivity (pTriplo: 0.28).[40] Altogether, it is at most possible to speculate that heterozygous LoF variants may represent a susceptibility factor characterized by incomplete penetrance, and, in any case, further studies are needed to better address this point.

Our *in vitro* tubulogenesis assays recapitulated findings in null mice ($Esam^{-/-}$) where a decrease of tube formation and vascular density was documented, although without overt morphological defects of the vasculature.[32] On the contrary, we observed a high frequency of vascular manifestations (intracranial hemorrhage, cerebral calcifications, focal white matter lesions, hydrocephalus, and dilation of lateral ventricles), in agreement with the critical role of *ESAM* in endothelial homeostasis. The fact that both null mice and human subjects carrying homozygous LoF variants in *ESAM* show a non-lethal (although extremely severe) phenotype could potentially be explained considering the redundant functional role of ESAM in physiological angiogenesis and the presence of other endothelial adhesion molecules which may compensate for its dysfunction.

It is noteworthy that the phenotype associated with bi-allelic variants of *ESAM* overlaps very closely with other known conditions characterized by endothelial dysfunction due to mutation of genes encoding tight junction molecules, namely *JAM2*, *JAM3,* and *OCLN*. Bi-allelic LoF variants of *JAM2* are associated with idiopathic basal ganglia calcification (IBGC8 [MIM: 618824]), a condition commonly characterized by intracerebral calcifications, cognitive decline, learning difficulties, seizures, slurred speech, movement disorders, and psychiatric symptoms.[19,20] Homozygous LoF variants in *JAM3* have been identified in individuals with multifocal intraparenchymal hemorrhage and subependymal calcification as cardinal features (HDBSCC [MIM: 613730]), accompanied by reduced white matter volume, porencephaly, and massive cystic degeneration resulting in enlarged ventricles.[21,22] In most of the cases, this condition is extremely severe, with death occurring in the first few weeks of life, while alive individuals develop profound developmental

Please cite this article in press as: Lecca et al., Bi-allelic variants in the *ESAM* tight-junction gene cause a neurodevelopmental disorder associated with fetal intracranial hemorrhage, The American Journal of Human Genetics (2023), https://doi.org/10.1016/j.ajhg.2023.03.005

Table 1. Phenotypic comparison of individuals with bi-allelic variants in *JAM2*, *JAM3*, *OCLN*, and *ESAM*

Clinical features	*JAM2*	*JAM3*	*OCLN*	*ESAM*
Head and neck				
Microcephaly	no	yes	yes	yes
Cataracts	no	yes	yes (rare)	no
Facial dysmorphisms	no	no	yes: long philtrum, microretrognathia, low-set ears, anteverted nares, high arched palate	yes: bitemporal narrowing, highly arched eyebrow, bulbous nasal tip, long eyelashes, high narrow palate, wide nasal bridge, upslanted palpebral fissures, microretrognathia, anteverted nares
Neurologic				
Developmental delay	yes	yes (severe)	yes (severe)	yes (severe)
Seizures	yes (rare)	yes	yes	yes
Spasticity	yes	yes	yes	yes
Hypotonia	no	yes	yes	yes
Neuroimaging findings				
Intracranial calcifications	yes	yes	yes	yes
Intracranial hemorrhage	no	yes	no	yes
Ventriculomegaly	no	yes	yes	yes
Corpus callosum anomalies	no	yes	yes	yes
Abdomen				
Hepatomegaly	no	yes	yes	no
Genitourinary				
Renal anomalies	no	yes (rare)	yes (rare)	yes (rare) (renal medullary hemorrhage)
Disease onset	adulthood	neonatal	neonatal	antenatal/neonatal

delay, microcephaly, generalized spasticity, and seizures, as we observed in our cohort of affected individuals. Finally, bi-allelic LoF variants affecting a tight junction protein, occludin, have been reported in individuals showing profound developmental delay, early-onset seizures, microcephaly, ventriculomegaly, spasticity, polymicrogyria, loss of white matter, and intracranial calcifications, referred to as pseudo-TORCH syndrome 1 (PTORCH1 [MIM: 251290]).[23,24] Overall, neurological manifestations (namely DD, seizures, spasticity, and hypotonia) and intracranial calcifications represent the cardinal clinical features shared among these conditions, which we also observed in *ESAM*-affected individuals, whereas ICH seems specifically restricted to *JAM3*- and *ESAM*-related phenotypes (Table 1). Furthermore, individuals with pathogenic variants in *COL4A1*, encoding type IV collagen alpha-1 chain protein (a vascular basement membrane protein highly expressed in brain vessels), show intracerebral hemorrhage with calcifications, porencephaly, cystic brain lesions, hydrocephalus, seizures, and retinal arterial tortuosity.[41–45] In aggregate these data suggest that defective cell adhesion and solute flux through the paracellular spaces in the neurovascular unit (NVU) is a key pathomechanism of brain calcification and hemorrhage in this group of disorders, which we here propose to rename as "tightjunctionopathies."

Prenatal ICH is a strong risk factor for perinatal mortality and adverse neurodevelopmental outcome. In a recent systematic review, cerebral palsy (CP) was observed at postnatal follow-up in 32% of fetuses with prenatal diagnosis of ICH, in most cases accompanied by severe neurodevelopmental delay.[18] Accordingly, we observed varying degrees of spasticity (mainly consisting of spastic tetraparesis) in all children with bi-allelic *ESAM* variants, together with GDD/UID and other neurological manifestations. Genetic testing is rarely considered in the diagnostic workup of CP individuals, although a recent systematic review and meta-analysis reported non-negligible diagnostic yields of ES and CMA (23% and 5%, respectively).[46] Our findings add a piece of knowledge to the genetic causes of CP conferring vulnerability to brain injury during prenatal life and causing long-term neurodevelopmental complications.

In summary, our study reveals a rare Mendelian condition associated with bi-allelic *ESAM* variants and emphasizes the increasingly emerging role of brain endothelial dysfunction in neurodevelopmental disorders. In fact, the interplay between vascular and neuronal systems is critical for the normal growth and function of neurons, considering that brain development relies heavily on proper cerebrovascular maturation forming a complex system of multidirectional communication

Please cite this article in press as: Lecca et al., Bi-allelic variants in the *ESAM* tight-junction gene cause a neurodevelopmental disorder associated with fetal intracranial hemorrhage, The American Journal of Human Genetics (2023), https://doi.org/10.1016/j.ajhg.2023.03.005

known as the "neuro-glial-vascular" unit.[47–50] Therefore, it is conceivable that alterations in cerebrovascular processes during early development may have long-lasting neurodevelopmental consequences, including non-malformative conditions. In fact, endothelial dysfunction leading to impaired cerebral angiogenesis has been documented in a mouse model hemizygous for the ASD-related 16p11.2 deletion, pointing to a potential role for endothelial impairment in ASD pathogenesis.[51] In this regard, a previous ASD postmortem brain study suggested an impairment in cerebral angiogenesis,[52] whereas a functional imaging study proposed a possible link between ASD and altered cerebral blood flow.[53] Altogether, these data also open a new perspective on potential treatment options for mitigating endothelial cell dysfunction in neurodevelopmental conditions due to altered BBB function.

Data and code availability

The published article includes all datasets generated or analyzed during this study. In particular, details of the variants are publicly available in the ClinVar database (https://www.ncbi.nlm.nih.gov/clinvar/) with the following accession numbers: SCV002818296 (c.115del), SCV002818297 (c.287del), SCV002818299 (c.35T>A), and SCV002818300 (c.451+1G>A).

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.ajhg.2023.03.005.

Acknowledgments

The authors thank all probands and families for their participation in this study. More details are provided in the supplemental note.

Author contributions

M.L.: Data curation, Investigation, Validation, Visualization; D.P., D.H.S., K.W., T.C., M.Z., Y.O., N.D.-F.: Funding acquisition, Investigation, Validation, Supervision; V.R., M.P.B., A.M., R.Z., D.M., T.M., J.E.P., J.H.-R., E.A.C., E.M., C.V.B., E.S.: Investigation; G.C., M.P., R.P., A.N., C.F., S.E.B., A.G., J.R.-B., M.A.R., L.B.S., S.S., A. Begemann, S.U., S.G., S.H., Y.Z., M.J.: Resources; A. Balduini, O.Z., R.H., H.L., A.R., L.G., E.T.-L., R.S., J.R.L.: Supervision, Funding acquisition, Writing – review & editing; E.E.: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Project administration, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing

Declaration of interests

J.R.L. has stock ownership in 23andMe and is a paid consultant for the Regeneron Genetics Center.

Received: January 9, 2023
Accepted: March 7, 2023
Published: March 29, 2023

Web resources

ClinVar, https://www.ncbi.nlm.nih.gov/clinvar/
OMIM, https://www.omim.org/

References

1. Olusanya, B.O., Wright, S.M., Nair, M.K.C., Boo, N.Y., Halpern, R., Kuper, H., Abubakar, A.A., Almasri, N.A., Arabloo, J., Arora, N.K., et al. (2020). Global burden of childhood epilepsy, intellectual disability, and sensory impairments. Pediatrics *146*, e20192623. https://doi.org/10.1542/peds.2019-2623.
2. Parenti, I., Rabaneda, L.G., Schoen, H., and Novarino, G. (2020). Neurodevelopmental disorders: from genetics to functional pathways. Trends Neurosci. *43*, 608–621. https://doi.org/10.1016/j.tins.2020.05.004.
3. Engbers, H.M., Nievelstein, R.A.J., Gooskens, R.H.J.M., Kroes, H.Y., van Empelen, R., Braams, O., Wittebol-Post, D., Hendriks, M.M.W.B., and Visser, G. (2010). The clinical utility of MRI in patients with neurodevelopmental disorders of unknown origin. Eur. J. Neurol. *17*, 815–822. https://doi.org/10.1111/j.1468-1331.2009.02927.x.
4. Retterer, K., Juusola, J., Cho, M.T., Vitazka, P., Millan, F., Gibellini, F., Vertino-Bell, A., Smaoui, N., Neidich, J., Monaghan, K.G., et al. (2016). Clinical application of whole-exome sequencing across clinical indications. Genet. Med. *18*, 696–704. https://doi.org/10.1038/gim.2015.148.
5. Mitani, T., Isikay, S., Gezdirici, A., Gulec, E.Y., Punetha, J., Fatih, J.M., Herman, I., Akay, G., Du, H., Calame, D.G., et al. (2021). High prevalence of multilocus pathogenic variation in neurodevelopmental disorders in the Turkish population. Am. J. Hum. Genet. *108*, 1981–2005. https://doi.org/10.1016/j.ajhg.2021.08.009.
6. Srivastava, S., Love-Nichols, J.A., Dies, K.A., Ledbetter, D.H., Martin, C.L., Chung, W.K., Firth, H.V., Frazier, T., Hansen, R.L., Prock, L., et al. (2019). Meta-analysis and multidisciplinary consensus statement: exome sequencing is a first-tier clinical diagnostic test for individuals with neurodevelopmental disorders. Genet. Med. *21*, 2413–2421. https://doi.org/10.1038/s41436-019-0554-6.
7. Thevenon, J., Duffourd, Y., Masurel-Paulet, A., Lefebvre, M., Feillet, F., El Chehadeh-Djebbar, S., St-Onge, J., Steinmetz, A., Huet, F., Chouchane, M., et al. (2016). Diagnostic odyssey in severe neurodevelopmental disorders: toward clinical whole-exome sequencing as a first-line diagnostic test. Clin. Genet. *89*, 700–707. https://doi.org/10.1111/cge.12732.
8. Vrijenhoek, T., Middelburg, E.M., Monroe, G.R., van Gassen, K.L.I., Geenen, J.W., Hövels, A.M., Knoers, N.V., van Amstel, H.K.P., and Frederix, G.W.J. (2018). Whole-exome sequencing in intellectual disability; cost before and after a diagnosis. Eur. J. Hum. Genet. *26*, 1566–1571. https://doi.org/10.1038/s41431-018-0203-6.
9. Geisheker, M.R., Heymann, G., Wang, T., Coe, B.P., Turner, T.N., Stessman, H.A.F., Hoekzema, K., Kvarnung, M., Shaw, M., Friend, K., et al. (2017). Hotspots of missense mutation identify neurodevelopmental disorder genes and functional domains. Nat. Neurosci. *20*, 1043–1051. https://doi.org/10.1038/nn.4589.
10. Lindstrand, A., Eisfeldt, J., Pettersson, M., Carvalho, C.M.B., Kvarnung, M., Grigelioniene, G., Anderlid, B.M., Bjerin, O., Gustavsson, P., Hammarsjö, A., et al. (2019). From cytogenetics to cytogenomics: whole-genome sequencing as a

first-line test comprehensively captures the diverse spectrum of disease-causing genetic variation underlying intellectual disability. Genome Med. *11*, 68. https://doi.org/10.1186/s13073-019-0675-1.

11. Martinez-Granero, F., Blanco-Kelly, F., Sanchez-Jimeno, C., Avila-Fernandez, A., Arteche, A., Bustamante-Aragones, A., Rodilla, C., Rodríguez-Pinilla, E., Riveiro-Alvarez, R., Tahsin-Swafiri, S., et al. (2021). Comparison of the diagnostic yield of aCGH and genome-wide sequencing across different neurodevelopmental disorders. NPJ Genom. Med. *6*, 25. https://doi.org/10.1038/s41525-021-00188-7.
12. Hiz Kurul, S., Oktay, Y., Töpf, A., Szabó, N.Z., Güngör, S., Yaramis, A., Sonmezler, E., Matalonga, L., Yis, U., Schon, K., et al. (2022). High diagnostic rate of trio exome sequencing in consanguineous families with neurogenetic diseases. Brain *145*, 1507–1518. https://doi.org/10.1093/brain/awab395.
13. Dingemans, A.J.M., Hinne, M., Jansen, S., van Reeuwijk, J., de Leeuw, N., Pfundt, R., van Bon, B.W., Vulto-van Silfhout, A.T., Kleefstra, T., Koolen, D.A., et al. (2022). Phenotype based prediction of exome sequencing outcome using machine learning for neurodevelopmental disorders. Genet. Med. *24*, 645–653. https://doi.org/10.1016/j.gim.2021.10.019.
14. Vohr, B., Allan, W.C., Scott, D.T., Katz, K.H., Schneider, K.C., Makuch, R.W., and Ment, L.R. (1999). Early-onset intraventricular hemorrhage in preterm neonates: incidence of neurodevelopmental handicap. Semin. Perinatol. *23*, 212–217. https://doi.org/10.1016/s0146-0005(99)80065-2.
15. Moretti, R., Pansiot, J., Bettati, D., Strazielle, N., Ghersi-Egea, J.F., Damante, G., Fleiss, B., Titomanlio, L., and Gressens, P. (2015). Blood-brain barrier dysfunction in disorders of the developing brain. Front. Neurosci. *9*, 40. https://doi.org/10.3389/fnins.2015.00040.
16. De Haan, T.R., Langeslag, J., van der Lee, J.H., and van Kaam, A.H. (2018). A systematic review comparing neurodevelopmental outcome in term infants with hypoxic and vascular brain injury with and without seizures. BMC Pediatr. *18*, 147. https://doi.org/10.1186/s12887-018-1116-9.
17. Gilard, V., Chadie, A., Ferracci, F.X., Brasseur-Daudruy, M., Proust, F., Marret, S., and Curey, S. (2018). Post hemorrhagic hydrocephalus and neurodevelopmental outcomes in a context of neonatal intraventricular hemorrhage: an institutional experience in 122 preterm children. BMC Pediatr. *18*, 288. https://doi.org/10.1186/s12887-018-1249-x.
18. Sileo, F.G., Zöllner, J., D'Antonio, F., Islam, S., Papageorghiou, A.T., and Khalil, A. (2022). Perinatal and long-term outcome of fetal intracranial hemorrhage: systematic review and meta-analysis. Ultrasound Obstet. Gynecol. *59*, 585–595. https://doi.org/10.1002/uog.24766.
19. Cen, Z., Chen, Y., Chen, S., Wang, H., Yang, D., Zhang, H., Wu, H., Wang, L., Tang, S., Ye, J., et al. (2020). Biallelic loss-of-function mutations in *JAM2* cause primary familial brain calcification. Brain *143*, 491–502. https://doi.org/10.1093/brain/awz392.
20. Schottlaender, L.V., Abeti, R., Jaunmuktane, Z., Macmillan, C., Chelban, V., O'Callaghan, B., McKinley, J., Maroofian, R., Efthymiou, S., Athanasiou-Fragkouli, A., et al. (2020). Biallelic *JAM2* variants lead to early-onset recessive primary familial brain calcification. Am. J. Hum. Genet. *106*, 412–421. https://doi.org/10.1016/j.ajhg.2020.02.007.
21. Mochida, G.H., Ganesh, V.S., Felie, J.M., Gleason, D., Hill, R.S., Clapham, K.R., Rakiec, D., Tan, W.H., Akawi, N., Al-Saffar, M., et al. (2010). A homozygous mutation in the tight-junction protein JAM3 causes hemorrhagic destruction of the brain, subependymal calcification, and congenital cataracts. Am. J. Hum. Genet. *87*, 882–889. https://doi.org/10.1016/j.ajhg.2010.10.026.
22. Akawi, N.A., Canpolat, F.E., White, S.M., Quilis-Esquerra, J., Morales Sanchez, M., Gamundi, M.J., Mochida, G.H., Walsh, C.A., Ali, B.R., and Al-Gazali, L. (2013). Delineation of the clinical, molecular and cellular aspects of novel *JAM3* mutations underlying the autosomal recessive hemorrhagic destruction of the brain, subependymal calcification, and congenital cataracts. Hum. Mutat. *34*, 498–505. https://doi.org/10.1002/humu.22263.
23. O'Driscoll, M.C., Daly, S.B., Urquhart, J.E., Black, G.C.M., Pilz, D.T., Brockmann, K., McEntagart, M., Abdel-Salam, G., Zaki, M., Wolf, N.I., et al. (2010). Recessive mutations in the gene encoding the tight junction protein occludin cause band-like calcification with simplified gyration and polymicrogyria. Am. J. Hum. Genet. *87*, 354–364. https://doi.org/10.1016/j.ajhg.2010.07.012.
24. Abdel-Hamid, M.S., Abdel-Salam, G.M.H., Issa, M.Y., Emam, B.A., and Zaki, M.S. (2017). Band-like calcification with simplified gyration and polymicrogyria: report of 10 new families and identification of five novel OCLN mutations. J. Hum. Genet. *62*, 553–559. https://doi.org/10.1038/jhg.2017.4.
25. Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015). GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum. Mutat. *36*, 928–930. https://doi.org/10.1002/humu.22844.
26. Jin, S.C., Dong, W., Kundishora, A.J., Panchagnula, S., Moreno-De-Luca, A., Furey, C.G., Allocco, A.A., Walker, R.L., Nelson-Williams, C., Smith, H., et al. (2020). Exome sequencing implicates genetic disruption of prenatal neurogliogenesis in sporadic congenital hydrocephalus. Nat. Med. *26*, 1754–1765. https://doi.org/10.1038/s41591-020-1090-2.
27. Saudi Mendeliome Group (2015). Comprehensive gene panels provide advantages over clinical exome sequencing for Mendelian diseases. Genome Biol. *16*, 134. https://doi.org/10.1186/s13059-015-0693-2.
28. Abouelhoda, M., Sobahy, T., El-Kalioby, M., Patel, N., Shamseldin, H., Monies, D., Al-Tassan, N., Ramzan, K., Imtiaz, F., Shaheen, R., and Alkuraya, F.S. (2016). Clinical genomics can facilitate countrywide estimation of autosomal recessive disease burden. Genet. Med. *18*, 1244–1249. https://doi.org/10.1038/gim.2016.37.
29. Fattahi, Z., Beheshtian, M., Mohseni, M., Poustchi, H., Sellars, E., Nezhadi, S.H., Amini, A., Arzhangi, S., Jalalvand, K., Jamali, P., et al. (2019). Iranome: A catalog of genomic variations in the Iranian population. Hum. Mutat. *40*, 1968–1984. https://doi.org/10.1002/humu.23880.
30. Scott, E.M., Halees, A., Itan, Y., Spencer, E.G., He, Y., Azab, M.A., Gabriel, S.B., Belkadi, A., Boisson, B., Abel, L., et al. (2016). Characterization of Greater Middle Eastern genetic variation for enhanced disease gene discovery. Nat. Genet. *48*, 1071–1076. https://doi.org/10.1038/ng.3592.
31. Kars, M.E., Başak, A.N., Onat, O.E., Bilguvar, K., Choi, J., Itan, Y., Çağlar, C., Palvadeau, R., Casanova, J.L., Cooper, D.N., et al. (2021). The genetic structure of the Turkish population reveals high levels of variation and admixture. Proc. Natl. Acad. Sci. USA *118*. e2026076118. https://doi.org/10.1073/pnas.2026076118.
32. Ishida, T., Kundu, R.K., Yang, E., Hirata, K.i., Ho, Y.D., and Quertermous, T. (2003). Targeted disruption of endothelial

cell-selective adhesion molecule inhibits angiogenic processes in vitro and in vivo. J. Biol. Chem. *278*, 34598–34604. https://doi.org/10.1074/jbc.M304890200.

33. Wegmann, F., Petri, B., Khandoga, A.G., Moser, C., Khandoga, A., Volkery, S., Li, H., Nasdala, I., Brandau, O., Fässler, R., et al. (2006). ESAM supports neutrophil extravasation, activation of Rho, and VEGF-induced vascular permeability. J. Exp. Med. *203*, 1671–1677. https://doi.org/10.1084/jem.20060565.
34. Hirata, K.I., Ishida, T., Penta, K., Rezaee, M., Yang, E., Wohlgemuth, J., and Quertermous, T. (2001). Cloning of an immunoglobulin family adhesion molecule selectively expressed by endothelial cells. J. Biol. Chem. *276*, 16223–16231. https://doi.org/10.1074/jbc.M100630200.
35. Nasdala, I., Wolburg-Buchholz, K., Wolburg, H., Kuhn, A., Ebnet, K., Brachtendorf, G., Samulowitz, U., Kuster, B., Engelhardt, B., Vestweber, D., and Butz, S. (2002). A transmembrane tight junction protein selectively expressed on endothelial cells and platelets. J. Biol. Chem. *277*, 16294–16303. https://doi.org/10.1074/jbc.M111999200.
36. Hara, T., Ishida, T., Cangara, H.M., and Hirata, K.i. (2009). Endothelial cell-selective adhesion molecule regulates albuminuria in diabetic nephropathy. Microvasc. Res. *77*, 348–355. https://doi.org/10.1016/j.mvr.2009.01.002.
37. Kobayashi, H., Looker, H.C., Satake, E., Saulnier, P.J., Md Dom, Z.I., O'Neil, K., Ihara, K., Krolewski, B., Galecki, A.T., Niewczas, M.A., et al. (2022). Results of untargeted analysis using the SOMAscan proteomics platform indicates novel associations of circulating proteins with risk of progression to kidney failure in diabetes. Kidney Int. *102*, 370–381. https://doi.org/10.1016/j.kint.2022.04.022.
38. Gallardo-Vara, E., Ruiz-Llorente, L., Casado-Vela, J., Ruiz-Rodríguez, M.J., López-Andrés, N., Pattnaik, A.K., Quintanilla, M., and Bernabeu, C. (2019). Endoglin Protein Interactome Profiling Identifies TRIM21 and Galectin-3 as New Binding Partners. Cells *8*, 1082. https://doi.org/10.3390/cells8091082.
39. Xu, B., Roos, J.L., Dexheimer, P., Boone, B., Plummer, B., Levy, S., Gogos, J.A., and Karayiorgou, M. (2011). Exome sequencing supports a de novo mutational paradigm for schizophrenia. Nat. Genet. *43*, 864–868. https://doi.org/10.1038/ng.902.
40. Collins, R.L., Glessner, J.T., Porcu, E., Lepamets, M., Brandon, R., Lauricella, C., Han, L., Morley, T., Niestroj, L.M., Ulirsch, J., et al. (2022). A cross-disorder dosage sensitivity map of the human genome. Cell *185*, 3041–3055.e25. https://doi.org/10.1016/j.cell.2022.06.036.
41. de Vries, L.S., Koopman, C., Groenendaal, F., Van Schooneveld, M., Verheijen, F.W., Verbeek, E., Witkamp, T.D., van der Worp, H.B., and Mancini, G. (2009). *COL4A1* mutation in two preterm siblings with antenatal onset of parenchymal hemorrhage. Ann. Neurol. *65*, 12–18. https://doi.org/10.1002/ana.21525.
42. Yoneda, Y., Haginoya, K., Kato, M., Osaka, H., Yokochi, K., Arai, H., Kakita, A., Yamamoto, T., Otsuki, Y., Shimizu, S.i., et al. (2013). Phenotypic spectrum of *COL4A1* mutations: porencephaly to schizencephaly. Ann. Neurol. *73*, 48–57. https://doi.org/10.1002/ana.23736.
43. Meuwissen, M.E.C., Halley, D.J.J., Smit, L.S., Lequin, M.H., Cobben, J.M., de Coo, R., van Harssel, J., Sallevelt, S., Woldringh, G., van der Knaap, M.S., et al. (2015). The expanding phenotype of *COL4A1* and *COL4A2* mutations: clinical data on 13 newly identified families and a review of the literature. Genet. Med. *17*, 843–853. https://doi.org/10.1038/gim.2014.210.
44. Alavi, M.V., Mao, M., Pawlikowski, B.T., Kvezereli, M., Duncan, J.L., Libby, R.T., John, S.W.M., and Gould, D.B. (2016). *Col4a1* mutations cause progressive retinal neovascular defects and retinopathy. Sci. Rep. *6*, 18602. https://doi.org/10.1038/srep18602.
45. Zagaglia, S., Selch, C., Nisevic, J.R., Mei, D., Michalak, Z., Hernandez-Hernandez, L., Krithika, S., Vezyroglou, K., Varadkar, S.M., Pepler, A., et al. (2018). Neurologic phenotypes associated with *COL4A1/2* mutations: Expanding the spectrum of disease. Neurology *91*, e2078–e2088. https://doi.org/10.1212/WNL.0000000000006567.
46. Srivastava, S., Lewis, S.A., Cohen, J.S., Zhang, B., Aravamuthan, B.R., Chopra, M., Sahin, M., Kruer, M.C., and Poduri, A. (2022). Molecular diagnostic yield of exome sequencing and chromosomal microarray in cerebral palsy: a systematic review and meta-analysis. JAMA Neurol. *79*, 1287–1295. https://doi.org/10.1001/jamaneurol.2022.3549.
47. Lacoste, B., Comin, C.H., Ben-Zvi, A., Kaeser, P.S., Xu, X., Costa, L.d.F., and Gu, C. (2014). Sensory-related neural activity regulates the structure of vascular networks in the cerebral cortex. Neuron *83*, 1117–1130. https://doi.org/10.1016/j.neuron.2014.07.034.
48. Andreone, B.J., Lacoste, B., and Gu, C. (2015). Neuronal and vascular interactions. Annu. Rev. Neurosci. *38*, 25–46. https://doi.org/10.1146/annurev-neuro-071714-033835.
49. Kugler, E.C., Greenwood, J., and MacDonald, R.B. (2021). The "neuro-glial-vascular" unit: the role of glia in neurovascular unit formation and dysfunction. Front. Cell Dev. Biol. *9*, 732820. https://doi.org/10.3389/fcell.2021.732820.
50. Zisis, E., Keller, D., Kanari, L., Arnaudon, A., Gevaert, M., Delemontex, T., Coste, B., Foni, A., Abdellah, M., Calì, C., et al. (2021). Digital reconstruction of the neuro-glia-vascular architecture. Cereb. Cortex *31*, 5686–5703. https://doi.org/10.1093/cercor/bhab254.
51. Ouellette, J., Toussay, X., Comin, C.H., Costa, L.D.F., Ho, M., Lacalle-Aurioles, M., Freitas-Andrade, M., Liu, Q.Y., Leclerc, S., Pan, Y., et al. (2020). Vascular contributions to 16p11.2 deletion autism syndrome modeled in mice. Nat. Neurosci. *23*, 1090–1101. https://doi.org/10.1038/s41593-020-0663-1.
52. Azmitia, E.C., Saccomano, Z.T., Alzoobaee, M.F., Boldrini, M., and Whitaker-Azmitia, P.M. (2016). Persistent angiogenesis in the autism brain: an immunocytochemical study of postmortem cortex, brainstem and cerebellum. J. Autism Dev. Disord. *46*, 1307–1318. https://doi.org/10.1007/s10803-015-2672-6.
53. Jann, K., Hernandez, L.M., Beck-Pancer, D., McCarron, R., Smith, R.X., Dapretto, M., and Wang, D.J.J. (2015). Altered resting perfusion and functional connectivity of default mode network in youth with autism spectrum disorder. Brain Behav. *5*, e00358. https://doi.org/10.1002/brb3.358.

The American Journal of Human Genetics, Volume 110

Supplemental information

Bi-allelic variants in the *ESAM* tight-junction gene cause a neurodevelopmental disorder associated with fetal intracranial hemorrhage

Mauro Lecca, Davut Pehlivan, Damià Heine Suñer, Karin Weiss, Thibault Coste, Markus Zweier, Yavuz Oktay, Nada Danial-Farran, Vittorio Rosti, Maria Paola Bonasoni, Alessandro Malara, Gianluca Contrò, Roberta Zuntini, Marzia Pollazzon, Rosario Pascarella, Alberto Neri, Carlo Fusco, Dana Marafi, Tadahiro Mitani, Jennifer Ellen Posey, Sadik Etka Bayramoglu, Alper Gezdirici, Jessica Hernandez-Rodriguez, Emilia Amengual Cladera, Elena Miravet, Jorge RoldanBusto, María Angeles Ruiz, Cristofol Vives Bauzá, Liat Ben-Sira, Sabine Sigaudy, Anaïs Begemann, Sheila Unger, Serdal Güngör, Semra Hiz, Ece Sonmezler, Yoav Zehavi, Michael Jerdev, Alessandra Balduini, Orsetta Zuffardi, Rita Horvath, Hanns Lochmüller, Anita Rauch, Livia Garavelli, Elisabeth Tournier-Lasserve, Ronen Spiegel, James R. Lupski, and Edoardo Errichiello

Supplemental Note

Supplemental clinical reports

Family A

Individual 1

The individual is a 10-year-old girl born at term of pregnancy (40 weeks) by spontaneous delivery with normal neonatal parameters (weight 3570 g, length 49 cm, head circumference 34 cm) and normal Apgar score (9 at 1 minute, 10 at 5 minutes). Parents were healthy and declared first cousins from the same geographic area in Turkey (Golbasi Adiyaman). The prenatal ultrasound scans were regular except for the last trimester check when a borderline (13 mm) bilateral ventricular dilatation was noted. Brain ultrasound performed at birth confirmed the ventricular dilatation associated with irregular ventricular edges, hypoplastic corpus callosum ad hyperechogenic spots located in the third ventricle and midline structures. Brain MRI and angio-MRI revealed a wide necrotic area located at the left temporo-parietooccipital regions with compression on the posterior-inferior areas of the lateral ventricle, slight dilatation of the third ventricle and narrowing of the distal portion of the aqueduct. In the following days, individual's reactivity, tone, and spontaneous motility worsened. Subsequent ultrasound scans and MRI demonstrated an increase in size of the lateral ventricles and brain lesions, specifically intraparenchymal hemorrhagic ones located in the left parieto-occipital region as well as a secondary hemorrhagic lesion in the right occipital-parietal region. Further hemorrhagic spots appeared in the right occipital and left frontal areas. The ophthalmological evaluation highlighted the presence of tortuosity of the perimacular retinal vessels and bilateral hemorrhages. Hydrocephalus was treated by shunting at 30 days of life. Seizures (blinking, chewing, and right sided clonus) occurred from 20 days of age. She is currently diagnosed with drug-resistant focal epilepsy (treated with carbamazepine, levetiracetam and clobazam). Ophthalmologic evaluation revealed worsening of the vascular anomalies, with increased tortuosity and hemorrhages, and appearance of ghost vessels. Fluorescein angiography highlighted wide areas of retinal ischemia, which was also found in the central retina, involving the macula of both eyes. Thus, laser treatment of the ischemic areas of the retina was performed to prevent fibrovascular proliferation. The individual never reached normal developmental milestones. Currently she suffers from spasticity and absence of speech. She has been PEG-fed from 2 years of age. She developed microcephaly and short stature over time. At the last evaluation, at 9 years and 8 months of age, weight was 27 kg (15^{th} percentile), height 124 cm (3^{rd} percentile) and OFC 45 cm (-5.7 SD). Furthermore, at clinical evaluation various dysmorphisms were noted, including bitemporal narrowing, high-arched and thick eyebrows, synophrys, up-slanting palpebral fissures, strabismus, bulbous nasal tip, full cheeks, diffuse hypertrichosis. She has been treated with multiple antiepileptic drugs but only achieving suboptimal seizure control. Recently, she has been admitted to the Pediatric Unit for recurrent episodes of pancreatitis with no clear etiology (re-analysis of the exome data did not identify any causative variants in pancreatitis-related genes). Chromosomal microarray performed at the age of 6 months was normal. Trio-based exome sequencing revealed the homozygous c.115del (Arg39Glyfs*33) frameshift variant in the *ESAM* gene. The healthy consanguineous parents resulted carrier of the same variant. AOH analysis revealed the variant was within a ~4 Mb AOH block on chromosome 11.

Individual 10

Individual 10 is the product of conception (POC) of the third pregnancy of consanguineous parents, sibling of individual 1. Prenatal ultrasound with echocardiography performed at 30 weeks of pregnancy identified cardiomegaly with dilatation of the atrium and right ventricle, tricuspid

insufficiency, pulmonary artery hypoplasia with valvular atresia, and reduced aortic diameter; in addition, small lung volume, cystic hygroma of the neck and talipes equinovarus were also noted. Cephalic biometry was in the lower limits and a massive cerebral edema was documented. Because of these clinical features, amniocentesis was performed with normal karyotype (46, XY) and array-CGH results. The pregnancy resulted in intrauterine death at 31 weeks of gestation. The autopsy highlighted diffuse periventricular leukomalacia with white matter gliosis and confirmed the previously observed heart abnormalities. After molecular diagnosis in the affected sister (individual 1), Sanger sequencing was carried out on the autopsy tissue samples (FFPE), which confirmed the presence of the same homozygous variant of the *ESAM* gene.

Family B

Individual 2

The index individual is a male individual born at term (37th gestational week) via normal spontaneous vaginal delivery after an unremarkable pregnancy and delivery. His birth measurements were as follows: weight 2150 g, height 44 cm, and head circumference 35 cm. He was the first child of a consanguineous union (Turkey). He had one female affected younger sibling (individual 3) and a healthy younger female sibling. He started having seizures at one month of life. He was last evaluated at 4 years 9 months, when anthropometric measurements revealed weight 12 kg (-3.2 SD), height 100 cm (6th percentile), head circumference 47 cm (-2.7 SD). He had severe global developmental delay with no meaningful words or walking and central hypotonia. His reflexes were increased globally with spastic tetraparesis. He had facial dysmorphic features including triangular face, smooth philtrum, and prominent chin. Brain MRI showed cerebral atrophy, dilated lateral ventricles with hydrocephalus *ex-vacuo*, thin corpus callosum, periventricular leukomalacia and abnormal T1 signal increase suggestive of hemorrhage. Ophthalmologic evaluation revealed bilateral retinal detachment and iris coloboma. Quad exome sequencing on both affected siblings (individuals 2 and 3) and unaffected parents revealed the homozygous c.115del (p.Arg39Glyfs*33) frameshift variant of *ESAM* in both affected individuals. Parents were heterozygous for the same variant. AOH analysis revealed the variant was within a 10.6 Mb AOH block shared in both affected siblings.

Individual 3

Individual 3 is the affected sister of individual 2. She was born at 35th gestational week via C-section with a birth weight of 2730 g, height 47 cm, and head circumference 35 cm. She developed first seizures in the infantile period. Anthropometric measurements at last evaluation (2 years 3 months) showed: weight 7.6 kg (-4.3 SD), height 78 cm (-2.8 SD) and head circumference 43.5 cm (-2.9 SD). She shared the same facial dysmorphic features of the affected brother, including triangular face, smooth philtrum, and prominent chin. Speech production was completely absent. She was hypotonic and her reflexes were increased. Brain MRI showed cerebral atrophy, dilated lateral ventricles with hydrocephalus *ex-vacuo*, thin corpus callosum, periventricular leukomalacia and abnormal T1 signal increase suggestive of hemorrhage. Ophthalmologic evaluation, performed on the 2nd postnatal day, revealed arterial tortuosity, venous dilatation, peripheral avascularity, retinal ischemia and severe extraretinal fibrovascular proliferation that induces preretinal and vitreous hemorrhage.

Family C

Individuals 4 and 5 are two siblings (male and female, respectively) who manifested severe prenatal cerebral hemorrhages and after birth showed, among main clinical features, spastic tetraparesis, epilepsy, hydrocephalus, dysphagia, respiratory failure, and cortical blindness.

Individual 4

At month eight of gestation, fetal MRI revealed post-hemorrhagic hydrocephalus secondary to germinal intraventricular matrix bleeding in choroid plexus, associated with hemorrhagic infarction. Cerebral MRI three days after birth documented hemorrhage of the germinal-intraventricular matrix grade IV bilateral (right predominance) with supratentorial post-hemorrhagic hydrocephalus. The largest parenchymal bleeding was right frontal and there was an area of cystic encephalomalacia (EMQ) with bleeding in the right upper parietal interior, probably communicating with the ventricular system. Cerebellum and neurohypophysis showed no alterations. Ultrasound fifteen days after birth also detected a left frontal EMQ zone with traces of bleeding-clot inside, communicating with the ipsilateral frontal horn, and a right parietal EMQ zone with traces of bleeding-clot inside and with communication with the ipsilateral lateral ventricle. Additionally, the ultrasound highlighted an alteration of the echogenicity of the right frontal periventricular white matter compatible with already known hemorrhage, traces of bleeding in occipital antlers and hydrocephalus affecting the lateral ventricles with extensive intraventricular bleeding.

Individual 5

During the second pregnancy, ventriculomegaly was detected at 26 weeks of gestation. At that time, MRI documented impaired cortico-subcortical differentiation and altered signal intensity of the white matter, suggestive of left frontoparietal extensive hemorrhagic infarction, and extra-axial bleeding.
Quad exome sequencing on both affected siblings (individuals 4 and 5) and unaffected parents detected the homozygous c.115del (p.Arg39Glyfs*33) frameshift variant in both affected siblings. Parents were heterozygous for the same variant, consistent with recessive disease traits. AOH analysis revealed the variant was within a ~2 Mb AOH block shared in both affected siblings.

Family D

Individual 6

The proband was a male individual from a third spontaneous pregnancy of healthy consanguineous parents of Algerian descent. He was born at term (37 weeks of gestation) via normal spontaneous vaginal delivery. At 3rd trimester ultrasound examination, a cerebral parenchymal destruction was observed and confirmed at birth by cerebral MRI. At birth, anthropometric measurements were as follows: weight 3040 g, height 46.5 cm and head circumference 35.5 cm. Cerebral MRI showed massive dilatation of lateral ventricles with global cerebral parenchymal destruction corresponding to hydranencephaly, diffuse meningeal and intraventricular hemorrhage with clotting, and focal destruction of the septum pellucidum. He had severe global developmental delay, spastic tetraparesis and epilepsy. Ophthalmologic evaluation revealed bilateral retinal hemorrhages. Due to the very poor prognosis, it was decided to initiate supportive palliative care with a progressive increase of morphine adapted to the pain of the symptoms. He died at 13 months of age. Trio-based exome sequencing revealed a homozygous nonsense variant (c.35T>A) in *ESAM* leading to a premature stop codon (p.Leu12*). Parents were carrier of the same variant. AOH analysis revealed the variant was located within a ~5.2 Mb AOH block.

Individual 11

Individual 11 is the affected sibling of individual 6 (first pregnancy of the couple). A parenchymal destruction with ventricular dilatation was noted at 21 weeks of gestation by ultrasound examination. Fetal MRI showed intraventricular hemorrhage involving lateral ventricles associated with diffuse intraparenchymal, periventricular hemorrhages and cerebral calcifications. Given the very poor prognosis, a medical termination of pregnancy (TOP) was proposed and accepted by the parents at 32 weeks of gestation. The parents denied their consent for post-mortem examination and, thus, no biological material was available for confirmatory genetic testing. The second pregnancy of the couple lead to a spontaneous abortion at 17 weeks of gestation.

Family E

Individual 7

Individual 7 was born to healthy consanguineous parents (first cousins originating from Kahramanmaras, Turkey) by Cesarian section at 39 weeks gestation. His birth weight was 3900 g and head circumference was 37.5 cm.

The neonatal period was described as unremarkable. The mother first developed concerns at age 4 months when head control was not established. The child was noted to be hypotonic and developmentally delayed; EEG recordings at this age were also abnormal. Further investigations at the age of 6 months included a CT scan, which documented diffuse cortical atrophy and sub-ependymal, periventricular and lenticular nucleus calcifications, suggesting cerebral microangiopathy. Based on these findings, AicardiGoutier syndrome (AGS) was suspected as differential diagnosis, but pleiocytosis was excluded in cerebral fluid and alpha interferon was normal. The child developed generalized tonic-clonic seizures at the age of 2 years. At last clinical evaluation (4 years of age), he showed severe developmental delay (with sitting, walking and speech never achieved), severe axial hypotonia, spastic tetraparesis, and progressive macrocephaly. His growth parameters were as follows: weight 16.5 kg (54^{th} percentile), length 98.5 cm (19^{th} percentile), head circumference 56.7 cm (+4.7 SD). At physical examination, several facial dysmorphisms were noted, including wide nasal bridge, flat philtrum, upslanted palpebral fissures, and broad forehead with bitemporal narrowing. Individual 7 was part of a study including 63 unrelated individuals with early-onset developmental and/or epileptic encephalopathy and remained without etiological diagnosis (Individual 73705 reported in Papuc et al., 2019). Trio-based exome sequencing revealed a rare homozygous variant (c.115del, p.Arg39Glyfs*33) in the *ESAM* gene. Both healthy parents were heterozygous for the same variant. SNP-array analysis revealed the variant to be located within a ~14 Mb AOH block.

Family F

Individual 8

The individual was born by vaginal delivery at 38 weeks of gestation from healthy consanguineous parents of Turkish descent. Two previous spontaneous abortions were recorded in the family history, although without further clinical details. The proband also has an older healthy brother. After birth, the proband remained in the neonatal intensive care unit for 27 days and required 2 days assisted ventilation due to respiratory distress. He presented with motor retardation and febrile seizures since the age of 16 months. He had no head control, and he was quadriplegic. The seizure semiology was in the form of blinking in the eyes and contractions in the whole body for 30 minutes. Therefore, Levetiracetam, Valproic acid and Clobazam treatments were started since seizures also manifested during afebrile periods. The EEG showed a generalized pattern. The individual showed right hemiparesis with no head control. He received full oral feeding with occasional coughing. He had rotatory nystagmus and high-arched palate. Neurological examination documented joint contractures (severe grade at ankles and knee), severe lower extremities spasticity, hyperactive deep tendon reflexes, hammer heels, dystonia of the hands and big toe.

Family G

Individual 9

The individual is a 19 months-old male born to healthy parents who are first-degree cousins of Arab Bedouin descent. Pregnancy was reported as unremarkable. In particular, the second trimester ultrasound scan was normal. He was born at term by elective Cesarean section. His Apgar scores were normal (8/9), and birth parameters were: weight 3140, height 49 cm and head circumference 35 cm. At the age of 3 months, the child presented poor feeding, low weight, truncal hypotonia with markedly head lag and dystonic postures, and limb hypertonia. He had abnormal eye contact and interrupted

pursuit. In addition, he showed facial dysmorphic features including bitemporal narrowing, bulbous nasal tip and low posteriorly rotated ears. Cranial ultrasound revealed marked ventriculomegaly. Brain MRI at 5 months showed prominent ventriculomegaly with external hydrocephalus compatible with cortical atrophy, thinning of the corpus callosum and delayed myelination. In addition, periventricular and cortical (mainly left frontal) hemosiderin tracers were evident suggesting previous intracranial hemorrhage.

Disease course was marked by severe psychomotor delay, inability to gain motor and cognitive milestones and gradual development of quadriparesis. At the age of 11 months, the individual manifested prolonged episode of focal seizures. Brain CT showed multiple periventricular calcifications in addition to ventriculomegaly. His EEG revealed right temporal spike and wave activity with secondary generalization. Thus, he started treatment with levetiracetam; nevertheless, he had other episodes of seizures requiring hospital admission.

On his last examination, at 18 months of age, the individual was unable to roll or sit; he was able to follow objects partially with saccadic pursuit and associated nystagmus. He still had head lag and inability to raise his head against gravity. He had limb hypertonia and dystonic postures tendon hyperreflexia with bilateral Babinski sign. He was able to feed orally but had severe failure to thrive and progressive microcephaly. Anthropometric parameters were as follows: weight 7.9 kg (-3.4 SD), length 73 cm (-2.9 SD), head circumference 45 cm (-2.1 SD). Extended metabolic studies and chromosomal microarray analysis were normal. On the other hand, trio-based exome sequencing revealed a homozygous splice variant (c.451+1G>A) in the *ESAM* gene. Both parents carried the same variant in the heterozygous state. AOH analysis revealed the variant was within a ~4 Mb AOH block.

Family H

Individual 12

Individual 12 is a female pregnancy achieved through *in vitro* fertilization (IVF) and pre-gestational testing (PGT). The parents are healthy and first-degree cousins of Arab Bedouin descent. They have two healthy daughters and a history of two pregnancy losses. Fetal sonography of individual 12 (the first affected fetus) was normal at 14 and 20 weeks of gestation. Ultrasound scan performed at 30 weeks of gestation revealed porencephaly and diffuse brain calcifications, thus the parents decided to terminate the pregnancy at 32 weeks of gestation. Post-mortem autopsy revealed normal gross appearance of the brain, but further histological examination revealed evidence of white matter lesions, especially ischemic hemorrhagic foci, and diffuse basal ganglia calcifications.

Individual 13

Individual 13 is the male sibling of individual 12. The pregnancy was terminated at 39 weeks of gestation due ultrasound scan and brain MRI findings like those of the previous TOP. Following the last pregnancy, trio-based exome sequencing was performed and identified a homozygous splice variant in the *ESAM* gene (c.451+1G>A). Both parents carried the same variant in the heterozygous state. SNP-array analysis revealed the variant was within a ~12 Mb AOH block shared by the two affected fetuses. Further segregation analysis by Sanger sequencing confirmed the presence of the c.451+1G>A variant in the autopsy material obtained from the second affected fetus (individual 13), whereas the two healthy daughters were either heterozygous or wild type.

Supplemental figures and legends

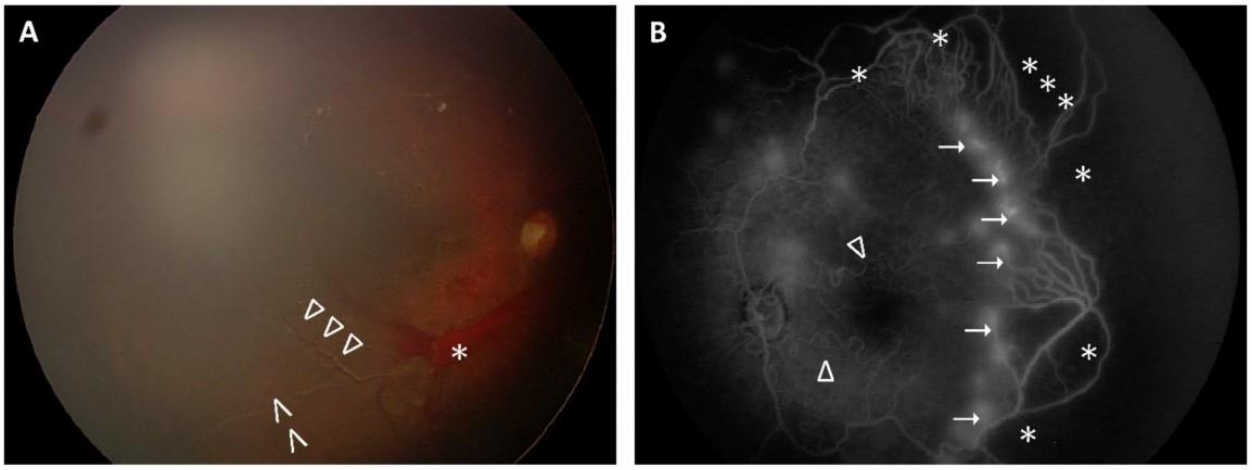


Figure S1. Retinal abnormalities in individual 1. (A) Color photography of the fundus oculi (right eye), showing retinal hemorrhages (∗), vascular tortuosity (ΔΔΔ) and ghost vessels (ΛΛ), i.e., occluded retinal vessels. (B) Fluorescein angiography of the retina (left eye), showing marked vascular tortuosity (Δ), retinal ischemia (∗) and new vessels (→).

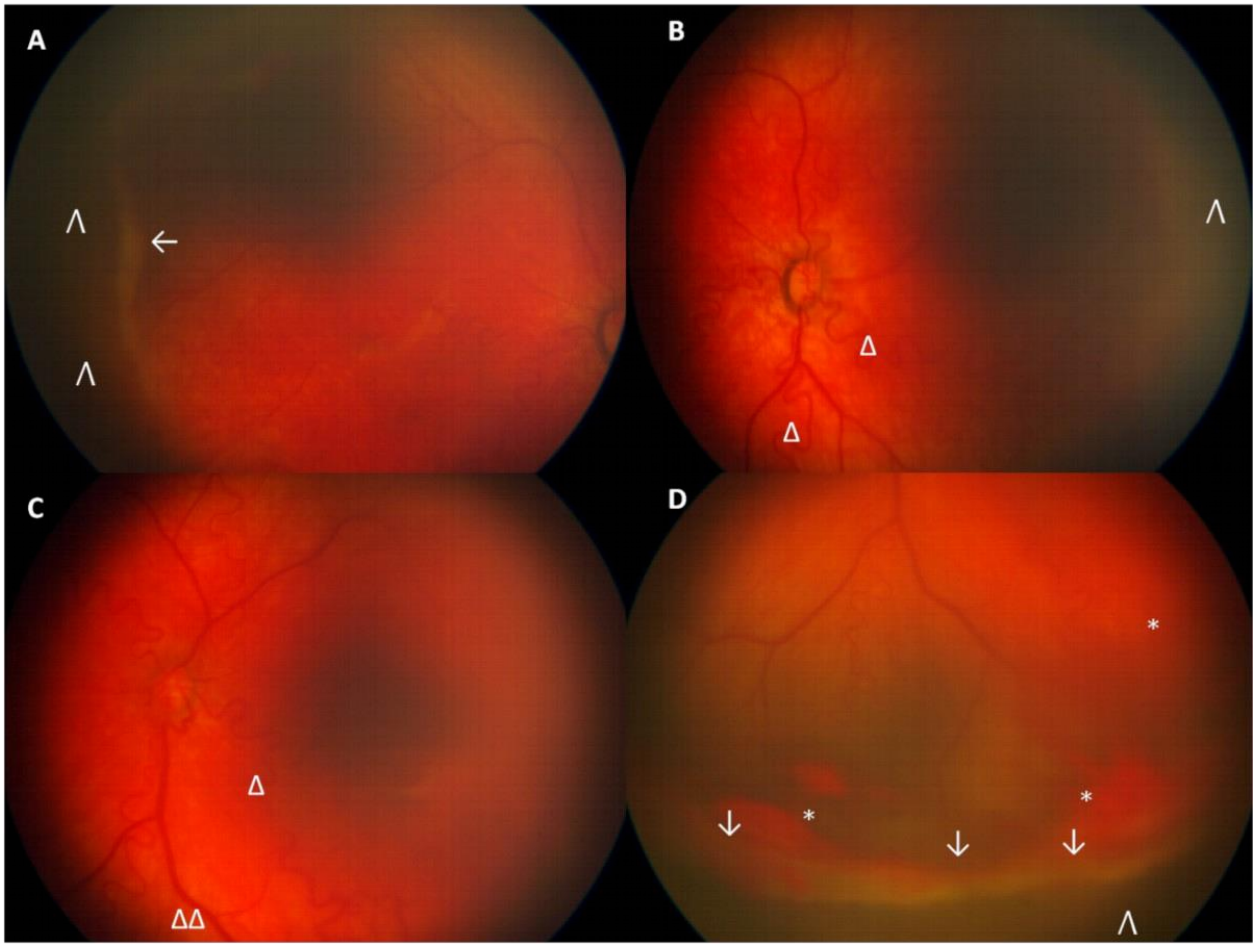


Figure S2. Retinal abnormalities in individual 3. Wide-field fundus photos captured on the 2nd postnatal day of both eyes (A, B: right eye; C, D: left eye) of individual 3 presenting venous and arterial dilatation, arterial tortuosity, and severe extraretinal fibrovascular proliferation. These are frequently encountered findings in the retinopathy of prematurity. (Δ) remarks arterial tortuosity, (ΔΔ) remarks venous dilatation, (Λ) remarks peripheral avascularity and ischemia, (∗) remarks preretinal and vitreous hemorrhage, (←) and (↓) remarks extraretinal fibrovascular proliferation.

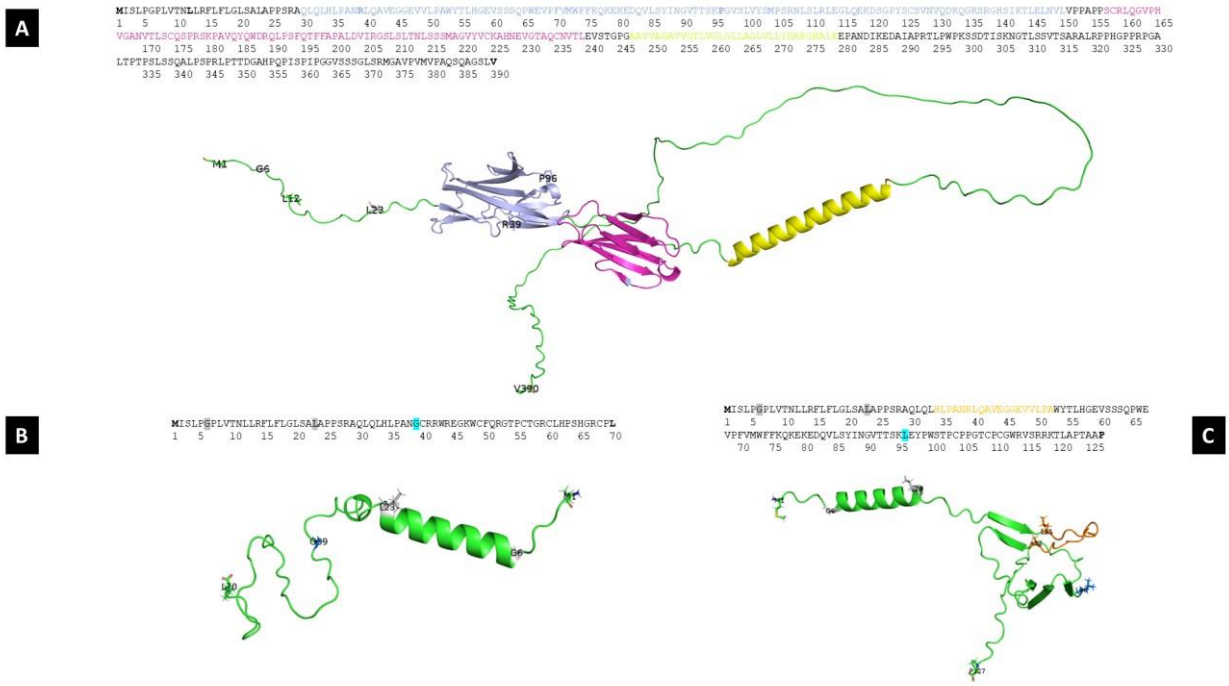


Figure S3. 3D models of mutant ESAM protein products. (A) Wild-type ESAM 3D structure: the first Iglike domain (aa 30-150) is indicated in violet, the second Ig-like domain (aa 157-239) in purple, and the helical transmembrane domain (aa 247-279) in yellow. These domains are removed in the p.Arg39Glyfs*33 (B) and p.Pro96Leufs*33 (C) truncated protein products. AlphaFold predicted the formation of 3D motifs that are absent in the wild-type protein: 1) a helical domain (aa 6-23) in both p.Arg39Glyfs*33 (B) and p.Pro96Leufs*33 (C) mutants and 2) a few non-contiguous β-sheets, separated by a disordered linker region (aa 35-52), in the p.Pro96Leufs*33 mutant (C). AlphaFold failed to predict the 3D structure of the p.Leu12*mutant because of minimal aminoacid input sequence. The predicted 3D motifs of p.Arg39Glyfs*33 and p.Pro96Leufs*33 mutants did not match with any known protein signature explored through different databases and bioinformatic tools (e.g., Motif Scan, ScanProsite, BlastP).

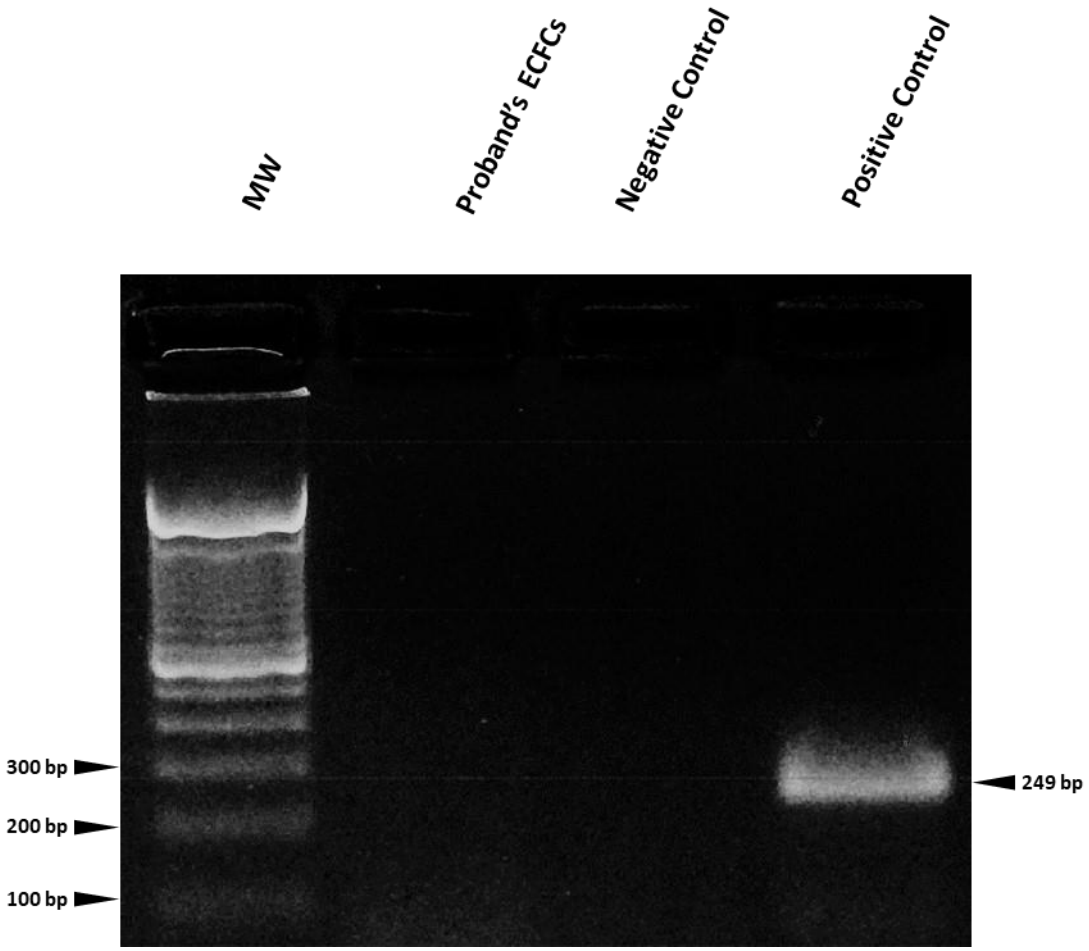


Figure S4. Reverse transcription polymerase chain reaction (RT-PCR) in endothelial colony forming cells (ECFCs) of individual 1. *ESAM* expression was tested by RT-PCR with specific primers designed across exon 1-3 junctions to avoid genomic DNA amplification, generating a 249-bp amplicon. No detectable cDNA product was observed in the proband's ECFCs in respect to a positive control. MW: 100-bp allelic ladder. Negative control: no template control (NTC); Positive control: age- and sex-matched control sample with wild-type *ESAM*. Experiments were performed as previously described with minor modifications
(Errichiello et al., 2017).

Supplemental methods

Genetic analysis and variant assessment

Candidate *ESAM* variants were identified by ES on genomic DNA samples extracted from peripheral blood lymphocytes of the proband and parents (trio ES) or with additional family members (quad ES). Variants were validated by Sanger dideoxy sequencing and tested for co-segregation in available family members who had not undergone ES. Variant frequency in control populations was assessed through the Genome Aggregation Database (gnomAD v.2.1.1) (https://gnomad.broadinstitute.org/) and TOPMed/BRAVO browser (https://bravo.sph.umich.edu/freeze8/hg38/), as well as by using in-house databases. Variants were classified according to the recommendations of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG/AMP) and Association for Clinical Genomic Science (ACGS) best practice guidelines (Richards et al., 2015; Ellard et al., 2020). Notably, the pathogenic PVS1 criterion was applied to predicted loss-of-function (LoF) variant alleles following the recommendations of the ClinGen Sequence Variant Interpretation (SVI) Workgroup's (Abou Tayoun et al., 2018). *ESAM* variants are described according to the reference sequence NM_138961.3. The presence of potential pathogenic CNVs (Copy Number Variations) was assessed by either chromosomal microarray analysis (CMA) and/or computational tools for CNV detection based on ES data. Regions/runs of homozygosity (ROH), also known as absence of heterozygosity (AOH), were detected by either SNP-based CMA or ES data thought specific computational pipelines. Three-dimensional (3D) models of protein structures were predicted with AlphaFold v2 (https://alphafold.ebi.ac.uk/) and visualized by using PyMOL (https://pymol.org/2/).

In vitro tube formation assay

Blood samples (40 mL) from individual 1 (index case) and her healthy biological parents were collected in EDTA-containing tubes. Endothelial colony-forming cells (ECFCs) were isolated from circulating mononuclear cells as previously described (Zuccolo et al., 2018). Early passage ECFCs (P2-P3, 15-18 days) were cultured in basal medium EBM-2 supplemented with 2% FBS in Cultrex (Trevigen)-coated 96-well plates. Capillary network formation was assessed starting from 4 up to 24 hours later by evaluating both dimensional and topological parameters. The length of endothelial tube-like structures, the number of polygon structures established by tubules, referred to as “meshes” and indicative of endothelial cell migration, and the number of master junctions were measured from acquired bright-field pictures by using the Angiogenesis Analyzer plugin of ImageJ, as previously reported (Balducci et al., 2021). Micrographs were captured by using an Olympus IX71-inverted microscope (Olympus) equipped with a CPlan F1 10 ×/0.30 objective. The formation of capillary networks was quantitatively evaluated by measuring the average tubular number in 5 serial microscopic fields (HPF) and total capillary tube length (µm) in 5 view fields using LC Micro software version 5.1 (Olympus). Data were collected from 3 independent experiments and statistically significant differences were evaluated by one-way ANOVA and Tukey post-hoc tests.

ESAM immunohistochemistry

Four micrometer thick whole slide sections were prepared from archival paraffin blocks of formalin-fixed (FFPE) fetal brain of individual 10 and an age- and sex-matched brain tissue control. Preliminary validation of two rabbit polyclonal anti-ESAM antibodies (ab73988 and ab74777; Abcam, Cambridge, UK) was performed on fetal kidney tissues where *ESAM* is also highly expressed. The ab73988 gave optimal performance results, in terms of sensitivity and specificity, and was thus adopted for experiments on brain tissues. Unstained FFPE sections were incubated with ab73988 diluted 1:200 and counterstained with hematoxylin and eosin. The reactions were revealed by using the peroxidase-conjugated biotinstreptavidin complex method (Dako LSAB2 System-HRP; Dako Denmark A/S, Glostrup, Denmark) and 3,3'diaminobenzidine-tetrahydrochloride-dihydrate (DAB) as chromogenic

substrate. Each reaction included a negative control obtained by substituting the primary antibody with equal amount of dilution buffer.

Supplemental references

Papuc, S.M., Abela, L., Steindl, K., Begemann, A., Simmons, T.L., Schmitt, B., Zweier, M., Oneda, B., Socher, E., Crowther, L.M., et al. (2019). The role of recessive inheritance in early-onset epileptic encephalopathies: a combined whole-exome sequencing and copy number study. Eur J Hum Genet. 27, 408-421. 10.1038/s41431-018-0299-8

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al. (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. Genet Med. *17*, 405-424. 10.1038/gim.2015.30

Ellard, S., Baple, E.L., Callaway, A., Berry, I., Forrester, N., Turnbull, C., Owens, M., Eccles, D.M., Abbs, S., Scott, R., et al. (2020). ACGS Best Practice Guidelines for Variant Classification in Rare Disease 2020. https://www.acgs.uk.com/media/11631/uk-practice-guidelines-for-variant-classification-v4-01-2020.pdf

Abou Tayoun, A.N., Pesaran, T., DiStefano, M.T., Oza, A., Rehm, H.L., Biesecker, L.G., Harrison, S.M.; ClinGen Sequence Variant Interpretation Working Group (ClinGen SVI). (2018). Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. Hum Mutat. *39*, 1517-1524. 10.1002/humu.23626

Zuccolo, E., Di Buduo, C., Lodola, F., Orecchioni, S., Scarpellino, G., Kheder, D.A., Poletto, V., Guerra, G.,
Bertolini, F., Balduini, A., et al. (2018). Stromal Cell-Derived Factor-1α Promotes Endothelial ColonyForming Cell Migration Through the Ca^{2+}-Dependent Activation of the Extracellular Signal-Regulated Kinase 1/2 and Phosphoinositide 3-Kinase/AKT Pathways. Stem Cells Dev. *27*, 23-34. 10.1089/scd.2017.0114

Balducci, V., Faris, P., Balbi, C., Costa, A., Negri, S., Rosti, V., Bollini, S., and Moccia, F. (2021). The human amniotic fluid stem cell secretome triggers intracellular Ca^{2+} oscillations, NF-κB nuclear translocation and tube formation in human endothelial colony-forming cells. J Cell Mol Med. *25*, 8074-8086.
10.1111/jcmm.16739

Errichiello, E., Mustafa, N., Vetro, A., Notarangelo, L.D., de Jonge, H., Rinaldi, B., Vergani, D., Giglio, S.R., Morbini, P., and Zuffardi, O. (2017). SMARCA4 inactivating mutations cause concomitant Coffin-Siris syndrome, microphthalmia and small-cell carcinoma of the ovary hypercalcaemic type. J Pathol. 243, 9-
15. 10.1002/path.4926

Supplemental acknowledgements

This work was supported by grants of the Italian Ministry of Health to RC (2022-2024). Individuals 1 and
10 were recruited in the framework of the European Reference Network for Rare Congenital Malformations and Intellectual Disability (ERN-ITHACA) [EU Framework Partnership Agreement ID: 3HPHP-FPA ERN-01-2016/739516]. E.E. was supported by “Fondo Ricerca e Giovani (FRG)” of the University of Pavia. The studies on individuals 2 and 3 were supported in part by the U.S. National Human Genome Research Institute (NHGRI) and National Heart Lung and Blood Institute (NHBLI) to the Baylor-Hopkins Center for Mendelian Genomics (BHCMG, UM1 HG006542, J.R.L); U.S. National Institute of Neurological Disorders and Stroke (NINDS) (R35NS105078 to J.R.L.). NHGRI U01 HG011758 to the Baylor College of Medicine Genomics Research Elucidate Genetics of Rare disease (BCM-GREGoR) consortium and grant K08 HG008986 to J.E.P.; D.P. was supported by Rett Syndrome Research Trust, and NINDS (1K23 NS125126-01A1). D.M. was supported by a Medical Genetics Research Fellowship Program through the United States National Institute of Health (T32 GM007526-42). T.M. was supported by the Uehara Memorial Foundation. The studies on individuals 4 and 5 were supported in part by the Carlos III Institute of Health (ISCIII), Ministry of Economy and Competitiveness (Spain) (grant PI18/00847; D.H.S. was funded by the Carlos III Institute of Health (ISCIII), Ministry of Economy and Competitiveness (Spain) (grant PI18/00847). J.H.R. was supported by an IDISBA fellowship program through the AETIB annual plan 2019 (Folium Program-INTRES Project). C.V.B. was funded by the Carlos III Institute of Health (ISCIII), Ministry of Economy and Competitiveness (Spain) (grant PI21/00890). The studies on individuals 6 and 11 were partly supported by the National Institutes of Health (R01NS096173 grant) and by the Leducq Foundation Transatlantic Network of Excellence grant “ReVAMP”. The studies on individual 7 were in part supported by the University of Zurich clinical research priority program praeclare and by the Swiss National Science Foundation grant 320030_179547 to A.R.. S.H. and Y.O. were supported by TUBITAK (The Scientific and Technological Research Council of Turkey) Project No. 216S771. Y.O. was supported by the Turkish Academy of Sciences’ Young Investigator award, TUBA-GEBIP. The studies on individual 8 were partly supported by the NIHR Cambridge Biomedical Research Centre (BRC-1215-20014). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. R.H. and Y.O. were supported by an MRC strategic award to establish an International Centre for Genomic Medicine in Neuromuscular Diseases (ICGNMD) MR/S005021/1. R.H. is a Wellcome Trust Investigator (109915/Z/15/Z), who received support from the Medical Research Council (UK) (MR/N025431/1 and MR/V009346/1), the European Research Council (309548), the Newton Fund (UK/Turkey, MR/N027302/1), the Addenbrookes Charitable Trust (G100142), the Evelyn Trust, the Stoneygate Trust, the Lily Foundation. H.L. received support from the Canadian Institutes of Health Research (Foundation Grant FDN-167281), the Canadian Institutes of Health Research and Muscular Dystrophy Canada (Network Catalyst Grant for NMD4C), the Canada Foundation for Innovation (CFI-JELF 38412), and the Canada Research Chairs program (Canada Research Chair in Neuromuscular Genomics and Health, 950-232279). The studies on individuals 12 and 13 were party funded through a pilot project of the Israeli Ministry of Health.

DISCUSSION

L'homéostasie cérébrale est conditionnée par une intégrité de la barrière hémato-encéphalique (BHE). Les cellules endothéliales sont des éléments structurels importants de cette barrière et sont reliées entre elles par des jonctions serrées permettant de créer une barrière physique et métabolique entre le contenu vasculaire et le SNC. Le développement de la BHE commence durant la période embryonnaire avec la formation des jonctions entre les cellules endothéliales (Anstrom *et al.*, 2007). Une altération de ces molécules pourrait entraîner une fragilité de la matrice germinale et une propension aux hémorragies chez le fœtus.

Dans cet article, il est décrit pour la première fois la présence de variants disruptifs biallélіques du gène *ESAM* (molécule de jonctions serrées) chez des fœtus et enfants présentant une HIC anténatale. Au niveau phénotypique, les patients porteurs de variants biallélіques d'*ESAM* présentaient tous une ventriculomégalie, une hémorragie intracérébrale sévère d'origine anténatale associée ou non à des calcifications intracrâniennes. Les enfants survivants présentaient un retard développemental global avec une déficience intellectuelle, une épilepsie et un retard d'acquisition de la parole. Un immunomarquage anti-ESAM sur cerveau d'un des fœtus porteur d'un variant homozygote c.115del (p.R39Gfs*33), ne montrait pas de marquage ESAM dans les cellules endothéliales capillaires, contrastant avec une expression intense d'ESAM dans l'endothélium cérébral d'un échantillon témoin apparié ; ces données confirmaient l'expression d'ESAM au niveau des capillaires cérébraux et la nature perte de fonction attendue du variant. Une étude sur la capacité de progéniteurs endothéliaux à former des structures de type capillaire *in vitro* a également été effectuée chez un individu et a montré des anomalies de prolifération, de migration et de tubulogénèse.

Les caractéristiques phénotypiques observées chez ces individus mutés *ESAM* sont proches de ceux rapportés pour d'autres maladies associées à des variations pathogènes d'autres gènes codant pour des molécules des jonctions serrées endothéliales à savoir *JAM2, JAM3, OCLN* et *CLDN5*. Cependant, avant la description des variations pathogènes d'*ESAM*, la présence d'une hémorragie intracrânienne n'était connue que pour le gène *JAM3* (Tableau 4).

	JAM2	***JAM3***	***OCLN***	***ESAM***	***CLDN5***
Transmission	AR	AR	AR	AR	AD
Age de début	Adulte	Anténatal Néonatal	Anténatal Néonatal	Anténatal Néonatal	Petite enfance
Anomalies cérébrales observées					
Microcéphalie	Oui	Oui	Oui	Oui	Oui
Calcifications intracrâniennes	Oui	Oui	Oui	Oui	Oui
Hémorragie intracrânienne	Non	Oui	Non	Oui	Non
Ventriculomégalie	Non	Oui	Oui	Oui	Non
Anomalies du corps calleux	Non	Oui	Oui	Oui	Non

Tableau 4. Principales caractéristiques des pathologies associées aux variations pathogènes des protéines des jonctions endothéliales

Les mécanismes conduisant à la survenue d'une hémorragie cérébrale dépendant du gène muté ne sont pas connus. En interrogeant une base de données sur l'expression des gènes dans les cellules vasculaires cérébrales de la souris (Vanlandewijck *et al.*, 2018 ; He *et al.*, 2018), nous observons un profil d'expression vasculaire différent pour ces protéines (Figure 28). Les protéines ESAM et JAM3 ont une expression au niveau des péricytes et des cellules musculaires lisses, ce qui n'est pas le cas pour les protéines JAM2, OCLN et CLDN5. Cette différence d'expression pourrait expliquer la variabilité phénotypique associée aux variations pathogènes des protéines jonctionnelles.

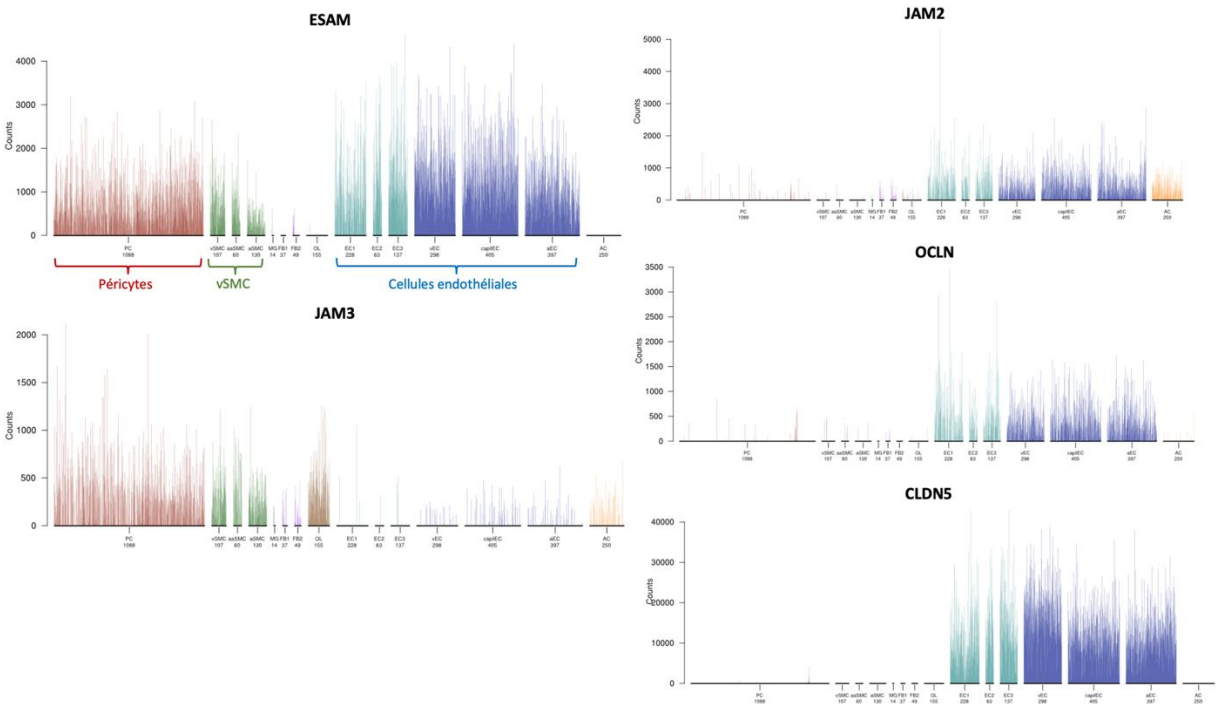


Figure 28. Profil d'expression des molécules de jonction serrées dans les cellules vasculaires cérébrales de la souris. vSMC = cellules musculaire lisses.
D'après https://betsholtzlab.org/VascularSingleCells/database.html.

Par ailleurs, après un AVC hémorragique, il a été montré dans une cohorte d'une vingtaine de patients adultes, des taux significativement élevés des molécules des jonction serrées (de Claudin-5, d'Occludine et de zonula occludens 1 (ZO-1)) dans le LCR de ces patients, témoins de rupture de la cohésion des cellules endothéliales (Jiao *et al.*, 2015). L'étude des molécules jonctionnelles est d'une grande importance pour le développement de thérapeutiques afin de prévenir un dysfonctionnement de la BHE et de restaurer sa fonction normale après une hémorragie cérébrale. Un certain nombre d'approches thérapeutiques ont montré leur efficacité dans la réduction de la rupture de la BHE et des modifications des jonctions serrées après une HIC dans des modèles précliniques. Cependant, il n'est pas certain que ces effets s'exercent directement sur l'endothélium cérébral ou s'il s'agit d'un effet indirect par réduction du nombre de lésions hémorragiques. Des agents qui ciblent spécifiquement l'endothélium et les protéines jonctionnelles n'ont pas été examinés (Keep *et al.*, 2018).

D- Maladie métabolique rare mimant le phénotype cérébral associé aux gènes *COL4A1/COL4A2* chez le fœtus

INTRODUCTION

L'analyse de nos données d'exome entier met en évidence en moyenne entre 150 et 200 variants hétérozygotes non synonymes rares (MAF < 1/10 000) chez un individu. L'analyse individuelle de chaque variant afin d'explorer une causalité potentielle est impossible dans une étude portant sur plus d'une centaine de sujets mais une étude d'enrichissement en variants rares à l'échelle du gène permet de classer ces derniers selon leur charge en variants rares et prédits délétères, et de comparer cette charge avec celles observées dans des cohortes contrôles appariées.

Nous avons ainsi réalisé un test de charge à l'échelle du gène basé sur l'ensemble des données de WES des 113 fœtus index. Avec les critères de sélection des variants candidats choisis, un gène présentait un enrichissement significatif (*genome-wide*) en variants candidats. Il s'agissait du gène *PDHA1* pour lequel trois fœtus de sexe féminin étaient porteurs de variants dont la nature a permis d'établir la causalité.

Les résultats sont décrits dans l'article ci-après.

Ultrasound Obstet Gynecol 2022; **60**: 805–811
Published online in Wiley Online Library (wileyonlinelibrary.com). **DOI:** 10.1002/uog.26046

Rare metabolic disease mimicking *COL4A1/COL4A2* fetal brain phenotype

T. COSTE[1,2], C. ALOUI[1#], F. PETIT[3#], S. MOUTTON[4#], L. DEVISME[5], C. F. WELLS[6], N. LEBOUCQ[7], P. VERPILLAT[8], M. YVERT[4], F. RIVIER[9*] and E. TOURNIER-LASSERVE[1,2*]

[1]*Université Paris Cité, Inserm, NeuroDiderot, Paris, France;* [2]*Service de Génétique Moléculaire Neurovasculaire, AP-HP, Hôpital Saint-Louis, Paris, France;* [3]*CHU Lille, Clinique de Génétique Guy Fontaine, Lille, France;* [4]*Centre Pluridisciplinaire de Diagnostic Prénatal, MSP Bordeaux Bagatelle, Talence, France;* [5]*CHU Lille, Institut de Pathologie, Lille, France;* [6]*CHU Montpellier, Département de Génétique Médicale et Fœtopathologie, Montpellier, France;* [7]*CHU Montpellier, Département de Neuroradiologie, Neuroradiologie Diagnostique Pédiatrique, Montpellier, France;* [8]*CHU Lille, Service de Radiologie, Lille, France;* [9]*CHU Montpellier, Département de Neurologie Pédiatrique, PhyMedExp, Université de Montpellier, INSERM, CNRS, Montpellier, France*

KEYWORDS: *COL4A1/COL4A2*; corpus callosal dysgenesis; exome sequencing; fetal intracerebral hemorrhage; *PDHA1*; ventriculomegaly

ABSTRACT

*Pathogenic variants of collagen type IV alpha 1 and 2 (*COL4A1/COL4A2*) genes cause various phenotypic anomalies, including intracerebral hemorrhage and a wide spectrum of developmental anomalies. Only 20% of fetuses referred for* COL4A1/COL4A2 *molecular screening (fetuses with a suspected intracerebral hemorrhage) carry a pathogenic variant in these genes, raising questions regarding the causative anomaly in the remaining 80% of these fetuses. We examined, following termination of pregnancy or* in-utero *fetal death, a series of 113 unrelated fetuses referred for* COL4A1/COL4A2 *molecular screening, in which targeted sequencing was negative. Using exome sequencing data and a gene-based collapsing test, we searched for enrichment of rare qualifying variants in our fetal cohort in comparison to the Genome Aggregation Database (gnomAD) control cohort (*n = 71 702*). Qualifying variants in pyruvate dehydrogenase E1 subunit alpha 1 (*PDHA1*) were overrepresented in our cohort, reaching genome-wide significance (*$P = 2.11 \times 10^{-7}$*). Heterozygous* PDHA1 *loss-of-function variants were identified in three female fetuses. Among these three cases, we observed microcephaly, ventriculomegaly, germinolytic pseudocysts, agenesis/dysgenesis of the corpus callosum and white-matter anomalies that initially suggested cerebral hypoxic-ischemic and hemorrhagic lesions. However, a careful* a-posteriori *reanalysis of imaging and postmortem data showed that the observed lesions were also consistent with those observed in fetuses carrying* PDHA1 *pathogenic variants, strongly suggesting that these two phenotypes may overlap. Exome sequencing should therefore be performed in fetuses referred for* COL4A1/COL4A2 *molecular screening which are screen-negative, with particular attention paid to the* PDHA1 *gene. © 2022 International Society of Ultrasound in Obstetrics and Gynecology.*

CASE SERIES

Exome sequencing was performed in 113 fetuses referred initially (with a suspected intracerebral hemorrhage) for collagen type IV alpha 1 and 2 (*COL4A1/COL4A2*) gene screening, in which no pathogenic variation was identified. A gene-based collapsing burden test comparing qualifying variants in the 113 fetuses *vs* the Genome Aggregation Database (gnomAD) public control database ranked the pyruvate dehydrogenase E1 subunit alpha 1 (*PDHA1*) gene first, showing genome-wide significance ($P = 2.11 \times 10^{-7}$). (For further details of methods, see Appendix S1.) Three qualifying variants of *PDHA1* were identified, one in each of three heterozygous female fetuses (Table S1, Figure 1). Two were loss-of-function variants and one was a missense variant. None of these variants was present in public control databases of polymorphisms (gnomAD and the 1000Genomes Project Phase 3 database). The p.(Arg304*) and p.(Arg119Trp) variants have been reported previously, in a large series of patients with pyruvate dehydrogenase complex deficiency (PDCD) (MIM# 312170)[1]. The main ultrasound findings and careful *a-posteriori* reanalysis of magnetic resonance

Correspondence to: Dr T. Coste, AP-HP, Service de Génétique Moléculaire Neurovasculaire, Hôpital Saint-Louis, Paris, France (e-mail: thibault.coste@aphp.fr)

#C.A., F.P. and S.M. contributed equally to this work. *F.R. and E.T.-L. are co-last authors.

Accepted: 28 July 2022

imaging (MRI) features of the affected fetuses are summarized in Table 1. This study was approved by the INSERM ethics review committee (IRB00003888) and informed consent was obtained from parents and for the fetuses included in the study.

Case 1

Fetus F1 was the first child of unrelated and healthy parents, without familial history. The first trimester of pregnancy was unremarkable. Ultrasound examination at 21 gestational weeks revealed cerebral anomalies, with moderate microcephaly, suspected corpus callosal agenesis, cerebellar hypoplasia, thickened ventricular walls and dilated frontal horns. At 23 weeks, there was intrauterine growth restriction, with an estimated fetal weight on the 3rd centile. Fetal cerebral MRI at 26 weeks confirmed the agenesis of the posterior corpus callosum, revealed agyria, periventricular subependymal germinolytic pseudocysts and vermian and cerebellar hypoplasia (Figure 2a–f). Hypersignal of the white matter and early leukomalacia were suspected. Signal anomalies of the frontal horns were

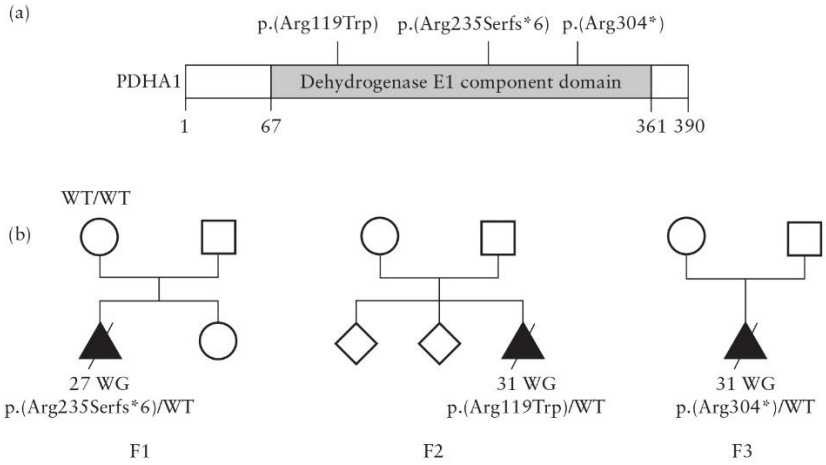


Figure 1 Location of pathogenic *PDHA1* variants and pedigrees of the three fetuses with qualifying mutations. (a) Protein structure and corresponding domain organization of human PDHA1 protein (protein NP_000275.1). Numbers below the figure represent the first and last coding amino acids for the only protein domain of *PDHA1*, and the last coding amino acid (AA 390). (b) Pedigrees of the three mutated probands. Males are represented by squares and females by circles, triangles indicate pregnancy not carried to term and diamonds indicate gender not specified. Filled symbols indicate affected individual (confirmed clinically and by magnetic resonance imaging) and empty symbols indicate clinically healthy individual. Diagonal black line indicates deceased individual. WG, gestational weeks; WT, wild type.

Table 1 Summary of main clinical and imaging findings in three fetuses with a *PDHA1* variant

Case	*GA (weeks)*	*Ultrasound findings*	*Brain MRI findings*	*Brain autopsy findings*	*Pathogenic variant in* PDHA1 *(NM_000284.4)*
F1	21	Microcephaly, corpus callosal agenesis, cerebellar hypoplasia, dilated frontal horns and thickened ventricular walls	Agenesis of posterior part of corpus callosum, agyria, periventricular germinolytic pseudocysts, cerebellar hypoplasia, leukomalacia	Microcephaly, delayed gyration, partial corpus callosal agenesis, subependymal germinolytic pseudocysts, occipital leukomalacia, negative Perls staining†	c.705_706del; p.(Arg235Serfs*6)
F2	24	Ventriculomegaly	Ventriculomegaly, periventricular germinolytic pseudocysts, thin corpus callosum, white-matter abnormalities	Microcephaly, ventriculomegaly, delayed gyration, subependymal germinolytic pseudocysts, thin corpus callosum, negative Perls staining†	c.355C>T; p.(Arg119Trp)
F3	23	Ventriculomegaly, corpus callosal agenesis, echogenic ribbon circumscribing ventricular walls	Microcephaly, cerebellar hypoplasia, ventriculomegaly, almost complete corpus callosal agenesis, parenchymal rarefaction with periventricular germinolytic pseudocysts, delayed gyration	Not performed	c.910C>T; p.(Arg304*)

†Negative Perls staining indicates that there was no sign of previous hemorrhage. GA, gestational age at first suspicious ultrasound examination; MRI, magnetic resonance imaging.

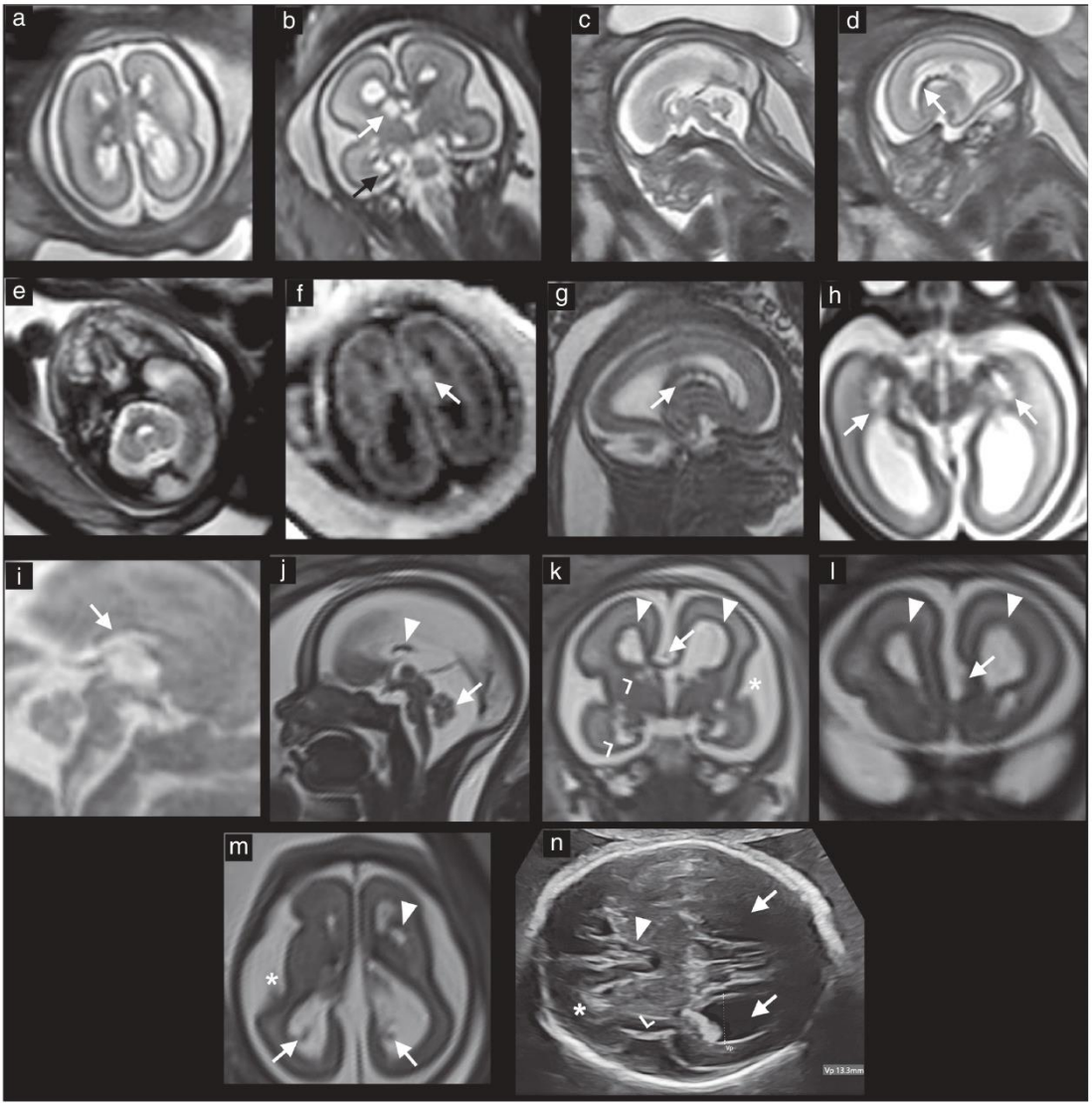


Figure 2 Brain imaging in three fetuses carrying pathogenic variants of *PDHA1* gene. (a–f) Fetal brain magnetic resonance imaging (MRI) performed at 26 weeks in fetus F1: (a–e) T2-weighted sequences and (f) T1-weighted sequence. (a) Axial image showing heterogeneous content of ventricles and hypointense signal of walls of lateral ventricles, with mild ventricular dilatation and irregular ventricular walls. (b) Coronal image showing severe gyration delay (appearance of agyria), bilateral periventricular pseudocysts (white arrow) and hypersignal of white matter on temporal lobe (black arrow). (c) Sagittal image showing partial agenesis of posterior part of corpus callosum. (d) Parasagittal plane showing highlighted ventricular walls, corresponding to suspected hemorrhage deposits (arrow). (e) Axial plane showing cerebellar hypoplasia. (f) Axial image showing heterogeneous content of ventricles with T1 hypersignals, suggestive of hemorrhagic suffusion (arrow). (g–i) Brain MRI at 30 weeks in fetus F2: axial and sagittal T2-weighted sequences. (g) Parasagittal section showing germinolytic microcysts localized in the ganglionic eminence and organized in 'string-of-beads' formation in right lateral ventricular wall (arrow). (h) Axial image showing microcysts present in anterior part of lateral ventricles (arrows) and hypersignal in left temporal white matter. (i) Sagittal section showing thin corpus callosum (arrow). (j–m) Brain MRI at 28 weeks in fetus F3. (j) T2-weighted sagittal sequence showing rudimentary segment of corpus callosum (8 mm, i.e. partial agenesis (arrowhead)) and overall reduction in parenchyma, especially affecting the vermis (hypoplasia (arrow)). (k) T2-weighted coronal sequence of middle part of frontal gyri: frontal horns are unusually enlarged (filled arrowheads) and there is overall reduction in parenchyma with expanded appearance of pericerebral spaces and significant gyration delay of sylvian sulci (*). Corpus callosum is thin (arrow) and there are cystic lesions of periventricular germinolysis under frontal horns and in temporal region (open arrowheads). (l) T2-weighted coronal sequence in anterior part of frontal gyri: frontal horns are enlarged and quadrangular in shape, with irregular walls (arrowheads). There is left frontal paramedian indentation, without schizencephaly (arrow). (m) T2-weighted axial sequence showing overall reduction in parenchyma, wide pericerebral spaces and ventriculomegaly affecting the occipital horns (arrows). Periventricular white matter appears heterogeneous and microcystic (arrowhead) and significant delay in gyration is evident (*). (n) Axial ultrasound image at 28 weeks in fetus F3, showing unusual anterior complex with septal agenesis (filled arrowhead), bilateral ventriculomegaly (arrows), heterogeneous microcystic hyperechoic appearance of periventricular white matter (*) and significant sylvian gyration delay (open arrowhead).

suggestive of bilateral ventricular hemorrhage, with hyposignal on T2- and hypersignal on T1-weighted sequences. Given these findings, we concluded that there had been several hypoxic-ischemic and hemorrhagic injuries associated with neurodevelopmental alterations. The fetal blood count and karyotype were normal. The pregnancy was terminated at 27 weeks. Autopsy showed a female fetus with a weight of 908 g, crown–heel length of 36 cm and moderate microcephaly (head circumference, 24 cm). Macroscopic cerebral study revealed global microcephaly with delayed gyration, partial corpus callosal agenesis and parietal and frontal subependymal pseudocysts. Histological analysis showed occipital leukomalacia, fragmentation of olivary and dentate nuclei and agenesis of the pyramidal tracts. Hemosiderin (a sign of previous hemorrhage) was not detected by Perls staining (Figure 3d). These findings were suggestive of clastic events.

Case 2

Fetus F2 was a female fetus from the third pregnancy of a 34-year-old woman. Her two previous children had normal neurological development. Ultrasound examination at 24 gestational weeks found dilatation of the lateral ventricles and a suspected blood clot in the left lateral ventricle without any other abnormality. Fetal brain MRI at 26 weeks confirmed slight asymmetric ventricular dilatation (width of cerebral right lateral ventricle, 14 mm; left lateral ventricle, 12.5 mm) and revealed the presence of three microcysts, which were interpreted as germinolytic cysts, in the left lateral ventricle. A possible left parietal parenchymal ischemic lesion was also suspected. Fetal MRI performed at 30 weeks showed no change in the dilatation of the lateral ventricles. A subependymal hemorrhage was suspected, without any clot in the ventricles. There were germinolytic microcysts in the right lateral ventricle apparently identical to those in the left one, localized in the ganglionic eminence and organized in a 'string-of-beads' formation. Microcysts were also present in the anterior part of the lateral ventricles with a septate appearance. The corpus callosum appeared thin. An abnormal signal was noted in the deep and superficial white matter of the left parietal hemisphere, on both T2 and diffusion sequences (Figure 2g–i). Cerebral maturation delay was also observed. Biometric data were between the 25th and 50th percentiles. According to the imaging findings, we concluded that the ventriculomegaly was probably the consequence of an intraventricular hemorrhage associated with ischemic lesions in the white matter and a delay in cerebral maturation. After termination of pregnancy, fetal postmortem examination at 31 weeks showed a eutrophic female fetus with a brain weight below the 5th percentile. Macroscopic brain analysis revealed delayed gyration and moderate dilatation of the lateral ventricles, with subependymal cysts at the ventricular junction and other confluent cysts in the right frontal pole (Figure 3e). The corpus callosum appeared thin without agenesis. Perls staining was negative, including around the different cysts.

Case 3

Fetus F3 was from the second pregnancy of an unrelated couple who already had a healthy child, without any familial history. Ultrasound examination at 23 gestational weeks found apparently complete corpus callosal agenesis with mild ventricular enlargement. Karyotyping and polymerase chain reaction analysis for cytomegalovirus were negative. A second ultrasound examination at 28 weeks confirmed the corpus callosal agenesis and dilatation of the lateral ventricles (Figure 2n). On brain MRI at 28 weeks (Figure 2j–m), the head circumference was below the 3rd percentile. The cerebral parenchyma showed overall rarefaction, with widened pericerebral spaces. The corpus callosum was almost entirely absent, although a small central segment could be observed (8 mm in length from the anteroposterior axis and 2 mm in width). The septal cavity was absent. The third ventricle was dilated and appeared ascended. Bilateral ventriculomegaly was slightly asymmetrical (left lateral ventricle, 15 mm; right lateral ventricle, 13 mm). The frontal ventricular horns were dilated and the left frontal ventricular horn presented relatively rectangular contours. The ventricular walls were irregular and on the left side they were relatively indented, without the appearance of nodular heterotopia. Several pseudocystic zones suggestive of germinolytic lesions were observed, located particularly within the left basifrontal white matter and to a lesser extent within the right frontal subependymal region, and bilaterally within the germinative zone in the periphery of the temporal horns. Sagittal sequences showed a slit-like appearance of the lateral ventricular walls, suggesting extensive germinolytic lesions, probably corresponding to the echogenic ribbon that surrounds the ventricular wall on ultrasound. Additionally, there was a clear indentation of the left anterior frontal cortex and adjacent white matter, located at the level of the interhemispheric fissure paramedially, with a lacunar appearance, without radiological evidence of schizencephaly. The T1- and T2-weighted gradient echo sequences were not suggestive of hemorrhagic stigma. Although it was still relatively early in gestation, a clear and global delay in gyration was noted. At the subtentorial level, the cerebellar growth parameters were below the 3rd percentile. The vermis appeared complete and the morphology of the fourth ventricle was normal. There was significant widening of all pericerebellar fluid spaces. Overall, a very early clastic event was hypothesized. The pregnancy was terminated at 31 weeks and the parents declined postmortem examination.

DISCUSSION

Our findings highlight that *PDHA1* pathogenic variants may lead to ultrasound and brain MRI lesions suggestive of those observed in fetuses carrying *COL4A1/COL4A2* pathogenic variants. In addition, our data strongly suggest the utility of exome sequencing in fetuses referred for *COL4A1/COL4A2* screening for which targeted sequencing is negative.

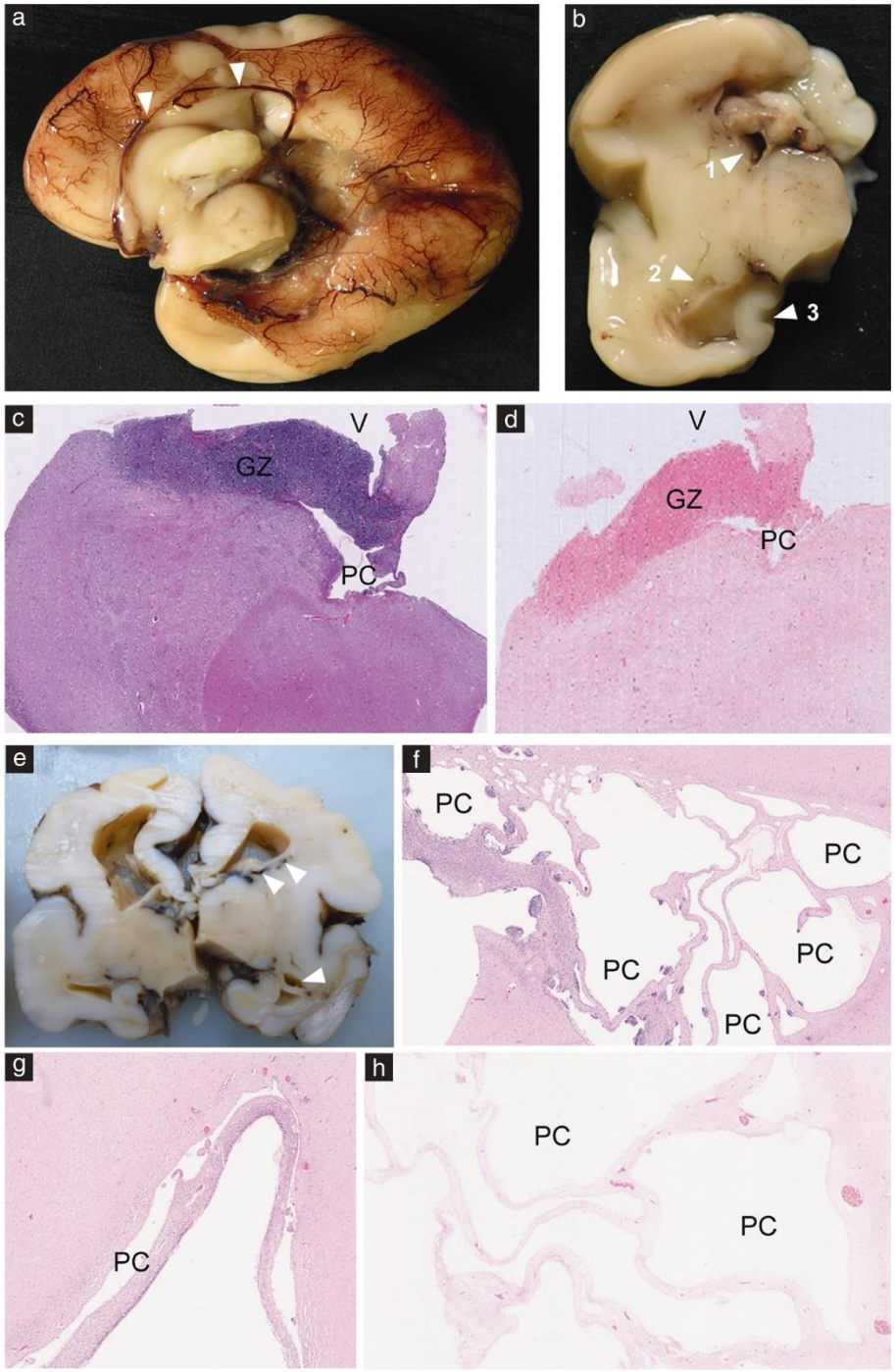


Figure 3 Macroscopic and histological brain features of fetal PDHA1 deficiency. (a–d) Fetus F1, at 27 weeks. (a) Median sagittal slice of right hemisphere showing partial agenesis of the corpus callosum (only third anterior part); arrowheads indicate pericallosal artery. (b) Transverse section of cerebral hemisphere showing left subependymal parietal (arrowhead 1) and temporal (arrowhead 2) pseudocysts (PCs) and temporal thinning (arrowhead 3). (c) Histological image showing subependymal PCs (hematoxylin & eosin staining (HES); original magnification, ×25). (d) Corresponding Perls staining, showing negative result (original magnification, ×25). (e–h) Fetus 2, at 31 weeks. (e) Coronal brain slice showing moderate ventricular dilatation and periventricular PCs (arrowheads). (f) Histological pattern showing multiple frontal PCs (HES; original magnification, ×20). (g) Periventricular subependymal PC (HES; original magnification, ×20). (h) Frontal PCs with negative result for Perls staining (original magnification, ×40). GZ, germinative zone; V, ventricle.

Fetal brain anomalies of *PDHA1* deficiency could lead to suspicion of lesions associated with *COL4A1/COL4A2* pathogenic variants

Primary PDCD is a mitochondrial disorder of carbohydrate oxidation. Molecular alterations in the *PDHA1* gene are the main cause of PDCD with an X-linked dominant inheritance pattern and a high frequency of *de-novo* variants (85–95% in female and 60–63% in male cases)[2]. The clinical phenotype and phenotype on imaging as well as the age at onset of PDCD are variable, the median age at onset being about 20 months[3,4]. PDCD may also present during fetal life, with prenatal ultrasound or MRI detecting microcephaly, ventriculomegaly, paraventricular pseudocysts, cerebellar hypoplasia, delayed gyration and/or dysgenesis of the corpus callosum. In addition, stroke-like lesions have been reported in some fetuses[5]. Fetal autopsy features have been described by Pirot *et al.*[6] in two unrelated female fetuses with *PDHA1* variants. The first patient had partial agenesis of the corpus callosum and microcephaly on routine ultrasound at 22 gestational weeks. Fetal MRI showed agenesis of the corpus callosum, microcephaly, colpocephaly, delayed gyral formation, cerebellar hypoplasia and microcavitary periventricular white-matter lesions. Macroscopic and histologic examination of the brain showed pachygyria, polymicrogyria, ventricular dilatation, pseudocysts of the subependymal germinal matrix and basal ganglia microcalcifications. The second fetus had agenesis of the corpus callosum, mild ventricular dilatation and delayed gyration at the 22-week ultrasound examination. Postmortem examination showed extensive subcortical cavitary lesions in the brain, corpus callosal hypoplasia, pseudocysts of the subependymal germinal matrix, dilatation of the ventricles, reactive gliosis and ischemic lesions.

Several clinical and imaging manifestations encountered in PDHA1 deficiency are also encountered in *COL4A1/COL4A2* collagenopathy. First, microcephaly can be observed in both conditions, although less consistently in *COL4A1/COL4A2* collagenopathy[7]. Second, ventriculomegaly is a common element in both cases[6,8], although it is often classified as moderate in metabolic abnormalities compared with those associated with a hemorrhagic event. Hypoplasia of the corpus callosum, cerebellar hypoplasia, schizencephaly, gyration abnormalities and periventricular leukomalacia are also elements that can be observed in both PDHA1 deficiency and *COL4A1/COL4A2*-related disorders[7,9,10]. Dysgenesis of the corpus callosum in *PDHA1* fetuses is thought to be developmental; this is in contrast to the corpus callosal lesions of *COL4A1* fetuses, which are thought to be the consequence of clastic lesions[6]. For this reason, an early anomaly of the corpus callosum should lead to suspicion of a *PDHA1* mutation. Finally, intracranial calcifications may be observed in both *COL4A1/COL4A2* collagenopathy and PDHA1 deficiency[6]. Some elements, however, tend to be specific to only one of the two conditions. To our knowledge, periventricular germinolytic pseudocysts have not been reported in patients with pathogenic variants in *COL4A1/COL4A2*. Hemorrhages are usually not detected on brain histological examination, even after Perls staining, in fetuses with PDHA1 deficiency, in contrast to observations in most *COL4A1/COL4A2* fetuses. Altogether, these data emphasize the overlap of *PDHA1* and *COL4A1/COL4A2* phenotypic features in fetuses carrying pathogenic variants.

Shared phenotypic features of fetuses in this series

In our three cases, antenatal brain imaging using ultrasound and particularly MRI raised the possibility of hypoxic-ischemic and/or hemorrhagic lesions, without demonstrating them clearly (non-specific white-matter abnormalities; paraventricular microcysts interpreted as germinolysis without certainty; absence of signs of intraventricular or intraparenchymal hemorrhage). In fetuses F1 and F3, the first sign of concern was an abnormal or absent corpus callosum (partial agenesis of the posterior part in F1 and almost complete agenesis in F3). In fetus F2, the corpus callosum appeared normal on ultrasound examination but a thin corpus callosum was then observed on fetal brain MRI. These corpus callosal abnormalities were confirmed by postmortem examination in two cases; postmortem was not carried out in fetus F3. All three fetuses presented the appearance of pseudo- and/or microcysts at the subependymal level, suggesting germinolysis. In F2, they appeared localized in the ganglionic eminence and organized in a 'string-of-beads' formation. In the same fetus, there were also microcysts in the anterior part of the lateral ventricles with a septate appearance, which was not suggestive of germinolysis. In fetuses F1 and F3, cerebellar hypoplasia was also observed. These different findings are rather unusual for lesions of vascular origin, such as those found in *COL4A1/COL4A2* collagenopathy. When hypoxic-ischemic/hemorrhagic lesions are suspected, it is crucial to examine carefully the appearance of microcysts that are not truly undergoing germinolysis, and to look for developmental abnormalities of the corpus callosum and cerebellum that may indicate PDHA1 deficiency, particularly in a female fetus.

Exome sequencing: efficient strategy for fetuses with suspected cerebral hypoxic-ischemic lesions and ventriculomegaly

This study emphasizes the difficulty in some cases in achieving accurate diagnosis based on fetal phenotypic analysis. In these three fetuses, imaging findings initially suggested a diagnosis based on a hypoxic-ischemic or hemorrhagic event, which led us to conduct first-line targeted sequencing of the *COL4A1/COL4A2* genes. The final diagnosis was achieved only after an exome-wide analysis performed using a proband-only strategy. Taken together, these data highlight the importance of genome-wide screening after targeted *COL4A1/COL4A2*

gene screening, coupled with sequential *in-silico* analysis of candidate genes, based on a multidisciplinary approach.

ACKNOWLEDGMENTS

We thank sincerely Dr Patricia Blanchet, Dr Alain Couture and Dr Anne Paris for their collaboration on this work. We also thank families for their participation in this study, and clinicians, geneticists and pathologists who referred fetuses enrolled in this study. Finally, we are greatly indebted to Dr Florence Marchelli for her excellent editing of figures. This work was supported by the National Institutes of Health (R01NS096173 grant).

REFERENCES

1. Imbard A, Boutron A, Vequaud C, Zater M, de Lonlay P, de Baulny HO, Barnerias C, Miné M, Marsac C, Saudubray JM, Brivet M. Molecular characterization of 82 patients with pyruvate dehydrogenase complex deficiency. Structural implications of novel amino acid substitutions in E1 protein. *Mol Genet Metab* 2011; **104**: 507–516.
2. Ganetzky R, McCormick EM, Falk MJ. Primary Pyruvate Dehydrogenase Complex Deficiency Overview. 2021 Jun 17. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Gripp KW, Mirzaa GM, Amemiya A (eds). GeneReviews® [Internet]. University of Washington: Seattle, WA, 1993–2022. https://www.ncbi.nlm.nih.gov/books/NBK571223/
3. DeBrosse SD, Okajima K, Zhang S, Nakouzi G, Schmotzer CL, Lusk-Kopp M, Frohnapfel MB, Grahame G, Kerr DS. Spectrum of neurological and survival outcomes in pyruvate dehydrogenase complex (PDC) deficiency: lack of correlation with genotype. *Mol Genet Metab* 2012; **107**: 394–402.
4. Sofou K, Dahlin M, Hallböök T, Lindefeldt M, Viggedal G, Darin N. Ketogenic diet in pyruvate dehydrogenase complex deficiency: short- and long-term outcomes. *J Inherit Metab Dis* 2017; **40**: 237–245.
5. Savvidou A, Ivarsson L, Naess K, Eklund EA, Lundgren J, Dahlin M, Frithiof D, Sofou K, Darin N. Novel imaging findings in pyruvate dehydrogenase complex (PDHc) deficiency - Results from a nationwide population-based study. *J Inherit Metab Dis* 2022; **45**: 248–263.
6. Pirot N, Crahes M, Adle-Biassette H, Soares A, Bucourt M, Boutron A, Carbillon L, Mignot C, Trestard L, Bekri S, Laquerrière A. Phenotypic and Neuropathological Characterization of Fetal Pyruvate Dehydrogenase Deficiency. *J Neuropathol Exp Neurol* 2016; **75**: 227–238.
7. Zagaglia S, Selch C, Nisevic JR, Mei D, Michalak Z, Hernandez-Hernandez L, Krithika S, Vezyroglou K, Varadkar SM, Pepler A, Biskup S, Leão M, Gärtner J, Merkenschlager A, Jaksch M, Møller RS, Gardella E, Kristiansen BS, Hansen LK, Vari MS, Helbig KL, Desai S, Smith-Hicks CL, Hino-Fukuyo N, Talvik T, Laugesaar R, Ilves P, Õunap K, Körber I, Hartlieb T, Kudernatsch M, Winkler P, Schimmel M, Hasse A, Knuf M, Heinemeyer J, Makowski C, Ghedia S, Subramanian GM, Striano P, Thomas RH, Micallef C, Thom M, Werring DJ, Kluger GJ, Cross JH, Guerrini R, Balestrini S, Sisodiya SM. Neurologic phenotypes associated with COL4A1 / 2 mutations: Expanding the spectrum of disease. *Neurology* 2018; **91**: e2078–2088.
8. Itai T, Miyatake S, Taguri M, Nozaki F, Ohta M, Osaka H, Morimoto M, Tandou T, Nohara F, Takami Y, Yoshioka F, Shimokawa S, Okuno-Yuguchi J, Motobayashi M, Takei Y, Fukuyama T, Kumada S, Miyata Y, Ogawa C, Maki Y, Togashi N, Ishikura T, Kinoshita M, Mitani Y, Kanemura Y, Omi T, Ando N, Hattori A, Saitoh S, Kitai Y, Hirai S, Arai H, Ishida F, Taniguchi H, Kitabatake Y, Ozono K, Nabatame S, Smigiel R, Kato M, Tanda K, Saito Y, Ishiyama A, Noguchi Y, Miura M, Nakano T, Hirano K, Honda R, Kuki I, Takanashi JI, Takeuchi A, Fukasawa T, Seiwa C, Harada A, Yachi Y, Higashiyama H, Terashima H, Kumagai T, Hada S, Abe Y, Miyagi E, Uchiyama Y, Fujita A, Imagawa E, Azuma Y, Hamanaka K, Koshimizu E, Mitsuhashi S, Mizuguchi T, Takata A, Miyake N, Tsurusaki Y, Doi H, Nakashima M, Saitsu H, Matsumoto N. Prenatal clinical manifestations in individuals with COL4A1/2 variants. *J Med Genet* 2021; **58**: 505–513.
9. Yoneda Y, Haginoya K, Kato M, Osaka H, Yokochi K, Arai H, Kakita A, Yamamoto T, Otsuki Y, Shimizu S, Wada T, Koyama N, Mino Y, Kondo N, Takahashi S, Hirabayashi S, Takanashi J, Okumura A, Kumagai T, Hirai S, Nabetani M, Saitoh S, Hattori A, Yamasaki M, Kumakura A, Sugo Y, Nishiyama K, Miyatake S, Tsurusaki Y, Doi H, Miyake N, Matsumoto N, Saitsu H. Phenotypic Spectrum of COL4A1 Mutations: Porencephaly to Schizencephaly. *Ann Neurol* 2013; **73**: 48–57.
10. Meuwissen MEC, Halley DJJ, Smit LS, Lequin MH, Cobben JM, de Coo R, van Harssel J, Sallevelt S, Woldringh G, van der Knaap MS, de Vries LS, Mancini GMS. The expanding phenotype of COL4A1 and COL4A2 mutations: clinical data on 13 newly identified families and a review of the literature. *Genet Med* 2015; **17**: 843–853.

SUPPORTING INFORMATION ON THE INTERNET

The following supporting information may be found in the online version of this article:

Appendix S1 Subjects and methods

Table S1 Variants identified in *PDHA1* gene in the cohort of 113 fetuses with suspicion of intracerebral hemorrhage

Supplemental data

Rare metabolic disease mimicking *COL4A1/COL4A2* fetal brain phenotype

T. Coste[1,2], C. Aloui[1]*, F. Petit[3]*, S. Moutton[4]*, L. Devisme[5], C. F. Wells[6], N. Leboucq[7], P. Verpillat[8], M. Yvert[4], F. Rivier[9†] and E. Tournier-Lasserve[1,2†]

*Authors contributed equally to this work. †Co-last authors.

[1]Université Paris Cité, Inserm, NeuroDiderot, F-75019 Paris, France

[2]Service de Génétique Neurovasculaire, AP-HP, Hôpital Saint-Louis, F-75010 Paris, France

[3]CHU Lille, Clinique de Génétique Guy Fontaine, F-59000 Lille, France

[4]Centre Pluridisciplinaire de Diagnostic Prénatal, MSP Bordeaux Bagatelle, Talence, France

[5]CHU Lille, Institut de Pathologie, Lille, France

[6]CHU Montpellier, Département de Génétique Médicale et Fœtopathologie, Montpellier, France

[7]CHU Montpellier, Département de Neuroradiologie, Neuroradiologie Diagnostique Pédiatrique, Montpellier, France

[8]CHU Lille, Service de Radiologie, F-59000 Lille, France

[9]CHU Montpellier, Département de Neurologie Pédiatrique, PhyMedExp, Université de Montpellier, INSERM, CNRS, Montpellier, France

Corresponding author:

Dr T. Coste

AP-HP, Service de Génétique Moléculaire Neurovasculaire, Hôpital Saint-Louis, France

e-mail: thibault.coste@aphp.fr

Supplemental data

Table S1. Variants identified in *PDHA1* gene in the cohort of 113 ICH fetuses.

Fetus	HGVSc position (NM_000284.4)	Protein change	Variant class	Status	gnomAD v.2.1	ACMG class
F1	c.705_706del	p.(Arg235Serfs*6)	Frameshift	HTZ	Absent	Pathogenic
F2	c.355C>T	p.(Arg119Trp)	Missense	HTZ	Absent	Pathogenic
F3	c.910C>T	p.(Arg304*)	Nonsense	HTZ	Absent	Pathogenic

HTZ: heterozygous

Supplemental data

Appendix S1: Material and methods

Inclusion criteria

These three fetuses were part of a series of 113 unrelated fetuses referred for *COL4A1* and *COL4A2* molecular screening. This study is a retrospective study and all fetuses were referred to the French national molecular genetics reference center for inherited cerebrovascular disorders in Saint Louis Hospital after termination of pregnancy (TOP) or intra-utero fetal death (IUFD). Fetal ICH was in most cases suspected at systematic second and third trimester ultrasound examinations and in some cases confirmed by fetal MRI and/or pathological examination. Systematic review of medical charts was performed in order to exclude fetuses with an identifiable cause or known risk factor for ICH. These included evidence of maternal trauma during pregnancy, cocaine or maternal drug use, maternal or neonatal infections and fetal alloimmune thrombocytopenia (FNAIT).

Written informed consent for fetal autopsy, genetic investigation and research was provided by parents in accordance with the declaration of Helsinki and the French law. This study has been approved by the Inserm Ethics Committee (INSERM IRB00003888). Genomic DNA was isolated from post mortem fetal tissue and from peripheral blood leukocytes of both parents and relatives when available.

Molecular screening of *COL4A1 / COL4A2* genes

Prior to exome sequencing, we searched for variants in *COL4A1* and *COL4A2* by performing a targeted high throughput sequencing with the SureSelect QXT® capture kit (Agilent technologies) and sequencing performed on a MiSeq® sequencer (Illumina) (conditions available upon request).

Supplemental data

Exome sequencing

Exome sequencing was performed as previously described.[9] Briefly, Exon capture was performed at the IntegraGen platform (Evry, France) for fetus probands and relatives using the SureSelect Human All Exon V5-UTR (Agilent technologies) or the Twist Human Core Exome Enrichment System (Twist Bioscience). Followed by 100 base pair paired-end sequencing using an Illumina NovaSeq platform. Data analysis was performed with an IntegraGen bioinformatic pipeline. Sequence reads were aligned to the human genome reference GRCh38/hg38 using BWA. Variant calling for the identification of SNVs (Single Nucleotide Variations) and small insertions/deletions was performed via the Broad Institute's GATK Haplotype Caller GVCF tool (GATK 3.8.1). Ensembl's VEP program was used to process variants for further annotations. Allele frequency annotations originated from the gnomAD (v2.1.1) and 1000Genomes project (Phase 3). PolyPhen-2 algorithm was used to predict deleteriousness of missense variants.

Qualifying variants and gene-based collapsing test

Several criteria were used to filter qualifying variants, including 1) quality control: a minimum of 10 reads was required for depth coverage and 25% for allelic balance, 2) variants nature: loss of function (including stop gained, frameshift and splice site variants) and missense variations, 3) allele frequency in external databases (gnomAD v2.1.1 and 1000Genomes Phase 3) with a requirement for minor allele frequency (MAF) $\leq$ 1/10 000, and 4*) in silico* pathogenicity prediction according to PolyPhen-2 algorithm.

Investigation of an enrichment of rare qualifying variants was performed by comparing our cohort (N=113) to the gnomAD v3.0 control cohort (N=71,702). We generated a combined VCF file including fetus and gnomAD control cohorts and we reannotated it with Ensembl's

Supplemental data

VEP version 104. For each gene, the number of individuals carrying at least one qualifying variant was compared between cases and controls using Fisher's exact test implemented in R (https://www.r-project.org). Bonferroni-corrected exome-wide statistical significance threshold was set to 2.5 x10-6.

Postmortem examination

Complete postmortem examination was performed with the informed consent of both parents in accordance with the French law and following standardized protocols including X-rays, photographs, macroscopical and histological examination of all viscera. Brains were fixed with formalin for 5 to 12 weeks. Macroscopic analysis was performed allowing the selection and conditioning of samples (paraffin embedding, 7-micron slicing, hematoxylin and eosin stain (HES)) of brain tissue for histological analysis.

DISCUSSION

Un test de charge en variants candidats à l'échelle du gène sur les données de WES a permis d'identifier des variants hétérozygotes causaux dans le gène *PDHA1* chez trois fœtus de sexe féminin. Deux de ces variants étaient de type perte de fonction entrainant une haploinsuffisance par dégradation de l'ARN messager, et le troisième était un variant faux-sens. Aucun de ces variants n'était présent dans les bases de données de contrôles (gnomAD et 1000Genomes), et le variant faux sens (p.R119W) avait déjà été rapporté chez un enfant présentant un déficit en pyruvate déshydrogénase confirmé biologiquement.

Les variations pathogènes de *PDHA1* sont responsables d'une maladie dominante liée à l'X avec déficit primaire du complexe de la pyruvate déshydrogénase (PDCD primaire). L'âge de début de la maladie est très variable. Généralement, l'âge moyen du diagnostic du PDCD primaire est d'environ 45 mois (âge médian ~20 mois) (Sofou *et al.*, 2017). La présentation peut être prénatale avec une échographie détectant une microcéphalie, une ventriculomégalie, des pseudokystes périventriculaires, une hypoplasie cérébelleuse, un retard de gyration, et/ou une dysgénésie du corps calleux. L'IRM fœtale permet de mettre en évidence ces lésions cérébrales, ainsi qu'une perte de volume du cerveau et/ou des hypersignaux périventriculaires en pondération T2 (Natarajan *et al.*, 2016 ; Pirot *et al.*, 2016). Les variants pathogènes *de novo* de *PDHA1* sont plus fréquents chez les sujets de sexe féminin que chez les sujets de sexe masculin : environ 85%-95% des femmes atteintes ont un variant pathogène *de novo* (Lissens *et al.*, 2000 ; Imbard *et al.*, 2011). L'analyse à postériori des ADN parentaux pour 2 des 3 des fœtus rapportés ici a permis de confirmer le caractère *de novo* des variants pathogènes.

Les trois fœtus rapportés ici présentaient une microcéphalie, une ventriculomégalie, une dysgénésie du corps calleux et des anomalies de la substance blanche qui ont initialement fait évoquer des lésions hypoxo-ischémiques et hémorragiques. Les données neuropathologiques avaient confirmé l'agénésie du corps calleux, la présence de pseudo-kystes pariétaux sous-épendymaires et la dilatation ventriculaire modérée. La coloration de Perls sur les coupes histologiques cérébrales ne retrouvait pas de pigments hémosidériniques en quantité significative.

A posteriori, après identification du variant, un « rétrophénotypage » a montré que les lésions observées étaient évocatrices de celles présentes chez des fœtus porteurs de variants pathogènes de *PDHA1*. Nous avons ainsi conclu que le phénotype cérébral d'un déficit en PDHA1 chez le fœtus peut mimer une collagénopathie COL4A1/A2.

Les résultats de cet article mettent en évidence les difficultés d'un diagnostic précis sur la base des données d'imagerie et de fœtopathologie chez un fœtus. Cette complexité vient de la similitude de certaines anomalies cérébrales observées dans les collagénopathies COL4A1/A2 et certaines maladies métaboliques comme le déficit en PDHA1. Comme mentionné dans l'article de synthèse en cours de soumission, nous avons détecté, dans notre cohorte de 113 fœtus, d'autres variants pathogènes dans divers gènes impliqués le métabolisme mitochondrial comme les gènes *COQ2, ATP5PO* et probablement le gène *NAXD.*

Cet article souligne là encore l'importance de l'analyse de l'exome entier particulièrement dans les cas de criblage *COL4A1/COL4A2* négatif.

E- Autres gènes candidats identifiés

1) Gène *PLOD3* impliqué dans la biosynthèse de COL4A1/COL4A2

Étant donné que les gènes *COL4A1* et *COL4A2* sont associés à des HIC chez le fœtus, les gènes impliqués dans la biosynthèse du collagène de type IV constituent de potentiels gènes candidats.

Une liste de 47 partenaires impliqués dans la synthèse de COL4A1 et COL4A2 (Figure 29) a été générée en collaboration avec l'équipe du Dr Vanacore aux USA (Projet NIH R01NS096173). Cette liste de gènes est disponible en Annexe 2.

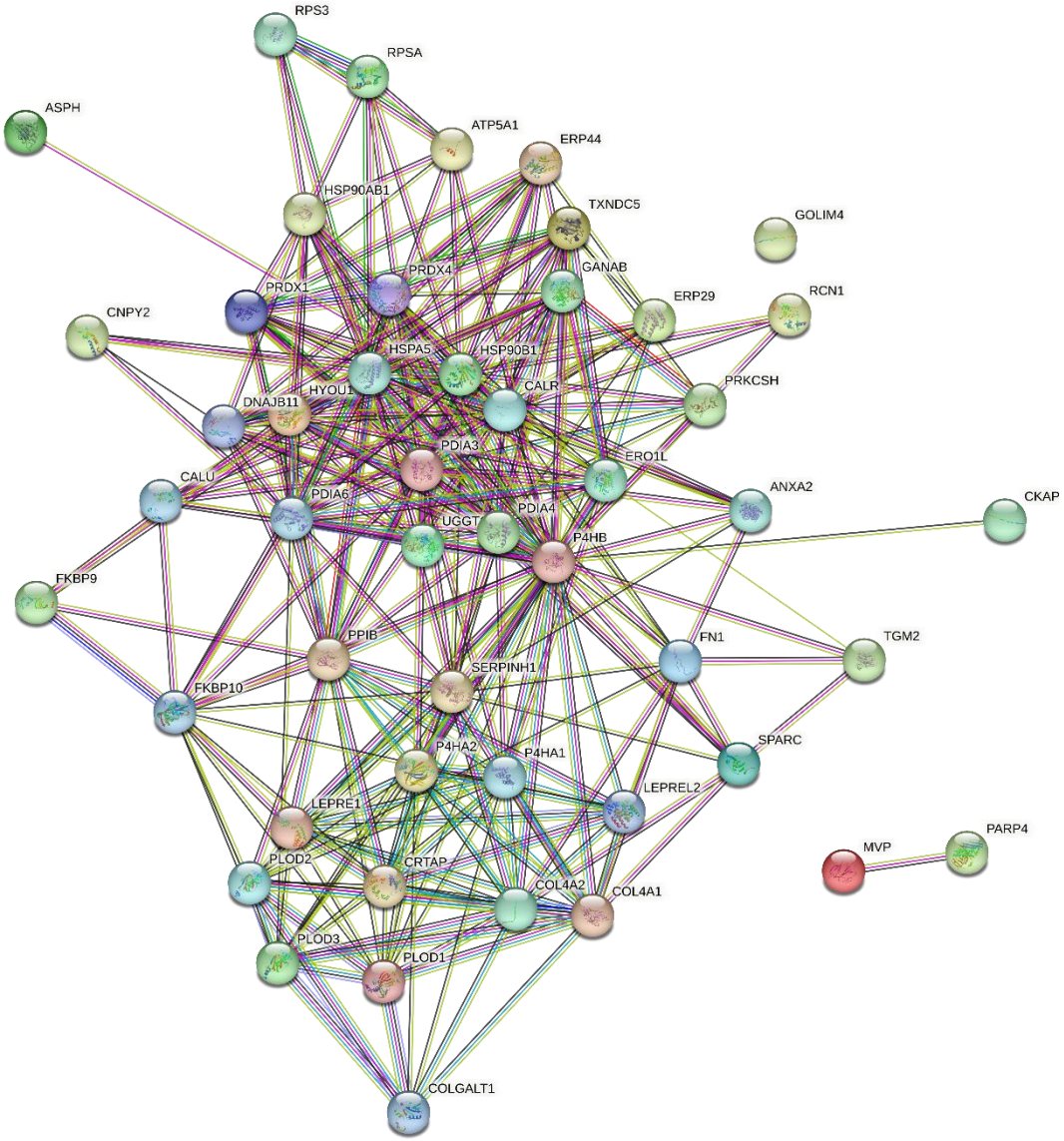


Figure 29. Illustration des interactions des 47 gènes partenaires intervenant dans la biosynthèse de COL4A1/COL4A2. Image générée via l'outil STRING v11.5 (https://string-db.org).

Sur les données de WES des 113 fœtus, nous avons cherché la présence de variants candidats dans cette liste de gènes. Nous avons pris comme filtre une fréquence allélique < 1/200 dans la population gnomAD sous l'hypothèse d'une transmission autosomique récessive (variants homozygotes et hétérozygotes composites) et 1/1000 en hypothèse dominante. Pour les variants de type faux-sens nous avons ajouté un critère de sélection supplémentaire, celui d'être prédit délétère par le logiciel de prédiction *in silico* PolyPhen-2.

Selon une hypothèse autosomique récessive (MAF <1/200), nous n'avons pas mis en évidence de variant homozygote candidat. Deux fœtus étaient toutefois porteurs de deux variants candidats dans le même gène :

- Fœtus F31 : variants p.Y547H et p.C282R dans le gène *PLOD2.* Il s'agit de deux variants faux sens dont la phase est inconnue car les ADN génomiques des parents n'étaient pas disponibles.
- Fœtus F12 : variants p.P489L et p.R197W dans le gène *PLOD3* (véritable hétérozygote composite confirmé grâce aux ADN génomiques parentaux)

Selon une hypothèse autosomique dominante (MAF<1/1000), 21 variants hétérozygotes distincts ont été détectés chez 20 fœtus (dont un fœtus (F12) portant deux variants hétérozygotes dans *PLOD3*) (Tableau 5). Tous ces variants étaient de type faux-sens à l'exception d'un variant de type disruptif (codon stop prématuré) dans le gène *TGM2*. Le gène *TGM2* est un gène qualifié de « tolérant » aux pertes de fonction hétérozygotes puisque sa probabilité d'être intolérant à la perte de fonction est très faible (pLI =0) ; il existe 143 individus dans la base de données gnomAD qui sont porteurs d'un variant disruptif hétérozygote dans ce gène.

Numéro d'identification Famille	Gène	Nomenclature HGVSc	Changement Protéine	Conséquence	Fréquence allélique gnomAD v2.1
F50	*ASPH*	NM_001164750.1:c.1211T>G	L404R	Faux sens	0.0076%
F53	*CKAP4*	NM_006825.3:c.1100C>T	P367L	Faux sens	0.0012%
F36	*CNPY2*	NM_014255.6:c.538G>A	D180N	Faux sens	0.034%
F108	*CNPY2*	NM_014255.6:c.538G>A	D180N	Faux sens	0.034%
F15	*COLGALT1*	NM_024656.3:c.976C>T	R326C	Faux sens	0.0094%
F42	*P3H1*	NM_001146289.1:c.426G>T	K142N	Faux sens	0.016%
F81	*P3H1*	NM_001146289.1:c.1501C>T	R501W	Faux sens	0.075%
F63	*P3H1*	NM_001146289.1:c.1501C>T	R501W	Faux sens	0.075%
F86	*P4HA2*	NM_001017973.1:c.695G>A	R232H	Faux sens	0.011%
F86	*P4HB*	NM_000918.3:c.709G>A	V237I	Faux sens	0.019%
F59	*PARP4*	NM_006437.3:c.1621A>G	T541A	Faux sens	0.0032%
F89	*PARP4*	NM_006437.3:c.2408A>G	H803R	Faux sens	0.034%
F14	*PDIA4*	NM_004911.4:c.497A>G	E166G	Faux sens	0.029%
F44	*PLOD1*	NM_000302.3:c.137G>A	R46H	Faux sens	0.025%
F92	*PLOD1*	NM_000302.3:c.613C>T	R205C	Faux sens	0.0036%
F69	*PLOD2*	NM_000935.2:c.844T>C	C282R	Faux sens	0.021%
F78	*PLOD2*	NM_000935.2:c.230G>T	G77V	Faux sens	Abs
F12	*PLOD3*	NM_001084.4:c.589C>T	R197W	Faux sens	0.0012%
F12	*PLOD3*	NM_001084.4:c.1466C>T	P489L	Faux sens	0.051%
F09	*PLOD3*	NM_001084.4:c.593T>C	L198P	Faux sens	0.0004%
F30	*TGM2*	NM_001323316.1:c.307C>T	Q103X	Codon stop	0.0068%
F29	*UGGT1*	NM_020120.3:c.331T>C	S111P	Faux sens	0.0008%
F15	*UGGT1*	NM_020120.3:c.1835A>C	Y612S	Faux sens	0.0018%

Tableau 5. Variants hétérozygotes faux-sens et disruptifs rares (MAF≤ 1/1000) des gènes impliqués dans la biosynthèse de COL4A1/COL4A2 identifiés chez les 113 fœtus.

Pour ces 21 variants identifiés, vous avons cherché une association génétique en comparant les fréquences de variants candidats rares (MAF <1/1000) entre les 113 fœtus atteints d'HIC et les individus contrôles de gnomAD à l'aide d'un test de charge à l'échelle du gène (Tableau 6).

Gène	Nombre de variants candidats chez fœtus	Nombre de variants candidats chez individus gnomAD v3.0	p-value
ASPH	1	580	0,6
CKAP4	1	404	0,47
CNPY2	2	85	0,008
COLGALT1	1	741	0,86
P3H1	3	900	0,17
P4HA2	1	339	0,42
P4HB	1	293	0,37
PARP4	2	1388	1
PDIA4	1	279	0,36
PLOD1	2	889	0,65
PLOD2	2	587	0,24
PLOD3	3	801	0,36
TGM2	1	630	1
UGGT1	2	795	0,36

Tableau 6. Analyse d'enrichissement en variants candidats des gènes de la biosynthèse de COL4A1/COL4A2 dans la cohorte de fœtus par rapport à la cohorte de contrôle gnomAD.

Aucun gène n'avait une valeur de p atteignant le seuil de significativité (seuil de significativité avec la correction de Bonferroni égal à 0,001 (0.05/47)).

Le gène *PLOD3* a fait l'objet d'une publication en collaboration avec l'équipe du Dr Gould. *PLOD3* code pour la lysyl-hydroxylase 3 (LH3), qui est une enzyme multifonctionnelle impliquée dans la modification du collagène mais dont la spécificité de substrat (type de collagène) n'était pas connue avant l'étude. L'équipe du Dr Gould a travaillé sur des cellules LH3 knock-out et a pu montrer que le déficit en LH3 altère drastiquement la sécrétion du collagène de type IV (α1α1α2), mais pas les autres types de collagènes (types I et III). L'analyse a de plus démontré que LH3 est une lysyl-hydroxylase sélective pour le collagène α1α1α2(IV), mais une glucosyltransférase générale pour les collagènes α1α1α2(IV), α1α1α2(I) et α1α1α1(III). C'est dans ce contexte que nous avons rapporté les données des deux fœtus avec des variants rares et prédits délétères dans le gène *PLOD3.* Il s'agissait du fœtus F09 porteur d'un variant candidat hétérozygote faux sens [p.L198P/Wild-type] et du fœtus F12 qui est hétérozygote composite [p.R197W/p.P489L]. Les variants p.(R197W) et p.(L198P) sont situés dans le domaine Glycosyltransférase et le variant p.(P489L) est situé dans le domaine accessoire (domaine similaire au domaine glycosyltransférase mais qui n'a pas d'activité catalytique) (Figure 30).

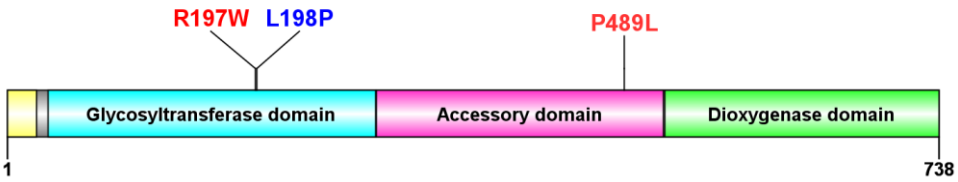


Figure 30. Représentation schématique de la protéine PLOD3 et positions des variants candidats identifiés chez les fœtus F09 (variant en bleu) et F12 (variants en rouge).

D'un point de vue structural, les variants p.(R197W) et p.(L198P) sont proches d'une séquence comportant plusieurs résidus d'acide aspartique chargés négativement (séquence poly-Asp) qui est très proche du site catalytique. Ces variants pourraient ainsi modifier la forme de la poche active (Figure 31). Concernant le variant p.(P489L), le résidu Proline est situé au début d'une hélice alpha dans le domaine accessoire qui est un domaine de reconnaissance du substrat. La proline assure normalement une "rotation" de la chaîne polypeptidique, le variant pourrait ainsi modifier la direction et affecter la reconnaissance du substrat.

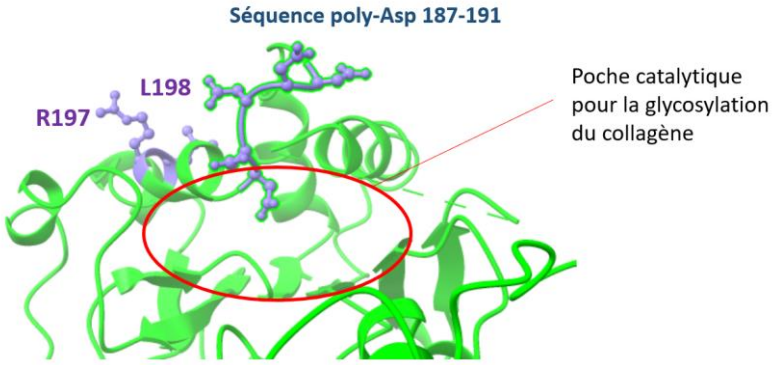


Figure 31. Représentation structurale des positions protéiques R197 et L198.

Les données des fœtus sont détaillées dans l'article d'Ishikawa *et al.,* (2022) **en annexe.** En dehors des lésions cérébrales ischémo-hémorragiques, les deux fœtus ne montraient pas de malformations externes ou viscérales ni d'anomalies du squelette.

Ces données suggèrent que des variations pathogènes de *PLOD3* pourraient être responsable d'un phénotype d'HIC fœtale mais ceci reste à démontrer formellement. La prochaine étape, en cours de réalisation par l'équipe du Dr Douglas Gould, consiste à tester fonctionnellement les variants candidats dans des cellules PFHR9 en utilisant une technique CRISPR pour introduire les variations spécifiques et tester si les activités glycosyltransférase et/ou hydroxylase sont altérées.

2) Gène *TIE1* impliqué dans l'angiogenèse

Lors de l'analyse des variants hétérozygotes rares et prédits délétères par le test de charge à l'échelle génique, le deuxième gène de la liste, après *PDHA1*, avec une p-value proche du seuil de significativité était le gène *TIE1*. Six variants hétérozygotes ont été identifiés chez six fœtus (Tableau 7). Il s'agit de variants faux-sens dont trois sont hérités de parents a priori asymptomatiques. Les variants ne sont pas regroupés dans un domaine particulier de la protéine TIE1 (Figure 32).

Fetus ID	Gène	Position cDNA (NM_005424.4)	Position protéine	Transmission	Fréquence dans gnomAD v2.1.1 (nombre d'allèles)	PolyPhen-2 (Score)	CADD score
F38	TIE1	c.1382G>A	p.(R461H)	ND	0.0028% (7)	Probably damaging (0.961)	25.2
F47	*TIE1*	c.207C>G	p.(I69M)	ND	0.0037% (7)	Probably damaging (0.997)	11.99
F55	*TIE1*	c.1297G>A	p.(G433S)	ND	0.0030% (7)	Probably damaging (1)	29.2
F101	*TIE1*	c.3053G>A	p.(R1018H)	Maternelle	0.00040% (1)	Probably damaging (1)	32
F103	*TIE1*	c.2462G>A	p.(R821Q)	Maternelle	0.0012% (3)	Probably damaging (0.994)	31
F112	*TIE1*	c.3052C>T	p.(R1018C)	Paternelle	0.00040% (1)	Probably damaging (1)	32

Tableau 7. Caractéristiques des variants candidats de *TIE1* identifiés chez les six fœtus index. ND : non déterminée.

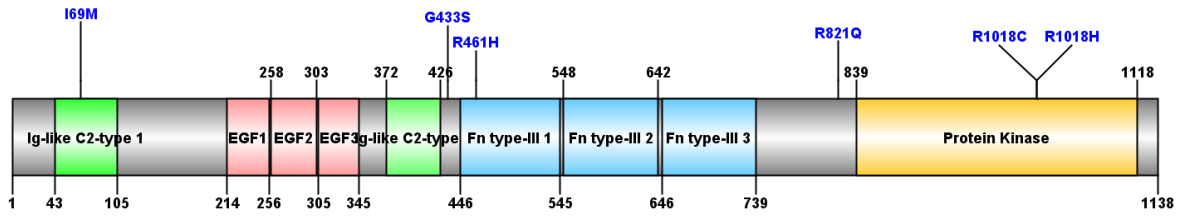


Figure 32. Représentation schématique de la protéine TIE1 et localisation des six variants identifiés chez les fœtus.

Les principales caractéristiques radiologiques et neuropathologiques de ces 6 fœtus sont décrites dans le tableau suivant (Tableau 8).

ID	Résultat échographie	Résultat IRM fœtale	Devenir grossesse	Examen neuropathologique	Autres anomalies extraneurologiques à l'examen foetopathologique
F38	Dilatation ventriculaire avec aspect échogène de l'hémisphère cérébral gauche	Nd	MFIU 28SA	Caillot sanguin dans ventricules / Nécrose corticale diffuse dans les régions fronto-pariétales / Hématome cérébral intraventriculaire avec perte de parenchyme cérébral par lésions d'autolyse	Epanchement séro-sanglant pleuro-péricardique et péritonéal
F47	Hydrocéphalie majeure avec rupture septale	Nd	IMG 34SA	Hydranencéphalie post-hémorragique	Poumons et reins congestifs, pâleur généralisée, macrocranie, pieds légèrement malposés
F55	Ventriculomégalie asymétrique avec lésions de porencéphalie	Ventriculomégalie bilatérale avec lésions de leucomalacie	IMG 35SA	Dilatation ventricule gauche avec lésions cavitaires et hémorragiques du centre semi ovale adjacent / Nombreux résidus hémorragiques intraventriculaires / Corps calleux fin	Avance staturo pondérale, hypertrophie rénale modérée, viscères globalement congestifs et hémorragiques
F101	Dilatation triventriculaire / Vermis et corps calleux non vus	Nd	IMG 25SA	Dilatation tétraventriculaire avec reliquats de remaniements post hémorragiques / Pas d'anomalie du vermis	Non
F103	Ventriculomegalie bilatérale majeur	Nd	IMG 30SA	Dilatation triventriculaire / Remaniements post-hémorragiques épendymaires et sous-épendymaires / Foyers d'hémorragies pétéchiales dans la substance blanche	Avance staturo-pondéral, dysmorphie facial
F112	RCIU / lésion hémisphère cérébelleux gauche	Asymétrie des hémisphères cérébelleux / Lésion ischémo-hémorragique de l'hémisphère cérébelleux gauche et d'une partie du vermis	IMG 31SA	Non effectué	Non effectué

Tableau 8. Principales caractéristiques radiologiques et neuropathologiques des fœtus porteurs d'un variant candidat de *TIE1*. Nd : non disponible.

- **Fonction du gène *TIE1***

Le gène *TIE1* est un gène de 22.22 kb présent sur le chromosome 1 (position 1p34.2). Il est constitué de 23 exons et code pour une protéine de 1138 acides aminés (126 kDa sous sa forme non glycosylée). La protéine TIE1 appartient à la famille des récepteurs à activité tyrosine

kinase impliqués dans les voies de signalisation intracellulaire. TIE1 est impliquée dans la croissance et la différenciation des cellules, l'angiogenèse et la réponse immunitaire.

Le gène *TIE1* est homologue au gène *TIE2* (également dénommé *TEK*), ils partagent en effet 80% d'homologie de séquence dans la partie intracellulaire (domaine tyrosine kinase).

TIE1 participe avec TIE2 et les angiopoïétines (ANGPT1 et ANGPT2) à la voie de signalisation angiopoïétine / TIE2 (Figure 33). Le récepteur TIE1 forme un hétérodimère avec TIE2. Il n'y a pas de ligand formellement identifié pour TIE1, raison pour laquelle il est qualifié de récepteur orphelin. Cependant, il a été récemment identifié une interaction in vitro entre TIE1 et SVEP1 (molécule de la matrice extracellulaire) par co-immunoprécipitation sur des cellules transfectées (Hußmann *et al.*, 2023). SVEP1 jouerait un rôle important dans la lymphangiogénèse chez la souris et le poisson zèbre ainsi que dans la formation des ansatomoses vasculaires (Karpanen *et al.*, 2017).

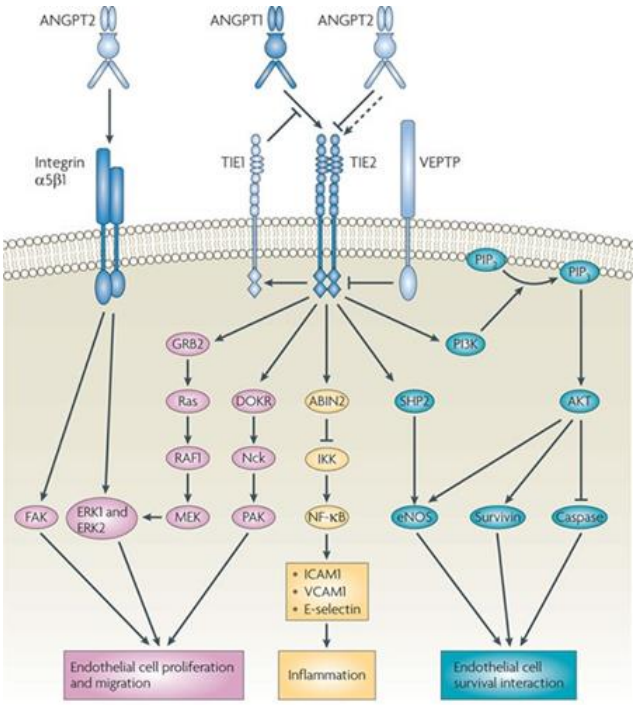


Figure 33. Voie de signalisation angiopoïétines / TIE2. Les multimères d'angiopoïétine 1 (ANGPT1) se lient à TIE2 sur les cellules endothéliales entrainant une dimérisation ou multimérisation de TIE2 puis son autophosphorylation dans sa partie intracellulaire. TIE1 est activée par l'autophosphorylation de TIE2 et le domaine extracellulaire de TIE1 interfère à son tour avec le complexe ANGPT1-TIE2 et antagonise ANGPT1. Extrait de Gillen *et al.*, 2019.

Par son action agoniste sur TIE2, il a été suggéré que l'ANGPT1 est nécessaire pour maintenir la quiescence des cellules endothéliales, favoriser la stabilité des vaisseaux, inhiber la fibrose

tissulaire et contribuer à la normalisation des vaisseaux. En revanche, l'ANGPT2 a une action qui peut être antagoniste dépendant du contexte sur cette voie, diminuant la stabilité vasculaire, favorisant l'activation endothéliale, la néoangiogenèse et le remodelage (Jeansson *et al.*, 2011).

- **Expression**

Les récepteurs TIE1 et TIE2 sont presque exclusivement exprimés dans les cellules endothéliales des vaisseaux sanguins et lymphatiques. *TIE1* est également exprimé, bien que dans une moindre mesure, dans certaines lignées de cellules hématopoïétiques (Eklund et al., 2016). Chez la souris, *TIE1* est fortement exprimée au niveau des cellules endothéliales des vaisseaux cérébraux (Figure 34).

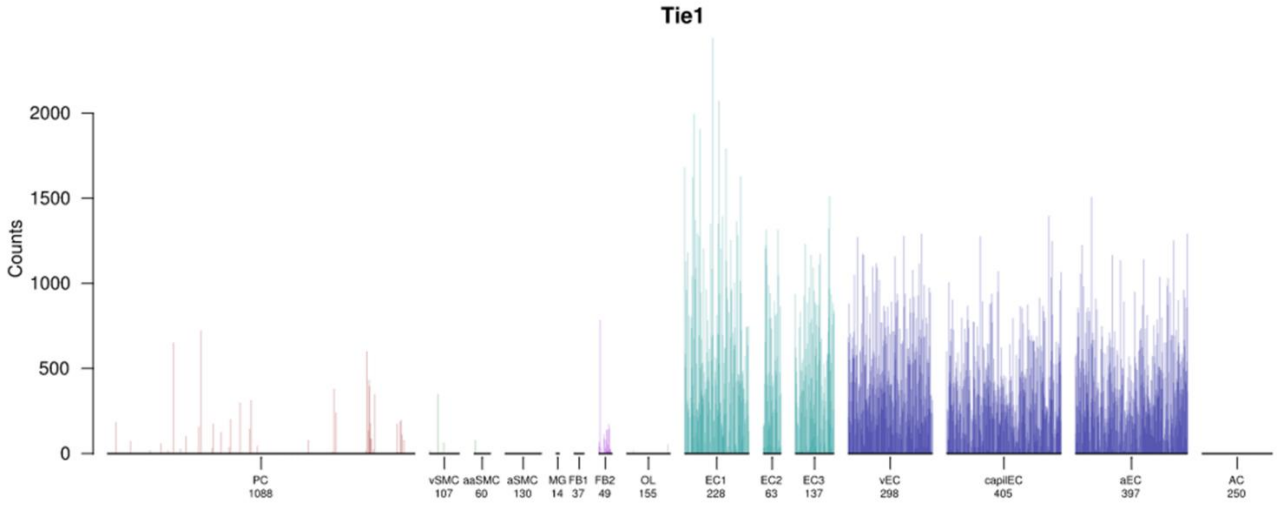


Figure 34. Profil d'expression de Tie1 dans les cellules vasculaires cérébrales de la souris. PC = péricyte ; SMC = cellules musculaire lisses ; EC= cellules endothéliales (vEC = venous EC ; aEC = arterial EC ; capilEC = capillary EC). ; AC = astrocyte. D'après https://betsholtzlab.org/VascularSingleCells/database.html.

- **Modèles animaux**

Chez le poisson zèbre, les embryons Tie1-/- présentent des défauts vasculaires (veine cardinale commune, formation altérée du plexus de la veine caudale) et une angiogenèse cérébrale réduite (Carlantoni *et al.*, 2021).

Chez la souris Tie1-/-, trois types d'anomalies ont été décrites :

- Un défaut du remodelage vasculaire. L'inactivation de Tie1 réduit la phosphorylation de Tie2 et l'activation d'Akt en aval empêchant le remodelage capillaire-veineux induit par ANGPT1 et

ANGPT2 (absence d'élargissement des vaisseaux sanguins au niveau de la trachée) (Korhonen *et al.*, 2016).

- Des hémorragies : les embryons homozygotes analysés à E18.5 présentent des hémorragies au niveau de l'extrémité de la queue, des orteils, du sinus sagittal de la tête et du dos, et dans les organes internes tels que les intestins et l'estomac. Tous les souriceaux décédèrent immédiatement après la naissance. 15% des embryons étatient décédés à E13.5 et présentaient également des hémorragies (Sato *et al.*, 1995).

- Des anomalies du système lymphatique : les embryons Tie-/- développent un œdème à E12.5 dans la zone dorsale du corps et du cou ainsi que des sacs lymphatiques de formes irrégulières sans signe de perte d'intégrité des vaisseaux sanguins à ce stade de développement. Ces anomalies vasculaires lymphatiques précèdent l'apparition d'hémorragies au stade embryonnaire E13.5 (D'Amico *et al.*, 2010).

- **Phénotype mendélien décrit chez l'homme**

TIE1 a été associé à une maladie autosomique dominante responsable de malformations lymphatiques (Michelini *et al.*, 2020). Ces auteurs ont mis en évidence des variants candidats chez 3 patients au sein d'une cohorte de 235 patients Italiens atteints de lymphœdème et négatifs pour les gènes connus déjà associés à un lymphœdème. Il s'agissait de variants hétérozygotes rares (MAF < 1/1000 dans gnomAD) et prédits délétères (selon les logiciels SIFT et PolyPhen-2) : p.(R436C), p.(E1016K) et p.(R1064H). Deux des patients présentaient une forme familiale de lymphœdème (une famille où l'index avaient 2 apparentés paucisymptomatiques et une famille dont la mère de l'index présentait également un lymphœdème). Un analyse structurale 3D pour ces variants concluait à une diminution du nombre d'interactions avec les résidus proches. Sur la base de ces résultats, les auteurs conclurent que *TIE1* est un gène candidat du lymphœdème. Cependant, il n'y a pas eu d'analyse statistique (recherche d'un enrichissement en variants candidats par rapport à une cohorte de contrôles) ni d'étude fonctionnelle pour établir la pathogénicité des trois variants.

Récemment, une possible implication de *TIE1* à l'état récessif dans une famille avec deux cas atteints de llymphœdème congénital et une lymphangiectasie intestinale a aussi été décrite. Le cas index était hétérozygote composite p.(G846R) / p.(S501N) mais son apparenté atteint n'avait pas été testé. Quant aux apparentés sains de cette famille, ils étaient hétérozygotes pour l'un des deux variants. Une étude fonctionnelle n'a pas non plus été réalisée (Abu Shtaya *et al.*,

2023). L'ensemble des ces données ne permet pas d'établir ni d'exclure la causalité de ces variants de *TIE1*.

- **Recherche d'anomalies vasculaires et lymphatiques chez les fœtus porteurs de variants de *TIE1***

Les anomalies vasculaires et lymphatiques rapportées chez la souris nous ont conduit à aller rechercher spécifiquement ces anomalies chez les fœtus porteurs de variants de *TIE1*.

Sur des coupes cérébrales d'un fœtus (F55) sur les six, il a été recherché des anomalies vasculaires de type anomalies de l'angiogenèse grâce à un immunomarquage anti-CD31, et sur des coupes de poumons il a été recherché des anomalies du système lymphatique à l'aide d'un immunomarquage D2-40. Il n'y avait pas de différence significative avec un témoin du même âge sur les coupes histologiques analysées (analyse réalisée par le Dr Jelena Martinovic).

L'analyse histologique des autres fœtus est en cours.

- **Recherche d'une correspondance phénotypique et génotypique avec d'autres patients via GeneMatcher**

GeneMatcher nous a permis de nous mettre en relation avec un chercheur ayant pour patients trois frères et sœurs d'une famille Allemande présentant des anévrismes intracrâniens et/ou une hémorragie sous-arachnoïdienne, et qui étaient porteurs d'un variant hétérozygote disruptif dans *TIE1* (p.Y1083X).

La nature de ce variant est différente de celle de nos variants (diruptif vs faux-sens) et il n'est actuellement pas possible de conclure en une similitude phénotypique. De plus, il s'avère que dans la base de données gnomAD il existe 134 individus porteurs de variants de type disruptif sur les 140 000 individus (fréquence voisine de 0.1%). Ces données de fréquence sont peu en faveur d'une implication de variants disruptifs à l'état hétérozygote dans cette pathologie anévrismale.

- **Perspectives d'étude**

Notre hypothèse de travail actuelle est la suivante : si les variants candidats observés chez nos fœtus à l'état hétérozygote (ou une partie d'entre eux) sont pathogènes ce serait très probablement au travers d'un gain de fonction que d'une perte de fonction, au vu de la fréquence

des variants perte de fonction dans les cohortes contrôles. Ceci requiert des explorations fonctionnelles avant que nous ne puissions l'établir.

Pour avancer dans l'interprétation de ces variants nous avons prévus plusieurs axes de travail :

1/ Une Analyse structurale pour prédire l'effet des variants candidats sur la protéine TIE1

La structure 3D déterminée par cristallographie de l'intégralité de la protéine TIE1 n'est pas disponible, seul le domaine Fibronectin-like 3 de TIE1 (résidus 642-738) a été cristallographié et aucun des variants candidats n'appartient à ce domaine. Néanmoins, une analyse in silico de la stabilité des six variants hétérozygotes de TIE1 par plusieurs logiciels de prédiction in silico montre une diminution de celle-ci. Une analyse structurale des variants avec une prédiction ou simulation de la structure 3D de TIE1 à partir de TIE2 est envisagée.

2/ Analyses fonctionnelles

Il est envisagé de réaliser une analyse de phosphorylation de la protéine TIE1 en présence ou en l'absence de ligand ANGPTL et ce dans des cellules exprimant TIE2. Il s'agit d'un test déjà décrit pour l'étude de variants dans le gène *TIE2* (Souma *et al.*, 2018).

Discussion

Les travaux réalisés m'ont conduit à tirer les conclusions suivantes :

- Les variations pathogènes de *COL4A1* et de *COL4A2* représentent la principale cause génétique des HIC fœtales avec environ 20% de fœtus porteurs d'une variation pathogène dans l'un de ces gènes. Dans 70% des cas il s'agit de variants apparus *de novo* et ces variants sont préférentiellement localisés dans la partie C-terminale du domaine de la triple hélice. Ces données ont une implication importante pour le conseil génétique.

- Outre *COL4A1* et *COL4A2*, il existe d'autres gènes responsables d'une HIC fœtale qui ne sont pas criblés de façon systémique comme les gènes associés à une thrombopénie sévère. C'est le cas des gènes *MPL* et *MECOM*. Pour *MPL*, antérieurement à ce travail de thèse, il avait déjà été rapporté des cas avec une HIC fœtale mais celle-ci n'apparaissait pas au premier plan dans le CAMT-MPL qui comme son nom l'indique est une pathologie congénitale détectée dans la plupart des cas dans la période néonatale. Pour *MECOM*, la présentation habituelle est celle d'une thrombocytopénie néonatale associée à des manifestations non hématologiques, et à notre de connaissance, il n'y avait pas eu de description de cas d'HIC fœtale. Étant donné que l'HIC fœtale n'est pas la présentation clinique habituelle de ces thrombopénies génétiques ces étiologies sont vraisemblablement sous-estimées.

- Il a été découvert une association entre les HIC fœtales et le gène *ESAM*, gène codant pour une protéine des jonctions serrées et qui n'était associé à aucune maladie monogénique connue auparavant. Cette découverte a permis d'élargir le spectre des anomalies cliniques associées aux molécules des jonction serrées.

- Les gènes associés aux HIC fœtales sont impliqués dans différentes voies biologiques. Ils codent pour des protéines intervenant dans la composition de la matrice extracellulaire, la synthèse des plaquettes, la coagulation / thrombose, les jonctions serrées et la réponse immunitaire.

- Certaines des anomalies phénotypiques observées chez les fœtus mutés dans les gènes *COL4A1/COL4A2* sont observées dans des pathologies fœtales non associées à une HIC. C'est le cas de plusieurs maladies mitochondriales.

- Cette analyse par WES sur une large cohorte de 113 fœtus, en plus de l'identification de plusieurs gènes causaux, a permis de retenir plusieurs gènes candidats qui nécessitent maintenant une validation de leur implication. Cette validation passera par l'établissement

d'une récurrence chez d'autres individus atteints d'HIC et des analyses in vitro sur des modèles cellulaires ou animaux adaptés.

- L'architecture génétique des HIC fœtales est complexe. Il existe une très grande hétérogénéité génétique avec pour chacun des gènes identifiés, hormis *COL4A1* et *COL4A2*, un très petit pourcentage de fœtus concernés.

- **Variants pathogènes de *COL4A1/COL4A2* chez le fœtus et implication pour le conseil génétique**

Il existe une grande variabilité interindividuelle et intrafamiliale des signes clinicoradiologiques associés aux variants pathogènes de *COL4A1* et de *COL4A2*. Ces signes peuvent se manifester aussi bien durant la vie fœtale qu'à l'âge adulte. Il existe par ailleurs une pénétrance incomplète qui n'est pas connue avec précision. Ces différents éléments rendent le conseil génétique difficile particulièrement chez les apparentés asymptomatiques. Par un travail de synthèse des données moléculaires obtenues ces 10 dernières années chez le fœtus, nous avons pu montrer un pourcentage élevé de variations *de novo* (70%) et une localisation différente par rapport aux variants identifiés chez l'adulte atteint d'une leucoencéphalopathie vasculaire. Le taux élevé de variants *de novo* permet de rassurer un couple quant au risque de récidive de l'évènement hémorragique lors d'une nouvelle grossesse.

Cependant, pour les 30% hérités d'un parent en général asymptomatique, les mécanismes sous tendant cette pénétrance incomplète sont inconnus. Une étude effectuée chez un enfant atteint d'une porencéphalie et chez son père asymptomatique, tous deux porteurs d'une variation pathogène de *COL4A2*, montra de manière intéressante des défauts de la membrane basale chez les deux individus, mais la rétention de COL4A2 dans le réticulum endoplasmique (RE) était seulement observée pour les cellules de l'enfant atteint. Ces cellules présentaient un stress et une apoptose accrue. Cela implique donc que la capacité des cellules à faire face à un défaut de repliement du collagène muté et à une rétention dans le RE est variable chez des individus pourtant porteurs d'une même variation pathogène et pourrait expliquer une partie de la variabilité dans l'expression de cette affection. En outre, le traitement des fibroblastes du patient avec une molécule chaperonne chimique (acide 4-phénylbutyrique (PBA)) réduisit efficacement l'accumulation intracellulaire de COL4A2, le stress du réticulum endoplasmique et l'apoptose (Murray *et al.*, 2014). Ces observations ont été confirmé par traitement avec du PBA de souris porteuses d'une variation faux-sens glycine de Col4a1 (Jones *et al.*, 2019) ; avec démonstration d'une réduction du stress du RE, d'une augmentation de collagène de type IV

dans la membrane basale et d'une diminution de l'étendue des lésions hémorragiques cérébrales sur les coupes histologiques. Ces données suggèrent que la modification de l'expression des protéines chaperonnes, de l'accumulation intracellulaire de collagène et du stress du RE sont des éléments pouvant expliquer la variabilité phénotypique des collagénopathie COL4A1/A2 et des options thérapeutiques potentielles chez les patients porteurs.

Un autre facteur modificateur possible est celui de la voie du TGFβ, jouant un rôle important dans le développement vasculaire (Arnold *et al.*, 2014). De données récentes chez les souris mutantes Col4a1 montrent que COL4A1/COLA42 module la signalisation TGFβ ce qui suggère la possibilité qu'une signalisation TGFβ altérée puisse contribuer aux lésions vasculaires cérébrales. En outre, il a été montré chez la souris mutante Col4a1 que l'inactivation de la signalisation TGFβ améliore la pathologie vasculaire cérébrale, avec prévention de la perte des cellules musculaires lisses et réduction significative des HIC (Branyan *et al.*, 2023). Des études chez l'humain sont nécessaires pour étudier la contribution de la voie du TGFβ aux anomalies vasculaires de *COLA41/COL4A2*.

- ## Architecture génétique actuelle des HIC fœtales

Cette analyse par WES d'une cohorte de 113 fœtus référés pour un diagnostic étiologique d'HIC après un criblage moléculaire négatif des gènes *COL4A1/COL4A2* ainsi qu'une revue de la littérature, nous permet de dresser un bilan actuel des déterminants génétiques. Les résultats moléculaires obtenus chez les 113 fœtus sont résumés dans le tableau en annexe 4. Différents gènes impliqués dans plusieurs voies biologiques peuvent être responsables d'une HIC fœtale, suggérant une grande hétérogénéité génétique. Les voies biologiques affectées concernent à la fois le contenu et le contenant des petits vaisseaux cérébraux (Figure 35).

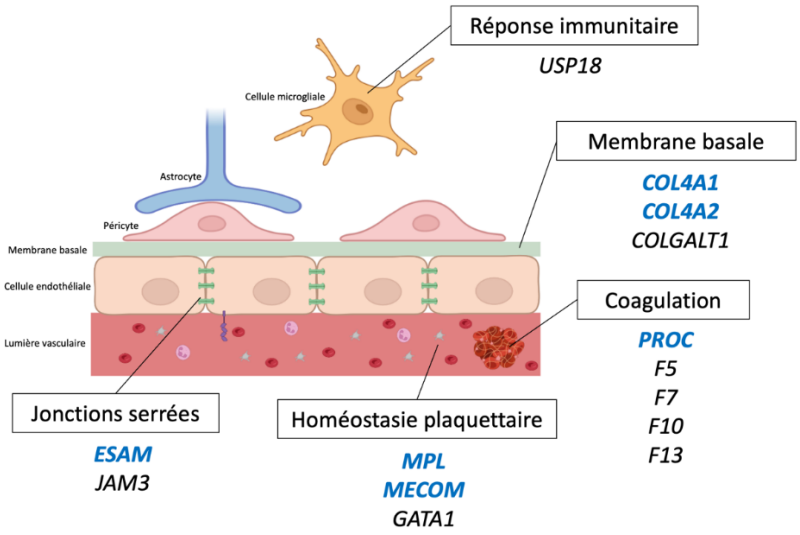


Figure 35. Voies biologiques et gènes impliqués dans les HIC fœtales. Les gènes non colorés en bleu sont ceux identifiés dans d'autres études.

Au niveau du contenu vasculaire, il peut s'agir d'un défaut de synthèse des plaquettes conduisant à une thrombopénie sévère. Nous avons identifié des variations pathogènes dans deux gènes essentiels à la mégacaryopoïèse, *MPL* et *MECOM*, dans deux familles avec deux fœtus atteints. Chez l'adulte, une thrombopénie sévère est responsable de saignements de type cutanéo-muqueux dans la majorité des cas et très rarement d'une HIC (Estcourt *et al.*, 2014). Chez le fœtus, ces saignements sont localisés au niveau intracrânien particulièrement au niveau de la matrice germinale car il s'agit d'une zone fragile hypervascularisée. Les mécanismes responsables de la survenue d'un saignement intracrânien chez seulement une partie des fœtus porteurs de ces variations pathogènes sont inconnus. Par exemple, dans le cas du CAMT-MPL, la thrombopénie sévère entraine une HIC fœtale dans 15% des cas mais pour la majorité des patients celle-ci passe inaperçue durant la grossesse et se manifeste chez le nouveau-né par un tableau de saignements cutanéo-muqueux (Germeshausen *et al.*, 2020). La thrombopénie à elle seule ne peut donc pas expliquer la présence d'un saignement spontané. Des hypothèses ont été émises pour comprendre cette différence d'expressivité. Certains auteurs ont évoqué le rôle de l'inflammation comme élément déclencheur d'hémorragies chez des souris thrombopéniques (George *et al.*, 2008). Un autre gène impliqué dans la mégacaryopoïèse est le gène *GATA1* responsable d'une macrothrombocytopénie sévère en cas de variation pathogène et déjà décrit comme responsable d'HIC fœtale (Bouchghoul *et al.*, 2018). Pour *GATA1*, la prévalence des variations pathogènes est extrêmement faible et il n'a pas été identifié de variant candidat chez les 113 fœtus. Pour ces gènes impliqués dans la synthèse plaquettaire, il est important de souligner que bien que l'hémorragie soit la principale complication, ils sont également associés

à des hémopathies malignes (Almazni *et al.,* 2019), ce qui est un élément capital pour le conseil génétique des apparentés.

A côté des anomalies de la synthèse plaquettaire, des anomalies des gènes impliqués dans la coagulation peuvent être responsables d'une HIC fœtale. Les déficits en facteurs de la coagulation sont des maladies récessives rares, d'expressivité variable et dont la prévalence chez le fœtus avec HIC pourrait être sous-estimée. Nous avons identifié une famille avec un fœtus (F12) porteur de variants pathogènes bialléliques dans le gène de la protéine C. La présence de variants bialléliques dans *PROC* est connue pour être responsable d'un déficit sévère en protéine C et est associée chez le nouveau-né à une coagulation intravasculaire disséminée et un purpura fulminans par thrombose microvasculaire suivie d'une hémorragie périvasculaire et d'une nécrose (Marlar *et al.,* 1990). Cependant le déficit homozygote en protéine C peut se manifester en période anténatale par une HIC avec ventriculomégalie secondaire à une thrombose (Martin *et al.,* 2021). En raison de son incidence relativement rare, ce déficit sévère pourrait également être sous-estimé chez le fœtus atteint d'une HIC (Fong *et al.,* 2010 ; Martin *et al.,* 2021).

Au niveau du <u>contenant vasculaire,</u> on peut retrouver des altérations des protéines des jonctions serrées assurant l'intégrité de la BHE. Il était précédemment connu que des variants pathogènes dans le gène *JAM3* pouvaient entrainer une HIC chez le fœtus. D'autres gènes codant pour des protéines de jonctions serrées ont été associées à un phénotype cérébral anténatal mais sans signe hémorragique. Par l'identification d'une famille avec un variant perte de fonction biallélique dans le gène *ESAM* puis un travail collaboratif, le spectre des maladies associées aux molécules de jonction a pu être étendu. Il n'est pas encore élucidé pourquoi certaines molécules jonctionnelles sont responsables d'une HIC et pas d'autres (Mochida *et al.,* 2010) mais vraisemblablement une différence dans leur profil d'expression et leurs implications biologiques annexes pourraient expliquer cette particularité. Il a été montré par exemple, dans des expériences d'inactivation génique chez la souris, que la protéine ESAM, outre le fait d'assurer la cohésion intercellulaire, joue un rôle dans l'angiogenèse (Ishida *et al.,* 2003). De plus, sur les données de la souris, *ESAM* a une forte expression au niveau des péricytes et des cellules musculaires vasculaires lisses à la différence des autres molécules des jonctions serrées pour lesquelles il n'y a pas de manifestions hémorragiques cérébrales.

Au sein de notre cohorte nous avons par ailleurs identifié des variants candidats dans le gène *PLOD3,* un gène intervenant dans la biosynthèse de *COL4A1* et *COL4A2* et donc participant à

l'élaboration de la membrane basale (ou matrice extracellulaire). La biosynthèse et l'assemblage du collagène impliquent de multiples modifications post-traductionnelles des chaînes polypeptidiques, notamment l'hydroxylation des résidus proline, l'hydroxylation des résidus lysine et la glycosylation. Des variations pathogènes dans les gènes intervenant dans ces étapes sont responsables de diverses maladies du tissu conjonctif. L'hydroxylation des résidus lysine, leur glycosylation ultérieure avant l'association des chaînes et la formation d'une structure en triple hélice constituent l'une des modifications post-traductionnelles importantes des collagènes, et sont catalysées par les enzymes LH1, LH2 et LH3 (codées par les gènes *PLOD1*, *PLOD2* et *PLOD3*). En outre, LH3 semble être la seule responsable des activités de galactosyltransférase (GT) et de glucosyltransférase. LH3 exerce ses activités de lysyltransférase et de GT sur plusieurs collagènes, en particulier les types IV, VI, XI et XII ; les variants pathogènes de *PLOD3* peuvent ainsi affecter une variété de protéines, expliquant ainsi les caractéristiques multisystémiques observées chez les personnes affectées (Ewans *et al.*, 2019). Les variants pathogènes de *PLOD3* sont à l'origine d'une maladie récessive du tissu conjonctif dont les caractéristiques cliniques sont proches du syndrome de Stickler (Salo *et al.*, 2008 ; Ewans *et al.*, 2019). Notre étude nous a conduit à suspecter son implication dans les HIC fœtales mais une étude fonctionnelle des variants candidats identifiés est indispensable pour conclure.

Nous n'avons pas observé de variants pathogènes ou candidats dans le gène *COLGALT1*, autre gène intervenant dans la biosynthèse de COL4A1/COL4A2.

La réponse immunitaire est aussi une voie impliquée dans les HIC fœtales. Dans cette voie, on y trouve le gène *USP18* codant pour une protéine régulatrice négative de la signalisation à l'interféron de type I (IFN-I). Le contrôle du système IFN-I par USP18 est essentiel à la quiescence de la microglie et à la prévention de lésions tissulaires incontrôlées (Goldmann *et al.*, 2015). En l'absence d'USP18, les cellules microgliales subissent une activation prolongée de la voie de signalisation JAK/STAT entrainant une microgliopathie (Honke *et al.*, 20016). Les variants pathogènes bialléliques d'USP18 sont responsables du pseudo-TORCH de type 2 qui est une interféronopathie caractérisée par une hémorragie intracrânienne, des calcifications, des malformations cérébrales, un dysfonctionnement hépatique, des présentations de type choc septique et une thrombocytopénie sévère (Meuwissen *et al.*, 2016). L'origine de la thrombopénie dans les interféronopathies n'est pas connue (Adang *et al.*, 2022) et semble par ailleurs inconstante (Misk *et al.*, 2022). Nous n'avons pas identifié de fœtus porteur de variants pathogènes ou candidats dans le gène *USP18*. La prévalence du pseudo-TORCH de type 2 est

de 1 /1 000 000 (Sehrish *et al.,* 2022) laissant à penser qu'il s'agit de cas extrêmement rares comme pour ceux associés aux gènes *GATA1* et *COLGALT1*. La taille notre cohorte, bien que conséquente, n'est probablement pas assez grande pour identifier de nouveaux cas associés à ces gènes.

Outres les gènes causaux, nous avons identifiés plusieurs gènes candidats appartenant également à diverses voies biologiques : angiogenèse (*TIE1*), système ubiquitine-protéasome (*DCAF5*), hémostase (*ITGA2B*), réparation de l'ADN (*EXD3*), cytosquelette (*ZFYVE28*), inflammation (*RNPEP*). Ceci suggère une architecture moléculaire complexe. Parmi les gènes candidats identifiés certains pourraient être responsables d'une HIC chez le fœtus mais ils ne sont actuellement pas démontrés comme étant responsables de cette condition. Des études complémentaires sont nécessaires afin de mieux préciser leur implication ou non dans la survenue d'une HIC chez le fœtus. Une des premières étapes est la recherche d'une récurrence. Cette récurrence est un élément important pour affirmer la causalité des gènes candidats. Pour augmenter la puissance statistique, nous disposons d'une cohorte d'enfants âgés de moins de 4 ans au moment du prélèvement et dont les ADN ont été référés au laboratoire de génétique neurovasculaire de l'hôpital Saint Louis devant la présence de lésions ischémo-hémorragiques cérébrales ayant fait suspecter une collagénopathie COL4A1/A2. Au préalable, comme pour la cohorte de fœtus, les ADN de ces enfants ont eu un séquençage ciblé de *COL4A1* et de *COL4A2* et aucune variation pathogène ou probablement pathogène n'a été identifiée. Un séquençage exonique d'un panel de gènes candidats est en cours de réalisation pour cette cohorte d'enfants.

- **Similitudes phénotypiques entre collagénopathies COL4A1/A2 et mitochondriopathies chez le fœtus**

Plusieurs variants pathogènes dans des gènes impliqués dans le métabolisme mitochondrial ont été identifiés parmi les 113 fœtus référés pour HIC. Nous avons identifié chez un fœtus consanguin un variant perte de fonction homozygote dans le gène *COQ2,* codant pour une 4-hydroxybenzoate polyprényltransférase mitochondriale. Les variants pathogènes bialléliques de COQ2 sont associées à un déficit en coenzyme Q10 (Jakobs *et al.,* 2013) dont l'apparition de la maladie va de la période néonatale à la fin de l'âge adulte. Cette condition entraîne des symptômes d'encéphalopathie. Certains patients peuvent présenter des lésions cérébrales ressemblant au syndrome de Leigh ou au syndrome de MELAS (avec des épisodes ressemblant à des accidents vasculaires cérébraux) (Munch *et al.,* 2023). En prénatal, les lésions cérébrales

décrites sont très variables et peuvent comprendre une leucoencéphalopathie, des kystes périventriculaires, une atrophie cérébelleuse, une agénésie du corps calleux et une ventriculomegalie, symptômes pouvant être présents chez les foetus mutés *COL4A1/COL4A2* (Hashemi *et al.,* 2020 ; Munch *et al.,* 2023).

Un autre fœtus consanguin était porteur d'un variant homozygote entrainant une anomalie d'épissage dans le gène *ATP5PO*. Ce gène code pour une sous-unité du complexe ATP synthase responsable de la production d'adénosine triphosphate (ATP), principale source d'énergie utilisée par les cellules. Ce gène est associé à une maladie autosomique récessive caractérisée par une hypotonie et un retard de développement dès la naissance. Peu de cas sont rapportés dans la littérature et les patients décrits présentent des lésions cérébrales du type atrophie, ventriculomégalie, hypoplasie du corps calleux, anomalies de la fosse postérieure, kystes périventriculaires et leucomalacie (Zech *et al.,* 2022 ; Ganapathi *et al.,* 2022). Chez un patient avec un variant pathogène de *ATP5PO* il a également été observé la présence de zones d'hémorragie anciennes bilatérales dans les lobes frontal et pariétal (Ganapathi *et al.,* 2022).

Par ailleurs, en utilisant une approche de type test de charge en variants candidats à l'échelle du gène, nous avons identifié trois fœtus féminins avec des variants pathogènes dans *PDHA1*, responsable d'un déficit en pyruvate déshydrogénase dont lésions cérébrales comprenaient ventriculomégalie, leucomalacie, kystes, dysgénésie du corps calleux et anomalies de la fosse postérieure.

Enfin nous avons identifié un variant homozygote faux-sens probablement pathogène dans le gène *NAXD*, qui est associé à une maladie neurodégénérative d'apparition infantile comprenant une encéphalopathie avec œdème cérébral et/ou une leucoencéphalopathie (Van Bergen *et al.,* 2018 ; Majethia *et al.,* 2021).

Dans toutes ces conditions, le spectre clinique et radiologique est très vaste et certains signes sont similaires à ceux rencontrés en contexte d'hémorragie cérébrale fœtale, notamment la ventriculomégalie et les lésions clastiques du parenchyme cérébral associées à une atteinte de la substance blanche. Pour ces raisons, les fœtus présentant des variants pathogènes dans les gènes du métabolisme mitochondrial avaient été initialement suspectés d'HIC fœtales et référés pour un criblage moléculaire des gènes *COL4A1* et *COL4A2*. La reprise à posteriori des éléments radiologiques et neuropathologiques (par les fœtopathologistes et les neuroradiologues) a montré que le phénotype de ces fœtus était compatible avec une maladie mitochondriale. Ces résultats illustrent la complexité de l'interprétation du phénotype chez le fœtus, ainsi que la difficulté de l'orientation étiologique lorsque certains signes associés de

différentes affections sont similaires. Ils soulignent également l'utilité des approches de criblage pangénomique (WES/WGS) pour le diagnostic moléculaire chez le fœtus et l'utilité du rétrophénotypage pour s'assurer de la cohérence avec le diagnostic moléculaire établi.

- ## Les déterminants génétiques des HIC chez le fœtus restent largement indéterminés

Une autre conclusion importante de cette étude est que les causes génétiques d'HIC sont hétérogènes et largement non déterminées. Nous avons été confrontés au problème d'une très faible récurrence pour chacun de nos gènes candidats. La faible récurrence rend difficile l'identification de gènes causaux et impose une analyse chronophage gène par gène ainsi que le partage de données via GeneMatcher.

Enfin, la présence d'une majorité de cas sporadiques associée à un nombre important de fœtus sans variant causal ou candidat après WES soulève plusieurs hypothèses non mutuellement exclusives :

i) Une hérédité complexe avec possiblement plusieurs déterminants génétiques impliqués. L'étude de modèles digéniques ou oligogéniques avec des approches basées sur des réseaux de gènes pour étudier les interactions complexes entre différents gènes pourrait identifier d'autres gènes causaux potentiels.

ii) L'existence de variations dans des régions introniques ou régulatrices, des variants structuraux ou des expansions que ne sont pas détectés par WES et qui nécessiterait d'utiliser une technique de WGS.

iii) Des anomalies épigénétiques telles que des anomalies de méthylation de l'ADN, des modifications des histones et/ou une dérégulation de l'expression de certains ARN non codants (Zuccarello *et al.*, 2021).

L'identification de ces gènes est essentielle pour le diagnostic, l'information et le conseil génétique des apparentés, mais aussi pour comprendre les mécanismes physiopathologiques sous-jacents à ces HIC fœtales et in fine développer des approches thérapeutiques adaptées.

Conclusion et Perspectives

Ce travail de thèse sur une large cohorte de fœtus référés pour suspicion de collagénopathie COL4A1/A2 nous a permis i) d'identifier plusieurs gènes causaux et candidats ; ii) de mettre en évidence la très grande hétérogénéité génétique des HIC ; iii) de montrer la difficulté des orientations diagnostiques sur la base des données phénotypiques neuroradiologiques et l'intérêt majeur d'une approche moléculaire pangénome. Malgré ce travail, il reste encore de nombreuses causes génétiques aux HIC fœtales à découvrir et l'analyse des données de WES/WGS sur de larges cohortes de fœtus dans le cadre de projets collaboratifs internationaux combinée à l'utilisation de plateformes d'échanges de données nous semble indispensable pour identifier de nouveaux acteurs moléculaires.

Ce travail a conduit au transfert au sein du service de génétique neurovasculaire de l'hôpital Saint Louis, des applications de ces données avec mise en place d'un panel de gènes à visée diagnostique incluant en plus des gènes *COL4A1/COL4A2* tous les gènes causaux et candidats identifiés dans cette thèse. Chez les fœtus avec HIC adressés pour diagnostic et chez qui ce panel se révèlera négatif, une analyse par WES ou WGS sera envisagée. Nous avons fait actuellement le choix du panel en première intention car il présente plusieurs avantages par rapport aux analyses pangénomiques : i) l'analyse est moins couteuse en ressources informatiques ; ii) la couverture des régions codantes est meilleure ; iii) l'analyse permet de se focaliser sur des variants dans des gènes déjà causaux/candidats qui auraient été négligés/noyés par les nombreux variants identifiés lors d'une analyse pangénomique.

La poursuite du projet comprendra différents axes : i) analyses fonctionnelles des variants identifiés dans certains gènes candidats, notamment le gène *TIE1* ; ii) augmentation du nombre d'index analysés (en particulier au travers de projets collaboratifs) ; iii) relecture des données de WES des 113 index en utilisant des filtres moins stringents particulièrement pour les variants bialléliques ; iv) analyse des gènes associés à un phénotype hémorragique chez la souris (liste en annexe 5) ; v) séquençage du génome entier pour les trios pour rechercher d'autres anomalies moléculaires non observables sur l'analyse d'exome entier ; vi) utilisation de techniques « omiques » sur tissu fœtal (si disponible) comme l'analyse du transcriptome permettant d'identifier des anomalies d'expression de gènes ou des anomalies d'épissage.

Références Bibliographiques

Abu Shtaya A, Sukenik-Halevy R, Bazak L, Lidzbarsky GA, Gonzaga-Jauregui C, Lagovsky I, et al. Possible biallelic inheritance in TIE1 in a family with congenital lymphedema, intestinal lymphangiectasia and cutis aplasia. Clin Genet. 2023 Aug;104(2):275-276.

Adang LA, Gavazzi F, D'Aiello R, Isaacs D, Bronner N, Arici ZS, et al. Hematologic abnormalities in Aicardi Goutières Syndrome. Mol Genet Metab. 2022 Aug;136(4):324-329.

Adiego B, Martínez-Ten P, Bermejo C, Estévez M, Recio Rodriguez M, Illescas T. Fetal intracranial hemorrhage. Prenatal diagnosis and postnatal outcomes. J Matern Fetal Neonatal Med. 2019 Jan;32(1):21-30.

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010 Apr;7(4):248-9.

Alamowitch S, Plaisier E, Favrole P, Prost C, Chen Z, Van Agtmael T, Marro B, Ronco P. Cerebrovascular disease related to COL4A1 mutations in HANAC syndrome. Neurology. 2009 Dec 1;73(22):1873-82.

Almazni I, Stapley R, Morgan NV. Inherited Thrombocytopenia: Update on Genes and Genetic Variants Which may be Associated With Bleeding. Front Cardiovasc Med. 2019 Jun 19;6:80.

Anstrom JA, Thore CR, Moody DM, Brown WR. Immunolocalization of tight junction proteins in blood vessels in human germinal matrix and cortex. Histochem Cell Biol. 2007 Feb;127(2):205-13.

Arnold TD, Niaudet C, Pang MF, Siegenthaler J, Gaengel K, Jung B, et al. Excessive vascular sprouting underlies cerebral hemorrhage in mice lacking αVβ8-TGFβ signaling in the brain. Development. 2014 Dec;141(23):4489-99.

Backenroth D, Homsy J, Murillo LR, Glessner J, Lin E, Brueckner M, et al. CANOES: detecting rare copy number variants from whole exome sequencing data. Nucleic Acids Res. 2014 Jul;42(12):e97.

Bastida JM, Lozano ML, Benito R, Janusz K, Palma-Barqueros V, Del Rey M, et al. Introducing high-throughput sequencing into mainstream genetic diagnosis practice in inherited platelet disorders. Haematologica. 2018 Jan;103(1):148-162.

Bellanné-Chantelot C, Mosca M, Marty C, Favier R, Vainchenker W, Plo I. Identification of MPL R102P Mutation in Hereditary Thrombocytosis. Front Endocrinol (Lausanne). 2017 Sep 20;8:235.

Berrut G, Ghali A, Quere I, Ternisien C, Gallois I, Roy PM, Marre M, Fressinaud P. La mutation C677T du gène de la 5,10-méthyltétrahydrofolate réductase est associée aux thromboses veineuses idiopathiques. Rev Med Interne. 2003 Sep;24(9):569-76.

Bouchghoul H, Quelin C, Loget P, Encha-Razavi F, Senat MV, Maheut L, et al. Fetal cerebral hemorrhage due to X-linked GATA1 gene mutation. Prenat Diagn. 2018 Sep;38(10):772-778.

Branyan K, Labelle-Dumais C, Wang X, Hayashi G, Lee B, Peltz Z, et al. Elevated TGFβ signaling contributes to cerebral small vessel disease in mouse models of Gould syndrome. Matrix Biol. 2023 Jan;115:48-70.

Breedveld G, de Coo IF, Lequin MH, Arts WF, Heutink P, Gould DB, et al. Novel mutations in three families confirm a major role of COL4A1 in hereditary porencephaly. J Med Genet. 2006 Jun;43(6):490-5.

Breysem L, Bosmans H, Dymarkowski S, Schoubroeck DV, Witters I, Deprest J, et al. The value of fast MR imaging as an adjunct to ultrasound in prenatal diagnosis. Eur Radiol. 2003 Jul;13(7):1538-48.

Briggs TA, Wolf NI, D'Arrigo S, Ebinger F, Harting I, Dobyns WB, et al. Band-like intracranial calcification with simplified gyration and polymicrogyria: a distinct "pseudo-TORCH" phenotype. Am J Med Genet A. 2008 Dec 15;146A(24):3173-80.

Buffon F. Collagénopathies COL4A1 et COL4A2. Sang Thrombose Vaisseaux. 2019;31(4):159-163.

Burstein J, Papile LA, Burstein R. Intraventricular hemorrhage and hydrocephalus in premature newborns: a prospective study with CT. AJR Am J Roentgenol. 1979 Apr;132(4):631-5.

Carlantoni C, Allanki S, Kontarakis Z, Rossi A, Piesker J, Günther S, Stainier DYR. Tie1 regulates zebrafish cardiac morphogenesis through Tolloid-like 1 expression. Dev Biol. 2021 Jan 1;469:54-67.

Cavaliere AF, Turrini I, Pallottini M, Vidiri A, Marchi L, Perelli F, et al. Genetic Profiling of Idiopathic Antenatal Intracranial Haemorrhage: What We Know? Genes (Basel). 2021 Apr 15;12(4):573.

Cavallin M, Mine M, Philbert M, Boddaert N, Lepage JM, Coste T, et al. Further refinement of COL4A1 and COL4A2 related cortical malformations. Eur J Med Genet. 2018 Dec;61(12):765-772.

Cen Z, Chen Y, Chen S, Wang H, Yang D, Zhang H, et al. Biallelic loss-of-function mutations in JAM2 cause primary familial brain calcification. Brain. 2020 Feb 1;143(2):491-502.

Chen J, Zhang P, Peng M, Liu B, Wang X, Du S, et al. An additional whole-exome sequencing study in 102 panel-undiagnosed patients: A retrospective study in a Chinese craniosynostosis cohort. Front Genet. 2022 Sep 2;13:967688.

Chung WK, Roberts TP, Sherr EH, Snyder LG, Spiro JE. 16p11.2 deletion syndrome. Curr Opin Genet Dev. 2021 Jun;68:49-56.

Cirulli ET. The Increasing Importance of Gene-Based Analyses. PLoS Genet. 2016 Apr 7;12(4):e1005852.

Cirulli ET, White S, Read RW, Elhanan G, Metcalf WJ, Tanudjaja F, et al. Genome-wide rare variant analysis for thousands of phenotypes in over 70,000 exomes from two cohorts. Nat Commun. 2020 Jan 28;11(1):542.

Colin E, Sentilhes L, Sarfati A, Mine M, Guichet A, Ploton C, et al. Fetal intracerebral hemorrhage and cataract: think COL4A1. J Perinatol. 2014 Jan;34(1):75-7.

Crespin M, Alhenc-Gelas M, Grangé G, Fallet-Bianco C, Fontenay M. Fetal intracerebral hemorrhage in familial thrombophilia. Pediatr Neurol. 2009 Oct;41(4):291-3.

Curtis C, Mineyko A, Massicotte P, Leaker M, Jiang XY, Floer A, Kirton A. Thrombophilia risk is not increased in children after perinatal stroke. Blood. 2017 May 18;129(20):2793-2800.

D'Amico G, Korhonen EA, Waltari M, Saharinen P, Laakkonen P, Alitalo K. Loss of endothelial Tie1 receptor impairs lymphatic vessel development-brief report. Arterioscler Thromb Vasc Biol. 2010 Feb;30(2):207-9.

de Sousa C, Clark T, Bradshaw A. Antenatally diagnosed subdural haemorrhage in congenital factor X deficiency. Arch Dis Child. 1988 Oct;63(10 Spec No):1168-70.

de Vries LS, Koopman C, Groenendaal F, Van Schooneveld M, Verheijen FW, Verbeek E, et al. COL4A1 mutation in two preterm siblings with antenatal onset of parenchymal hemorrhage. Ann Neurol. 2009 Jan;65(1):12-8.

Dicuonzo F, Palma M, Fiume M, Scarpello R, Lefons V, Maghenzani M, Carella A. Cerebrovascular disorders in the prenatal period. J Child Neurol. 2008 Nov;23(11):1260-6.

Downes K, Megy K, Duarte D, Vries M, Gebhart J, Hofer S, et al. Diagnostic high-throughput sequencing of 2396 patients with bleeding, thrombotic, and platelet disorders. Blood. 2019 Dec 5;134(23):2082-2091.

Durrani-Kolarik S, Manickam K, Chen B. COL4A1 Mutation in a Neonate With Intrauterine Stroke and Anterior Segment Dysgenesis. Pediatr Neurol. 2017 Jan;66:100-103.

Eklund L, Kangas J, Saharinen P. Angiopoietin-Tie signalling in the cardiovascular and lymphatic systems. Clin Sci (Lond). 2017 Jan 1;131(1):87-103.

Elchalal U, Yagel S, Gomori JM, Porat S, Beni-Adani L, Yanai N, Nadjari M. Fetal intracranial hemorrhage (fetal stroke): does grade matter? Ultrasound Obstet Gynecol. 2005 Sep;26(3):233-43.

Ellestad SC, Zimmerman SA, Thornburg C, Mitchell TE, Swamy GK, James AH. Severe factor V deficiency presenting with intracranial haemorrhage during gestation. Haemophilia. 2007 Jul;13(4):432-4.

England EC, Cornejo P, Neilson DE, Rao RP, Goncalves LF. Fetal brain small vessel disease 1 caused by a novel mutation in the COL4A1 gene. Pediatr Radiol. 2021 Mar;51(3):480-484.

Eshghi P, Cohan N, Naderi M, Karimi M. Factor XIII deficiency: a review of literature. 2012;4(2).

Estcourt LJ, Stanworth SJ, Collett D, Murphy MF. Intracranial haemorrhage in thrombocytopenic haematology patients--a nested case-control study: the InCiTe study protocol. BMJ Open. 2014 Feb 7;4(2):e004199.

Ewans LJ, Colley A, Gaston-Massuet C, Gualtieri A, Cowley MJ, McCabe MJ, et al. Pathogenic variants in PLOD3 result in a Stickler syndrome-like connective tissue disorder with vascular complications. J Med Genet. 2019 Sep;56(9):629-638.

Fong CY, Mumford AD, Likeman MJ, Jardine PE. Cerebral palsy in siblings caused by compound heterozygous mutations in the gene encoding protein C. Dev Med Child Neurol. 2010 May;52(5):489-93.

Gambin T, Akdemir ZC, Yuan B, Gu S, Chiang T, Carvalho CMB, et al. Homozygous and hemizygous CNV detection from exome sequencing data in a Mendelian disease cohort. Nucleic Acids Res. 2017 Feb 28;45(4):1633-1648.

Ganapathi M, Friocourt G, Gueguen N, Friederich MW, Le Gac G, Okur V, et al. A homozygous splice variant in ATP5PO, disrupts mitochondrial complex V function and causes Leigh syndrome in two unrelated families. J Inherit Metab Dis. 2022 Sep;45(5):996-1012.

Garel C, Rosenblatt J, Moutard ML, Heron D, Gelot A, Gonzales M, Miné E, Jouannic JM. Fetal intracerebral hemorrhage and COL4A1 mutation: promise and uncertainty. Ultrasound Obstet Gynecol. 2013 Feb;41(2):228-30.

Germeshausen M, Ancliff P, Estrada J, Metzler M, Ponstingl E, Rütschle H, et al. MECOM-associated syndrome: a heterogeneous inherited bone marrow failure syndrome with amegakaryocytic thrombocytopenia. Blood Adv. 2018 Mar 27;2(6):586-596.

Germeshausen M, Ballmaier M. CAMT-MPL: congenital amegakaryocytic thrombocytopenia caused by MPL mutations - heterogeneity of a monogenic disorder - a comprehensive analysis of 56 patients. Haematologica. 2021 Sep 1;106(9):2439-2448.

Ghi T, Simonazzi G, Perolo A, Savelli L, Sandri F, Bernardi B, et al. Outcome of antenatally diagnosed intracranial hemorrhage: case series and review of the literature. Ultrasound Obstet Gynecol. 2003 Aug;22(2):121-30.

Giansily-Blaizot M, Rallapalli PM, Perkins SJ, Kemball-Cook G, Hampshire DJ, Gomez K, Ludlam CA, McVey JH. The EAHAD blood coagulation factor VII variant database. Hum Mutat. 2020 Jul;41(7):1209-1219.

Gilissen C, Hoischen A, Brunner HG, Veltman JA. Disease gene identification strategies for exome sequencing. Eur J Hum Genet. 2012 May;20(5):490-7.

Gillen J, Richardson D, Moore K. Angiopoietin-1 and Angiopoietin-2 Inhibitors: Clinical Development. Curr Oncol Rep. 2019 Feb 26;21(3):22.

Goerge T, Ho-Tin-Noe B, Carbo C, Benarafa C, Remold-O'Donnell E, Zhao BQ, Cifuni SM, Wagner DD. Inflammation induces hemorrhage in thrombocytopenia. Blood. 2008 May 15;111(10):4958-64.

Goldenberg NA, Manco-Johnson MJ. Protein C deficiency. Haemophilia. 2008 Nov;14(6):1214-21.

Goldmann T, Zeller N, Raasch J, Kierdorf K, Frenzel K, Ketscher L, et al. USP18 lack in microglia causes destructive interferonopathy of the mouse brain. EMBO J. 2015 Jun 12;34(12):1612-29.

Gordeeva V, Sharova E, Babalyan K, Sultanov R, Govorun VM, Arapidi G. Benchmarking germline CNV calling tools from exome sequencing data. Sci Rep. 2021 Jul 13;11(1):14416.

Gould DB, Phalan FC, Breedveld GJ, van Mil SE, Smith RS, Schimenti JC, et al. Mutations in Col4a1 cause perinatal cerebral hemorrhage and porencephaly. Science. 2005 May 20;308(5725):1167-71.

Gould DB, Phalan FC, van Mil SE, Sundberg JP, Vahedi K, Massin P, Bousser MG, et al. Role of COL4A1 in small-vessel disease and hemorrhagic stroke. N Engl J Med. 2006 Apr 6;354(14):1489-96.

Guey S, Hervé D. Main features of COL4A1-COL4A2 related cerebral microangiopathies. Cereb Circ Cogn Behav. 2022 Mar 24;3:100140.

Guo MH, Plummer L, Chan YM, Hirschhorn JN, Lippincott MF. Burden Testing of Rare Variants Identified through Exome Sequencing via Publicly Available Control Data. Am J Hum Genet. 2018 Oct 4;103(4):522-534.

Gupta V, Schlatterer SD, Bulas DI, du Plessis AJ, Mulkey SB. Pregnancy and Child Outcomes Following Fetal Intracranial Hemorrhage. Pediatr Neurol. 2023 Mar;140:68-75.

Hashemi SS, Zare-Abdollahi D, Bakhshandeh MK, Vafaee A, Abolhasani S, Inanloo Rahatloo K, et al. Clinical spectrum in multiple families with primary COQ10 deficiency. Am J Med Genet A. 2021 Feb;185(2):440-452.

Hausman-Kedem M, Ben-Sira L, Kidron D, Ben-Shachar S, Straussberg R, Marom D, et al. Deletion in COL4A2 is associated with a three-generation variable phenotype: from fetal to adult manifestations. Eur J Hum Genet. 2021 Nov;29(11):1654-1662.

Hausman-Kedem M, Malinger G, Modai S, Kushner SA, Shiran SI, Ben-Sira L, et al. Monogenic Causes of Apparently Idiopathic Perinatal Intracranial Hemorrhage. Ann Neurol. 2021 Apr;89(4):813-822.

He L, Vanlandewijck M, Mäe MA, Andrae J, Ando K, Del Gaudio F, et al. Single-cell RNA sequencing of mouse brain and lung vascular and vessel-associated cell types. Sci Data. 2018 Aug 21;5:180160.

Heremans J, Freson K. High-throughput sequencing for diagnosing platelet disorders: lessons learned from exploring the causes of bleeding disorders. Int J Lab Hematol. 2018 May;40 Suppl 1:89-96.

Honke N, Shaabani N, Zhang DE, Hardt C, Lang KS. Multiple functions of USP18. Cell Death Dis. 2016 Nov 3;7(11):e2444.

Høy Hansen M, Steensboe Lang C, Abildgaard N, Nyvold CG. Comparative evaluation of the heterozygous variant standard deviation as a quality measure for next-generation sequencing. J Biomed Inform. 2022 Nov;135:104234.

Hußmann M, Schulte D, Weischer S, Carlantoni C, Nakajima H, Mochizuki N, et al. Svep1 is a binding ligand of Tie1 and affects specific aspects of facial lymphatic development in a Vegfc-independent manner. Elife. 2023 Apr 25;12:e82969.

Ilves N, Pajusalu S, Kahre T, Laugesaar R, Šamarina U, Loorits D, Kool P, Ilves P. High Prevalence of Collagenopathies in Preterm- and Term-Born Children With Periventricular Venous Hemorrhagic Infarction. J Child Neurol. 2023 Jul 10:8830738231186233.

Imbard A, Boutron A, Vequaud C, Zater M, de Lonlay P, de Baulny HO, et al. Molecular characterization of 82 patients with pyruvate dehydrogenase complex deficiency. Structural implications of novel amino acid substitutions in E1 protein. Mol Genet Metab. 2011 Dec;104(4):507-16.

Ishida T, Kundu RK, Yang E, Hirata K, Ho YD, Quertermous T. Targeted disruption of endothelial cell-selective adhesion molecule inhibits angiogenic processes in vitro and in vivo. J Biol Chem. 2003 Sep 5;278(36):34598-604.

Itai T, Miyatake S, Taguri M, Nozaki F, Ohta M, Osaka H, et al. Prenatal clinical manifestations in individuals with COL4A1/2 variants. J Med Genet. 2021 Aug;58(8):505-513.

Jakobs BS, van den Heuvel LP, Smeets RJ, de Vries MC, Hien S, Schaible T, Smeitink JA, et al. A novel mutation in COQ2 leading to fatal infantile multisystem disease. J Neurol Sci. 2013 Mar 15;326(1-2):24-8.

Jeanne M, Jorgensen J, Gould DB. Molecular and Genetic Analyses of Collagen Type IV Mutant Mouse Models of Spontaneous Intracerebral Hemorrhage Identify Mechanisms for Stroke Prevention. Circulation. 2015 May 5;131(18):1555-65.

Jeanne M, Labelle-Dumais C, Jorgensen J, Kauffman WB, Mancini GM, Favor J, et al. COL4A2 mutations impair COL4A1 and COL4A2 secretion and cause hemorrhagic stroke. Am J Hum Genet. 2012 Jan 13;90(1):91-101.

Jeansson M, Gawlik A, Anderson G, Li C, Kerjaschki D, Henkelman M, Quaggin SE. Angiopoietin-1 is essential in mouse vasculature during development and in response to injury. J Clin Invest. 2011 Jun;121(6):2278-89.

Jelin AC, Norton ME, Bartha AI, Fick AL, Glenn OA. Intracranial magnetic resonance imaging findings in the surviving fetus after spontaneous monochorionic cotwin demise. Am J Obstet Gynecol. 2008 Oct;199(4):398.e1-5.

Jiao X, He P, Li Y, Fan Z, Si M, Xie Q, Chang X, Huang D. The Role of Circulating Tight Junction Proteins in Evaluating Blood Brain Barrier Disruption following Intracranial Hemorrhage. Dis Markers. 2015;2015:860120.

Jones FE, Murray LS, McNeilly S, Dean A, Aman A, Lu Y, et al. 4-Sodium phenyl butyric acid has both efficacy and counter-indicative effects in the treatment of Col4a1 disease. Hum Mol Genet. 2019 Feb 15;28(4):628-638.

Karpanen T, Padberg Y, van de Pavert SA, Dierkes C, Morooka N, Peterson-Maduro J, et al. An Evolutionarily Conserved Role for Polydom/Svep1 During Lymphatic Vessel Formation. Circ Res. 2017 Apr 14;120(8):1263-1275.

Kaur J, Reinhardt DP. Extracellular Matrix (ECM) Molecules. In: Stem Cell Biology and Tissue Engineering in Dental Sciences . Elsevier; 2015. p. 25-45.

Keep RF, Andjelkovic AV, Xiang J, Stamatovic SM, Antonetti DA, Hua Y, Xi G. Brain endothelial cell junctions after cerebral hemorrhage: Changes, mechanisms and therapeutic targets. J Cereb Blood Flow Metab. 2018 Aug;38(8):1255-1275.

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

Korhonen EA, Lampinen A, Giri H, Anisimov A, Kim M, Allen B, et al. Tie1 controls angiopoietin function in vascular remodeling and inflammation. J Clin Invest. 2016 Sep 1;126(9):3495-510.

Kujovich JL. Factor V Leiden thrombophilia. Genet Med. 2011 Jan;13(1):1-16.

Kuo DS, Labelle-Dumais C, Mao M, Jeanne M, Kauffman WB, Allen J, Favor J, Gould DB. Allelic heterogeneity contributes to variability in ocular dysgenesis, myopathy and brain malformations caused by Col4a1 and Col4a2 mutations. Hum Mol Genet. 2014 Apr 1;23(7):1709-22.

Landau D, Rosenberg N, Zivelin A, Staretz-Chacham O, Kapelushnik J. Familial factor VII deficiency with foetal and neonatal fatal cerebral haemorrhage associated with homozygosis to Gly180Arg mutation. Haemophilia. 2009 May;15(3):774-8.

Lazzarotto T, Blázquez-Gamero D, Delforge ML, Foulon I, Luck S, Modrow S, Leruez-Ville M. Congenital Cytomegalovirus Infection: A Narrative Review of the Issues in Screening and Management From a Panel of European Experts. Front Pediatr. 2020 Jan 31;8:13.

Lee S, Kim HM, Kang J, Seong WJ, Kim MJ. Fetal intracranial hemorrhage and maternal vitamin K deficiency induced by total parenteral nutrition: A case report. Medicine (Baltimore). 2022 Jan 7;101(1):e28434.

Leutenegger AL, Prum B, Génin E, Verny C, Lemainque A, Clerget-Darpoux F, Thompson EA. Estimation of the inbreeding coefficient through use of genomic data. Am J Hum Genet. 2003 Sep;73(3):516-23.

Lichtenbelt KD, Pistorius LR, De Tollenaer SM, Mancini GM, De Vries LS. Prenatal genetic confirmation of a COL4A1 mutation presenting with sonographic fetal intracranial hemorrhage. Ultrasound Obstet Gynecol. 2012 Jun;39(6):726-7.

Lieberman L, Greinacher A, Murphy MF, Bakchoul T, Corke S, Tanael S, et al. Fetal-Neonatal Alloimmune Thrombocytopenia (FNAIT): Guidance to Reduce the Risk of Intracranial Bleeding. Blood. 29 nov 2018;132(Supplement 1):4717-4717.

Lissens W, De Meirleir L, Seneca S, Liebaers I, Brown GK, Brown RM, et al. Mutations in the X-linked pyruvate dehydrogenase (E1) alpha subunit gene (PDHA1) in patients with a pyruvate dehydrogenase complex deficiency. Hum Mutat. 2000;15(3):209-19.

Lou C, Jiang J, Chen W, Zhang Z, Xu G, Liu Y, et al. Structural and functional characterization of novel F7 mutations identified in Chinese factor VII-deficient patients. Br J Haematol. 2023 Mar 23.

Luo J, Luo Y, Zeng H, Reis C, Chen S. Research Advances of Germinal Matrix Hemorrhage: An Update Review. Cell Mol Neurobiol. 2019 Jan;39(1):1-10.

Majethia P, Mishra S, Rao LP, Rao R, Shukla A. NAD(P)HX dehydratase (NAXD) deficiency due to a novel biallelic missense variant and review of literature. Eur J Med Genet. 2021 Sep;64(9):104266.

Mao M, Alavi MV, Labelle-Dumais C, Gould DB. Type IV Collagens and Basement Membrane Diseases. In: Current Topics in Membranes. Elsevier; 2015. p. 61-116.

Marlar RA, Neumann A. Neonatal purpura fulminans due to homozygous protein C or protein S deficiencies. Semin Thromb Hemost. 1990 Oct;16(4):299-309.

Martin G, Thomas MA, Wei XC, Le D. Diffuse Intracerebral Hemorrhage in an Infant With a Novel Homozygous Variant Leading to Severe Protein C Deficiency. J Pediatr Hematol Oncol. 2021 Aug 1;43(6):e763-e765.

McVey JH, Boswell EJ, Takamiya O, Tamagnini G, Valente V, Fidalgo T, Layton M, Tuddenham EG. Exclusion of the first EGF domain of factor VII by a splice site mutation causes lethal factor VII deficiency. Blood. 1998 Aug 1;92(3):920-6.

Meroni A, Kalantari S, Arossa A, Spinillo A, Melito C, Scatigno AL, et al. De novo RANBP2 variant in a fetal demise case with cerebral intraparenchymal hemorrhage. Am J Med Genet A. 2023 Jul;191(7):1973-1977.

Meuwissen ME, de Vries LS, Verbeek HA, Lequin MH, Govaert PP, Schot R, et al. Sporadic COL4A1 mutations with extensive prenatal porencephaly resembling hydranencephaly. Neurology. 2011 Mar 1;76(9):844-6.

Meuwissen ME, Halley DJ, Smit LS, Lequin MH, Cobben JM, de Coo R, et al. The expanding phenotype of COL4A1 and COL4A2 mutations: clinical data on 13 newly identified families and a review of the literature. Genet Med. 2015 Nov;17(11):843-53.

Meuwissen ME, Schot R, Buta S, Oudesluijs G, Tinschert S, Speer SD, et al. Human USP18 deficiency underlies type 1 interferonopathy leading to severe pseudo-TORCH syndrome. J Exp Med. 2016 Jun 27;213(7):1163-74.

Michelini S, Ricci M, Veselenyiova D, Kenanoglu S, Kurti D, Baglivo M, et al. TIE1 as a Candidate Gene for Lymphatic Malformations with or without Lymphedema. Int J Mol Sci. 2020 Sep 16;21(18):6780.

Misk RA, Qawasme L, Abunejma FM, Abu Rahma BI, Abuawwad EM, Abu Iram RI, et al. A Case Report and Literature Review of Pseudo-TORCH Syndrome Type 2 (PTORCH2). Case Rep Pediatr. 2022 Oct 22;2022:3555532.

Mitani T, Isikay S, Gezdirici A, Gulec EY, Punetha J, Fatih JM, et al. High prevalence of multilocus pathogenic variation in neurodevelopmental disorders in the Turkish population. Am J Hum Genet. 2021 Oct 7;108(10):1981-2005.

Miyatake S, Schneeberger S, Koyama N, Yokochi K, Ohmura K, Shiina M, et al. Biallelic COLGALT1 variants are associated with cerebral small vessel disease. Ann Neurol. 2018 Dec;84(6):843-853.

Mochida GH, Ganesh VS, Felie JM, Gleason D, Hill RS, Clapham KR, et al. A homozygous mutation in the tight-junction protein JAM3 causes hemorrhagic destruction of the brain, subependymal calcification, and congenital cataracts. Am J Hum Genet. 2010 Dec 10;87(6):882-9.

Monteagudo A. Intracranial Hemorrhage. Am J Obstet Gynecol. 2020 Dec;223(6):B34-B37.

Münch J, Prasuhn J, Laugwitz L, Fung CW, Chung BH, Bellusci M, et al. Neuroimaging in Primary Coenzyme-Q10-Deficiency Disorders. Antioxidants (Basel). 2023 Mar 14;12(3):718.

Murray LS, Lu Y, Taggart A, Van Regemorter N, Vilain C, Abramowicz M, Kadler KE, Van Agtmael T. Chemical chaperone treatment reduces intracellular accumulation of mutant collagen IV and ameliorates the cellular phenotype of a COL4A2 mutation that causes haemorrhagic stroke. Hum Mol Genet. 2014 Jan 15;23(2):283-92.

Natarajan N, Tully HM, Chapman T. Prenatal presentation of pyruvate dehydrogenase complex deficiency. Pediatr Radiol. 2016 Aug;46(9):1354-7.

Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 2003 Jul 1;31(13):3812-4.

Noris P, Pecci A. Hereditary thrombocytopenias: a growing list of disorders. Hematology Am Soc Hematol Educ Program. 2017 Dec 8;2017(1):385-399.

Nurden AT, Fiore M, Nurden P, Pillois X. Glanzmann thrombasthenia: a review of ITGA2B and ITGB3 defects with emphasis on variants, phenotypic variability, and mouse models. Blood. 2011 Dec 1;118(23):5996-6005.

Nurden AT, Nurden P. High-throughput sequencing for rapid diagnosis of inherited platelet disorders: a case for a European consensus. Haematologica. 2018 Jan;103(1):6-8.

Nurden AT, Nurden P. Inherited thrombocytopenias: history, advances and perspectives. Haematologica. 2020 Aug;105(8):2004-2019.

O'Driscoll MC, Daly SB, Urquhart JE, Black GC, Pilz DT, Brockmann K, et al. Recessive mutations in the gene encoding the tight junction protein occludin cause band-like calcification with simplified gyration and polymicrogyria. Am J Hum Genet. 2010 Sep 10;87(3):354-64.

Ozduman K, Pober BR, Barnes P, Copel JA, Ogle EA, Duncan CC, Ment LR. Fetal stroke. Pediatr Neurol. 2004 Mar;30(3):151-62.

Pedersen BS, Quinlan AR. Who's Who? Detecting and Resolving Sample Anomalies in Human DNA Sequencing Studies with Peddy. Am J Hum Genet. 2017 Mar 2;100(3):406-413.

Petermann R, Bakchoul T, Curtis BR, Mullier F, Miyata S, Arnold DM; Subcommittee on Platelet Immunology. Investigations for fetal and neonatal alloimmune thrombocytopenia: communication from the SSC of the ISTH. J Thromb Haemost. 2018 Dec;16(12):2526-2529.

Pirot N, Crahes M, Adle-Biassette H, Soares A, Bucourt M, Boutron A, et al. Phenotypic and Neuropathological Characterization of Fetal Pyruvate Dehydrogenase Deficiency. J Neuropathol Exp Neurol. 2016 Mar;75(3):227-38.

Plaisier E, Gribouval O, Alamowitch S, Mougenot B, Prost C, Verpont MC, et al. COL4A1 mutations and hereditary angiopathy, nephropathy, aneurysms, and muscle cramps. N Engl J Med. 2007 Dec 27;357(26):2687-95.

Pöschl E, Schlötzer-Schrehardt U, Brachvogel B, Saito K, Ninomiya Y, Mayer U. Collagen IV is essential for basement membrane stability but dispensable for initiation of its assembly during early development. Development. 2004 Apr;131(7):1619-28.

Putbrese B, Kennedy A. Findings and differential diagnosis of fetal intracranial haemorrhage and fetal ischaemic brain injury: what is the role of fetal MRI? Br J Radiol. 2017 Feb;90(1070):20160253.

Quenez O, Cassinari K, Coutant S, Lecoquierre F, Le Guennec K, Rousseau S, et al. Detection of copy-number variations from NGS data using read depth information: a diagnostic performance evaluation. Eur J Hum Genet. 2021 Jan;29(1):99-109.

Ramenghi LA, Fumagalli M, Righini A, Triulzi F, Kustermann A, Mosca F. Thrombophilia and fetal germinal matrix-intraventricular hemorrhage: does it matter? Ultrasound Obstet Gynecol. 2005 Oct;26(5):574-6.

Ratelade J, Klug NR, Lombardi D, Angelim MKSC, Dabertrand F, Domenga-Denier V, et al. Reducing Hypermuscularization of the Transitional Segment Between Arterioles and Capillaries Protects Against Spontaneous Intracerebral Hemorrhage. Circulation. 2020 Jun 23;141(25):2078-2094.

Richards A, van den Maagdenberg AM, Jen JC, Kavanagh D, Bertram P, Spitzer D, et al. C-terminal truncations in human 3'-5' DNA exonuclease TREX1 cause autosomal dominant retinal vasculopathy with cerebral leukodystrophy. Nat Genet. 2007 Sep;39(9):1068-70.

Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American

College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015 May;17(5):405-24.

Ritchie KJ, Malakhov MP, Hetherington CJ, Zhou L, Little MT, Malakhova OA, et al. Dysregulation of protein modification by ISG15 results in brain cell injury. Genes Dev. 2002 Sep 1;16(17):2207-12.

Rodríguez MÁ de la R, Hernández-Suárez M, Padilla-Pérez AI, Dévora-Cabrera Y, Acevedo WP. Fetal intracranial hemorrhage in a case of 16p microdeletion. Case Reports in Perinatal Medicine. 1 janv 2022;11(1).

Rosenfeld JA, Coe BP, Eichler EE, Cuckle H, Shaffer LG. Estimates of penetrance for recurrent pathogenic copy-number variations. Genet Med. 2013 Jun;15(6):478-81.

Rouaud T, Labauge P, Tournier Lasserve E, Mine M, Coustans M, Deburghgraeve V, Edan G. Acute urinary retention due to a novel collagen COL4A1 mutation. Neurology. 2010 Aug 24;75(8):747-9.

Salo AM, Cox H, Farndon P, Moss C, Grindulis H, Risteli M, Robins SP, Myllylä R. A connective tissue disorder caused by mutations of the lysyl hydroxylase 3 gene. Am J Hum Genet. 2008 Oct;83(4):495-503.

Sanapo L, Whitehead MT, Bulas DI, Ahmadzia HK, Pesacreta L, Chang T, du Plessis A. Fetal intracranial hemorrhage: role of fetal MRI. Prenat Diagn. 2017 Aug;37(8):827-836.

Sarigecili E, Ucar HK, Havali C, Cansu A, Aydin K. Acute necrotizing encephalopathy associated with RANBP2 mutation: value of MRI findings for diagnosis and intervention. Acta Neurol Belg. 2023 Apr;123(2):571-582.

Sato TN, Tozawa Y, Deutsch U, Wolburg-Buchholz K, Fujiwara Y, Gendron-Maguire M, et al. Distinct roles of the receptor tyrosine kinases Tie-1 and Tie-2 in blood vessel formation. Nature. 1995 Jul 6;376(6535):70-4.

Schrenk DA, Nasrallah HA. Faulty fences: Blood-brain barrier dysfunction in schizophrenia. Current Psychiatry. 2022 October;21(10):28-32.

Schwabenland M, Mossad O, Peres AG, Kessler F, Maron FJM, Harsan LA, et al. Loss of USP18 in microglia induces white matter pathology. Acta Neuropathol Commun. 2019 Jul 4;7(1):106.

Sehrish I, Sunitha T, Srilekha A, Gupta A, Nallari P, Venkateshwari A. A Novel Familial Case Report of Genetic Syndrome Mimicking Congenital TORCH infections; Pseudo-TORCH Syndrome 2. J Reprod Infertil. 2022 Apr-Jun;23(2):135-138.

Seo A, Gulsuner S, Pierce S, Ben-Harosh M, Shalev H, Walsh T, et al. Inherited thrombocytopenia associated with mutation of UDP-galactose-4-epimerase (GALE). Hum Mol Genet. 2019 Jan 1;28(1):133-142.

Serre JL, Leutenegger AL, Bernheim A, Fellous M, Rouen A, Kunstmann JM, Hyon C, Siffroi JP. Le don anonyme de sperme n'augmente pas significativement les unions entre apparentés,

la consanguinité et l'incidence des maladies récessives. Gynecol Obstet Fertil. 2014 Apr;42(4):200-4.

Shannon P, Hum C, Parks T, Schauer GM, Chitayat D, Chong K, et al. Brain and Placental Pathology in Fetal COL4A1 Related Disease. Pediatr Dev Pathol. 2021 May-Jun;24(3):175-186.

Shimojima K, Inoue T, Fujii Y, Ohno K, Yamamoto T. A familial 593-kb microdeletion of 16p11.2 associated with mental retardation and hemivertebrae. Eur J Med Genet. 2009 Nov-Dec;52(6):433-5.

Shinawi M, Liu P, Kang SH, Shen J, Belmont JW, Scott DA, et al. Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. J Med Genet. 2010 May;47(5):332-41.

Sibon I, Coupry I, Menegon P, Bouchet JP, Gorry P, Burgelin I, et al. COL4A1 mutation in Axenfeld-Rieger anomaly with leukoencephalopathy and stroke. Ann Neurol. 2007 Aug;62(2):177-84.

Sileo FG, Zöllner J, D'Antonio F, Islam S, Papageorghiou AT, Khalil A. Perinatal and long-term outcome of fetal intracranial hemorrhage: systematic review and meta-analysis. Ultrasound Obstet Gynecol. 2022 May;59(5):585-595.

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

Sofou K, Dahlin M, Hallböök T, Lindefeldt M, Viggedal G, Darin N. Ketogenic diet in pyruvate dehydrogenase complex deficiency: short- and long-term outcomes. J Inherit Metab Dis. 2017 Mar;40(2):237-245.

Song U, Ryu YH, Hong K, Shim SY, Park S, Lee JS, Ju YS, Shin SH, Lee S. Severe protein C deficiency in a newborn caused by a homozygous pathogenic variant in the PROC gene: a case report. BMC Pediatr. 2021 Oct 16;21(1):453.

Souma T, Thomson BR, Heinen S, Carota IA, Yamaguchi S, Onay T, et al. Context-dependent functions of angiopoietin 2 are determined by the endothelial phosphatase VEPTP. Proc Natl Acad Sci U S A. 2018 Feb 6;115(6):1298-1303.

Stutterd C, Savoia H, Fink AM, Stark Z. Severe fetal ischaemic brain injury caused by homozygous protein C deficiency. Prenat Diagn. 2014 Feb;34(2):192-4.

Sundaramoorthy M, Meiyappan M, Todd P, Hudson BG. Crystal structure of NC1 domains. Structural basis for type IV collagen assembly in basement membranes. J Biol Chem. 2002 Aug 23;277(34):31142-53.

Szpecht D, Szymankiewicz M, Nowak I, Gadzinowski J. Intraventricular hemorrhage in neonates born before 32 weeks of gestation-retrospective analysis of risk factors. Childs Nerv Syst. 2016 Aug;32(8):1399-404.

Tabibian S, Shams M, Naderi M, Dorgalaleh A. Prenatal diagnosis in rare bleeding disorders-An unresolved issue? Int J Lab Hematol. 2018 Jun;40(3):241-250.

Takenouchi T, Ohyagi M, Torii C, Kosaki R, Takahashi T, Kosaki K. Porencephaly in a fetus and HANAC in her father: variable expression of COL4A1 mutation. Am J Med Genet A. 2015 Jan;167A(1):156-8.

Talbert D. Cerebral Venous Malformation as a Cause of Neonatal Intra-Ventricular Haemorrhage and Unexplained Infant Subdural Haemorrhage. Anat Physiol. 2016;06(02).

Teunissen MWA, Kamsteeg EJ, Sallevelt SCEH, Pennings M, Bauer NJC, Vermeulen RJ, Nicolai J. Biallelic Variants in the COLGALT1 Gene Causes Severe Congenital Porencephaly: A Case Report. Neurol Genet. 2021 Mar 9;7(2):e564.

Thibault M, Leydet J, Tournier-Lasserve E, Crow YJ, Rivier F, Echenne B, et al. Syndromes génétiques mimant les infections congénitales : à propos de 2 cas. Arch Pediatr. 2011 Dec;18(12):1297-1301.

Toughza J, Agadr A, Nejjari M, Ammari IA. Diagnostic et prise en charge d´une thrombopénie néonatale sévère par allo-immunisation materno-fœtale: rapport d´un cas et revue de la littérature. Pan Afr Med J. 2020 Dec 29;37:382.

Trillot N, Zawadzki C, Watel A, Jude B. La transition G20210A dans le gène de la prothrombine et la maladie thromboembolique veineuse. La Revue de Médecine Interne. 1 oct 2000;21(10):911-4.

Van Bergen NJ, Guo Y, Rankin J, Paczia N, Becker-Kettern J, Kremer LS, et al. NAD(P)HX dehydratase (NAXD) deficiency: a novel neurodegenerative disorder exacerbated by febrile illnesses. Brain. 2019 Jan 1;142(1):50-58.

van der Knaap MS, Smit LM, Barkhof F, Pijnenburg YA, Zweegman S, Niessen HW, Imhof S, Heutink P. Neonatal porencephaly and adult stroke related to mutations in collagen IV A1. Ann Neurol. 2006 Mar;59(3):504-11.

Vahedi K, Boukobza M, Massin P, Gould DB, Tournier-Lasserve E, Bousser MG. Clinical and brain MRI follow-up study of a family with COL4A1 mutation. Neurology. 2007 Oct 16;69(16):1564-8.

Vanlandewijck M, He L, Mäe MA, Andrae J, Ando K, Del Gaudio F, et al A molecular atlas of cell types and zonation in the brain vasculature. Nature. 2018 Feb 22;554(7693):475-480.

Verdu A, Cazorla MR, Moreno JC, Casado LF. Prenatal stroke in a neonate heterozygous for factor V Leiden mutation. Brain Dev. 2005 Sep;27(6):451-4.

Vergani P, Strobelt N, Locatelli A, Paterlini G, Tagliabue P, Parravicini E, Ghidini A. Clinical significance of fetal intracranial hemorrhage. Am J Obstet Gynecol. 1996 Sep;175(3 Pt 1):536-43.

Whitelaw A, Haines ME, Bolsover W, Harris E. Factor V deficiency and antenatal intraventricular haemorrhage. Arch Dis Child. 1984 Oct;59(10):997-9.

Winkelhorst D, Murphy MF, Greinacher A, Shehata N, Bakchoul T, Massey E, et al. Antenatal management in fetal and neonatal alloimmune thrombocytopenia: a systematic review. Blood. 2017 Mar 16;129(11):1538-1547.

Yang J, Mao H, Sun L. Congenital coagulation factor V deficiency with intracranial hemorrhage. J Clin Lab Anal. 2022 Nov;36(11):e24705.

Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. J Comput Biol. 2004;11(2-3):377-94.

Yoneda Y, Haginoya K, Arai H, Yamaoka S, Tsurusaki Y, Doi H, et al. De novo and inherited mutations in COL4A2, encoding the type IV collagen α2 chain cause porencephaly. Am J Hum Genet. 2012 Jan 13;90(1):86-90.

Yoneda Y, Haginoya K, Kato M, Osaka H, Yokochi K, Arai H, et al. Phenotypic spectrum of COL4A1 mutations: porencephaly to schizencephaly. Ann Neurol. 2013 Jan;73(1):48-57.

Yougbaré I, Lang S, Yang H, Chen P, Zhao X, Tai WS, et al. Maternal anti-platelet β3 integrins impair angiogenesis and cause intracranial hemorrhage. J Clin Invest. 2015 Apr;125(4):1545-56.

Zagaglia S, Selch C, Nisevic JR, Mei D, Michalak Z, Hernandez-Hernandez L, et al. Neurologic phenotypes associated with COL4A1/2 mutations: Expanding the spectrum of disease. Neurology. 2018 Nov 27;91(22):e2078-e2088.

Zech M, Kopajtich R, Steinbrücker K, Bris C, Gueguen N, Feichtinger RG, et al. Variants in Mitochondrial ATP Synthase Cause Variable Neurologic Phenotypes. Ann Neurol. 2022 Feb;91(2):225-237.

Zhao L, Liu H, Yuan X, Gao K, Duan J. Comparative study of whole exome sequencing-based copy number variation detection tools. BMC Bioinformatics. 2020 Mar 5;21(1):97.

Zuccarello D, Sorrentino U, Brasson V, Marin L, Piccolo C, Capalbo A, Andrisani A, Cassina M. Epigenetics of pregnancy: looking beyond the DNA code. J Assist Reprod Genet. 2022 Apr;39(4):801-816.

Annexes

Annexe 1. Versions des logiciels et des bases de données utilisées pour l'analyse bio-informatique des données de WES

Logiciels	
bcl2fastq	2.17
FastQC	0.11.4
BWA	0.7.15
GATK (SNVs)	3.8.1
DNAcopy	3.5.1
Circos	0.69.6
VEP	95.1
Picard	2.12.1

Bases de données	
Mirbase	22
Clinvar	20190219
Freq-IG	19-10-2018 (1 210 Exomes)
Kaviar	160204
Cosmic	87
PhyloP	100
PhastCons	100
DisGeNet	4.0
GnomAD	2.0.1
FATHMM	2.3
dbscSNV	1.1
dbNSFP	3.5
OMIM	19-02-2019
RegulomeDB	1.1
Ensembl database version	95
Genome assembly	GRCh38.p12
GENCODE	29
RefSeq	2018-07-10 (GCF_000001405.38_GRCh38.p12_genomic.gff)
Regulatory build	1.0
PolyPhen	2.2.2
SIFT	5.2.2
dbSNP	151
1000 Genomes	Phase 3 (remapped)

Annexe 2. Liste des 47 gènes candidats intervenant dans la biosynthèse de COL4A1 et COL4A2

Liste basée sur les données de la littérature et les données fournies par le Dr Roberto Vanacore dans le cadre d'un projet collaboratif financé par le NIH (subvention R01NS096173).

RPS3	*DNAJB11*	*HYOU1*	*PLOD1*	*PDIA4*
RPSA	*ERP29*	*MVP*	*PLOD2*	*PDIA6*
ANXA2	*ERP44*	*GANAB*	*PLOD3*	*TGM2*
ASPH	*HSP90B1*	*PPIB*	*P3H1*	*RCN1*
ATP5F1A	*ERO1A*	*FKBP10*	*P3H3*	*SPARC*
HSPA5	*FN1*	*FKBP9*	*P4HA1*	*TXNDC5*
CALR	*PRKCSH*	*PRDX1*	*P4HA2*	*UGGT1*
CALU	*GOLIM4*	*PRDX4*	*CNPY2*	
CRTAP	*HSP90AB1*	*PARP4*	*P4HB*	
CKAP4	*SERPINH1*	*COLGALT1*	*PDIA3*	

Annexe 3. Liste des CNV (délétions) homozygotes identifiés par le logiciel HMZDelFinder chez les fœtus consanguins

Fœtus	Chromosome	Début	Fin	Taille (en kilobases)	Gènes inclus dans la délétion homozygote
F01	20	38213041	38213743	0,702	*KIAA1755*
F58	14	32821401	33560210	738,8	*NPAS3*
F58	1	151762375	151763167	0,792	*MRPL9*
F81	1	207144543	207466912	322,4	*CD55*
F81	8	69638734	69679619	40,9	*SULF1, SLCO5A1*
F81	1	145385193	145390116	4,92	*NBPF20*
F81	17	75103969	75104180	0,211	*SLC16A5*
F81	7	95324401	95324475	0,074	*PON1*

Annexe 4. Synthèse des données pour les 113 fœtus index

N° fœtus	Sexe	WES Trio Famille	Forme familiale	Consanguinité interrogatoire	Consanguinité Fantasio	Origine ethnique	Gène Causal	Position HGVSc / Position protéine	Mode de transmission	Gène Candidat	Position HGVSc / Position protéine	Mode de transmission
F01	F	Oui	Non	Oui	Oui	Africaine	-	-	-	*NAXD*	NM_018210.3:c.748G>A / p.Asp250Asn	Autosomique récessif
F02	M	Oui	Non	Non	Non	Européenne	-	-	-	-	-	-
F03	M	Oui	Oui	Non	Non	Européenne	***MPL***	NM_005373.2:c.305G>C / p.Arg102Pro & NM_005373.2:c.1609C>T / p.Arg537Trp	Autosomique récessif	-	-	-
F04	F	Non	Non	ND	Non	Européenne	-	-	-	-	-	-
F05	F	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F06	F	Oui	Non	ND	Non	Européenne	-	-	-	*ITGA2B*	NM_000419.4:c.2297T>C / p.Ile766Thr	Autosomique dominant (de novo)
F07	F	Non	Oui	Non	Non	Multiple	-	-	-	-	-	-
F08	M	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F09	F	Non	Non	Non	Non	Européenne	-	-	-	*PLOD3*	NM_001084.4:c.593T>C / p.Leu198Pro	Autosomique récessif
F10	F	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F11	M	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F12	F	Oui	Non	ND	Non	Multiple	***PROC***	NM_000312.3:c.644_651del / p.Gly215Aspfs*40 & NM_000312.3:c.1163C>A / p.Ala388Glu	Autosomique récessif	*PLOD3*	NM_001084.4:c.589C>T / p.Arg197Trp & NM_001084.4:c.1466C>T / p.Pro489Leu	Autosomique récessif
F13	F	Non	Non	ND	Non	Admixed American	-	-	-	-	-	-
F14	M	Non	Non	ND	Non	Africaine	-	-	-	-	-	-
F15	M	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F16	M	Oui	Oui	Non	Non	Africaine	-	-	-	-	-	-
F17	M	Oui	Non	ND	Non	Européenne	-	-	-	-	-	-
F18	M	Non	Non	ND	Non	Européenne	-	-	-	-	-	-
F19	F	Non	Oui	Non	Non	Européenne	***MECOM***	NM_001105077.3:c.1268del / p.Ser423Leufs*6	Autosomique dominant	-	-	-

F20	F	Non	Non	ND	Non	Européenne	-	-	-	-	-	-
F21	F	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F22	F	Non	Non	Non	Non	Africaine	-	-	-	-	-	-
F23	M	Oui	Oui	Non	Non	Européenne	-	-	-	-	-	-
F24	M	Non	Non	ND	Non	Européenne	-	-	-	-	-	-
F25	F	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F26	M	Oui	Non	Non	Non	Multiple	-	-	-	-	-	-
F27	F	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F28	M	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F29	F	Non	Non	ND	Non	Européenne	-	-	-	-	-	-
F30	M	Oui	Oui	Non	Non	Européenne	-	-	-	-	-	-
F31	F	Non	Non	Non	Non	Admixed American	-	-	-	-	-	-
F32	F	Non	Oui	Non	Non	Européenne	***COQ2***	NM_015697.8:c.1197del / p.Asn401Ilefs*15	Autosomique récessif	-	-	-
F33	M	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F34	F	Oui	Non	Non	Non	Africaine	-	-	-	*DCAF5*	NM_003861.2:c.1433del / p.Asp478Alafs*17	Autosomique dominant (de novo)
F35	M	Oui	Non	Non	Non	Européenne	-	-	-	*RNPEP*	NM_020216.3:c.805G>T / p.Glu269*	Autosomique récessif
F36	M	Oui	Non	ND	Non	Africaine	-	-	-	-	-	-
F37	M	Non	Non	ND	Non	Multiple	-	-	-	-	-	-
F38	M	Non	Non	ND	Non	Européenne	-	-	-	*TIE1*	NM_005424.4:c.1382G>A / p.Arg461His	Autosomique dominant
F39	M	Non	Non	ND	Non	Européenne	-	-	-	-	-	-
F40	M	Non	Non	Non	Non	Multiple	-	-	-	-	-	-
F41	F	Non	Non	ND	Non	Européenne	-	-	-	-	-	-
F42	F	Non	Non	ND	Non	Européenne	-	-	-	-	-	-
F43	M	Non	Non	ND	Non	Européenne	-	-	-	-	-	-
F44	M	Non	Non	ND	Non	Multiple	-	-	-	-	-	-
F45	M	Non	Non	ND	Non	Européenne	-	-	-	-	-	-

F46	F	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F47	M	Non	Non	Non	Non	Européenne	-	-	-	*TIE1*	NM_005424.4:c.207C>G / p.Ile69Met	Autosomique dominant
F48	F	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F49	F	Non	Non	ND	Non	Européenne	***PDHA1***	NM_000284.3:c.355C>T / p.Arg119Trp	Dominant lié à l'X	-	-	-
F50	F	Non	Non	ND	Non	Multiple	-	-	-	-	-	-
F51	M	Non	Non	Non	Non	Admixed American	-	-	-	-	-	-
F52	M	Non	Non	ND	Non	Admixed American	-	-	-	-	-	-
F53	F	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F54	F	Non	Non	Non	Non	Européenne	***PDHA1***	NM_000284.3:c.705_706del / p.Arg235Serfs*6	Dominant lié à l'X	-	-	-
F55	F	Non	Non	Non	Non	Européenne	-	-	-	*TIE1*	NM_005424.4:c.1297G>A / p.Gly433Ser	Autosomique dominant
F56	M	Non	Non	Non	Non	Multiple	-	-	-	-	-	-
F57	F	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F58	F	Non	Non	Oui	Oui	Européenne	***ATP5PO***	NM_001697.2:c.87+3A>G	Autosomique récessif	-	-	-
F59	F	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F60	F	Non	Non	ND	Non	Européenne	-	-	-	-	-	-
F61	M	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F62	M	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F63	F	Non	Non	Non	Non	Africaine	-	-	-	-	-	-
F64	M	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F65	F	Non	Non	ND	Non	Admixed American	-	-	-	-	-	-
F66	M	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F67	M	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F68	F	Non	Non	ND	Non	Admixed American	-	-	-	-	-	-
F69	M	Non	Non	Non	Non	Africaine	-	-	-	-	-	-
F70	F	Non	Non	ND	Non	Africaine	-	-	-	-	-	-

F71	F	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F72	M	Non	Non	ND	Non	Européenne	-	-	-	-	-	-
F73	M	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F74	M	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F75	F	Non	Non	ND	Non	Européenne	-	-	-	-	-	-
F76	M	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F77	M	Non	Non	Non	Non	Africaine	-	-	-	*BCOR*	NM_001123385.1:c.4813G>C / p.Val1605Leu	Récessif lié à l'X
F78	M	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F79	M	Non	Non	Non	Non	Multiple	-	-	-	-	-	-
F80	M	Oui	Non	Non	Non	Européenne	-	-	-	-	-	-
F81	M	Oui	Non	Oui	Oui	Africaine	-	-	-	-	-	-
F82	M	Oui	Non	Non	Non	Européenne	-	-	-	-	-	-
F83	F	Non	Non	ND	Non	Européenne	-	-	-	-	-	-
F84	F	Non	Non	Non	Non	Africaine	-	-	-	-	-	-
F85	M	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F86	F	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F87	M	Oui	Non	Non	Non	Européenne	-	-	-	-	-	-
F88	M	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F89	F	Oui	Non	Non	Non	Multiple	-	-	-	-	-	-
F90	M	Non	Non	ND	Non	Européenne	-	-	-	-	-	-
F91	F	Non	Non	Non	Non	Européenne	-	-	-	-	-	-
F92	M	Oui	Oui	Oui	Oui	Européenne	***ESAM***	NM_138961.2:c.35T>A / p.Leu12*	Autosomique récessif	-	-	-
F93	M	Oui	Non	Non	Non	Européenne	-	-	-	-	-	-
F94	F	Oui	Non	Oui	Non	Européenne	-	-	-	-	-	-
F95	F	Non	Non	Non	Non	Européenne	***PDHA1***	NM_000284.3:c.910C>T / p.Arg304*	Dominant lié à l'X	-	-	-
F96	M	Oui	Non	Non	Non	Européenne	-	-	-	-	-	-
F97	M	Oui	Oui	Non	Non	Africaine	-	-	-	-	-	-
F98	M	Non	Non	Non	Non	Multiple	-	-	-	-	-	-

F99	M	Non	Non	ND	Non	Européenne	-	-	-	-	-	-
F100	M	Oui	Non	Non	Non	Européenne	-	-	-	-	-	-
F101	M	Oui	Non	Oui	Non	Asie du Sud	-	-	-	*TIE1*	NM_005424.4:c.3053G>A / p.Arg1018His	Autosomique dominant
F102	F	Oui	Non	Non	Non	Européenne	-	-	-	-	-	-
F103	F	Oui	Non	Non	Non	Européenne	-	-	-	*TIE1*	NM_005424.4:c.2462G>A / p.Arg821Gln	Autosomique dominant
F104	F	Oui	Non	Non	Non	Européenne	-	-	-	-	-	-
F105	F	Oui	Non	Non	Non	Européenne	-	-	-	-	-	-
F106	F	Oui	Non	Non	Non	Européenne	-	-	-	*ITGA2B*	NM_000419.4:c.1300C>T / p.Pro434Ser	Autosomique dominant (de novo)
F107	F	Oui	Non	Non	Non	Européenne	-	-	-	-	-	-
F108	F	Oui	Non	Non	Non	Africaine	-	-	-	*EXD3*	NM_017820.4:c.2015G>A / p.Gly672Glu	Autosomique dominant (de novo)
F109	F	Oui	Non	Non	Non	Européenne	-	-	-	*ZFYVE28*	NM_020972.2:c.1544C>T / p.Ala515Val	Autosomique dominant (de novo)
F110	F	Oui	Non	Non	Non	Européenne	-	-	-	-	-	-
F111	M	Oui	Non	ND	Non	Européenne	-	-	-	-	-	-
F112	M	Oui	Non	Non	Non	Européenne	-	-	-	*TIE1*	NM_005424.4:c.3052C>T / p.Arg1018Cys	Autosomique dominant
F113	F	Oui	Non	Non	Non	Européenne	-	-	-	-	-	-

ND : donnée non disponible

Annexe 5. Liste des gènes *knockout* dans le modèle souris associés à un phénotype d'hémorragie intracrânienne

Gène chez la souris (database MGI 6.22)	Nom du gène chez l'humain	Code MGI	Gène chez la souris (database MGI 6.22)	Nom du gène chez l'humain	Code MGI	Gène chez la souris (database MGI 6.22)	Nom du gène chez l'humain	Code MGI
Abca1	***ABCA1***	J:105902	Foxc1	***FOXC1***	J:5191, J:75733	Pnpla6	***PNPLA6***	J:87675
Adam10	***ADAM10***	J:159624	Foxc2	***FOXC2***	J:39636	Por	***POR***	J:74666
Adgra2	***ADGRA2***	J:171226	G3bp1	***G3BP1***	J:101358	Proc	***PROC***	J:50527
Aggf1	***AGGF1***	J:239252	Gfap	***GFAP***	J:124385	Pros1	***PROS1***	J:152423
Alkbh1	***ALKBH1***	J:166841	Gldc	***GLDC***	J:221782	Psen1	***PSEN1***	J:95663
Alpl	***ALPL***	J:39015	Glrx3	***GLRX3***	J:143577	Ptpn9	***PTPN9***	J:118880
Ankrd17	***ANKRD17***	J:152840	Grip1	***GRIP1***	J:79476	Rala	***RALA***	J:239583
Anxa7	***ANXA7***	J:58572	Hdac8	***HDAC8***	J:150709	Rara	***RARA***	J:21009
Apaf1	***APAF1***	J:41353	Hsf2	***HSF2***	J:83730	Rarg	***RARG***	J:21009
App	***APP***	J:241632	Hspg2	***HSPG2***	J:58700	Rhbdf1	***RHBDF1***	J:203951
Birc6	***BIRC6***	J:94169	Id1	***ID1***	J:86480	Runx1	***RUNX1***	J:118892
Braf	***BRAF***	J:41257	Id3	***ID3***	J:86480	Sec23a	***SEC23A***	J:237614
C1galt1	***C1GALT1***	J:96775	Itga7	***ITGA7***	J:100236	Serpina10	***SERPINA10***	J:136102
Casp9	***CASP9***	J:49087	Itgav	***ITGAV***	J:50951	Serpinc1	***SERPINC1***	J:65092
Cbfb	***CBFB***	J:36593	Itgb3	***ITGB3***	J:224675	Slc23a2	***SLC23A2***	J:76194
Ccdc134	***CCDC134***	J:256995	Itgb8	***ITGB8***	J:77682	Socs7	***SOCS7***	J:93449
Ccdc85c	***CCDC85C***	J:93608	Jag1	***JAG1***	J:115783	Sphk1	***SPHK1***	J:103755
Ccm2	***CCM2***	J:177584	Jam3	***JAM3***	J:191990	Sphk2	***SPHK2***	J:103755
Ccr2	***CCR2***	J:121703	Kansl2	***KANSL2***	J:291094	Ssr2	***SSR2***	J:239583
Cdc42	***CDC42***	J:250065	Kansl3	***KANSL3***	J:291094	Sufu	***SUFU***	J:101731
Cep41	***CEP41***	J:181179	Kdm2b	***KDM2B***	J:171291	Syk	***SYK***	J:181715
Cfap54	***CFAP54***	J:228444	Kdr	***KDR***	J:142352	Tek	***TEK***	J:99096
Cfh	***CFH***	J:240426	Kif27	***KIF27***	J:185566	Tfpi	***TFPI***	J:51137
Clec1b	***CLEC1B***	J:181715	Kif6	***KIF6***	J:269442	Tgfb1	***TGFB1***	J:136106

Cog6	***COG6***	J:239583	Lamc1	***LAMC1***	J:124247	Tgfb3	***TGFB3***	J:136106
Col18a1	***COL18A1***	J:92842	Lemd2	***LEMD2***	J:221495	Thbd	***THBD***	J:71269
Col1a1	***COL1A1***	J:129569	Llgl1	***LLGL1***	J:89028	Tmsb4x	***TMSB4Y***	J:195902
Col4a1	***COL4A1***	J:240798	Mapk7	***MAPK7***	J:107620	Tnfrsf21	***TNFRSF21***	J:181470
Col4A2	***COL4A2***	J:163225	Morc2a	***MORC2***	J:239583	Trp53bp2	***TP53BP2***	J:240594
Commd9	***COMMD9***	J:230747	Msi1	***MSI1***	J:81794	Trp73	***TP73***	J:60896
Crebbp	***CREBBP***	J:63763	Msx1	***MSX1***	J:173525	Tulp3	***TULP3***	J:70060
Ctnnb1	***CTNNB1***	J:142352	Msx2	***MSX2***	J:173525	Twsg1	***TWSG1***	J:82738
Cx3cr1	***CX3CR1***	J:129712	Ndufs4	***NDUFS4***	J:161393	Ube4b	***UBE4B***	J:103761
Cycs	***CYCS***	J:98949	Nme5	***NME5***	J:185566	Ulk4	***ULK4***	J:185566
Ddb1	***DDB1***	J:117820	Nr1h2	***NR1H2***	J:81790	Usp18	***USP18***	J:78821
E2f5	***E2F5***	J:47283	Nr1h3	***NR1H3***	J:81790	Vim	***VIM***	J:124385
Erg	***ERG***	J:188920	Odad3	***ODAD3***	J:278798	Vkorc1	***VKORC1***	J:207884
F11	***F11***	J:67398	Parl	***PARL***	J:269785	Wasf2	***WASF2***	J:84651
F2	***F2***	J:48274	Pcnt	***PCNT***	J:226309	Wnk1	***WNK1***	J:152906
F3	***F3***	J:41500	Pdgfb	***PDGFB***	J:89186	Wnt7a	***WNT7A***	J:142352
F5	***F5***	J:136102	Pdgfrb	***PDGFRB***	J:86542	Wnt7b	***WNT7B***	J:142352
Fbln1	***FBLN1***	J:137690	Plat	***PLAT***	J:89052	Ybx1	***YBX1***	J:99099
Fbxw7	***FBXW7***	J:88517	Plod3	***PLOD3***	J:106510	Zfp385a	***ZNF385A***	J:93829
Fermt3	***FERMT3***	J:133659	Plxnd1	***PLXND1***	J:143763	.	.	.
Fli1	***FLI1***	J:64178	Pnkp	***PNKP***	J:226703	.	.	.

Source : MGI (Mouse Genome Informatics) https://www.informatics.jax.org

Articles Annexes

- **Rôle du gène *PLOD3* dans la biosynthèse de COL4A1/COL4A2**

Shikawa Y, Taga Y, Coste T, Tufa SF, Keene DR, Mizuno K, Tournier-Lasserve E, Gould DB. Lysyl hydroxylase 3-mediated post-translational modifications are required for proper biosynthesis of collagen α1α1α2(IV). J Biol Chem. 2022 Dec;298(12):102713.

- **Lésions neuropathologiques observées chez les fœtus porteurs de variants pathogènes de *COL4A1***

Gubana F, Christov C, Coste T, Tournier-Lasserve E, Benachi A, Fallet-Bianco C, Encha-Razavi F, Martinovic J. Prenatal Diagnosis of *COL4A1* Mutations in Eight Cases: Further Delineation of the Neurohistopathological Phenotype. Pediatr Dev Pathol. 2022 Jul-Aug;25(4):435-446.

- **Prévalence des variants pathogènes de *COL4A1/COL4A2* chez les fœtus présentant des lésions cérébrales multifocales hémorragiques et/ou ischémiques sévères**

Maurice P, Guilbaud L, Garel J, Mine M, Dugas A, Friszer S, Maisonneuve E, Moutard ML, Coste T, Héron D, Tournier-Lasserve E, Garel C, Jouannic JM. Prevalence of *COL4A1* and *COL4A2* mutations in severe fetal multifocal hemorrhagic and/or ischemic cerebral lesions. Ultrasound Obstet Gynecol. 2021 May;57(5):783-789.

JBC RESEARCH ARTICLE

Lysyl hydroxylase 3–mediated post-translational modifications are required for proper biosynthesis of collagen α1α1α2(IV)

Received for publication, July 8, 2022, and in revised form, November 5, 2022 Published, Papers in Press, November 17, 2022,
https://doi.org/10.1016/j.jbc.2022.102713

Yoshihiro Ishikawa[1,*], Yuki Taga[2], Thibault Coste[3], Sara F. Tufa[4], Douglas R. Keene[4], Kazunori Mizuno[2], Elisabeth Tournier-Lasserve[3], and Douglas B. Gould[1,5,*]

From the [1]Department of Ophthalmology, University of California San Francisco, School of Medicine, California, USA; [2]Nippi Research Institute of Biomatrix, Ibaraki, Japan; [3]Université Paris Cité, Inserm Neurodiderot, AP-HP Paris, France; [4]Research Department, Shriners Hospital for Children, Portland, Oregon, USA; [5]Department Anatomy, Cardiovascular Research Institute, Bakar Aging Research Institute, and Institute for Human Genetics, University of California, San Francisco, California, USA

Edited by Ursula Jakob

Collagens are the most abundant proteins in the body and among the most biosynthetically complex. A molecular ensemble of over 20 endoplasmic reticulum resident proteins participates in collagen biosynthesis and contributes to heterogeneous post-translational modifications. Pathogenic variants in genes encoding collagens cause connective tissue disorders, including osteogenesis imperfecta, Ehlers–Danlos syndrome, and Gould syndrome (caused by mutations in *COL4A1* and *COL4A2*), and pathogenic variants in genes encoding proteins required for collagen biosynthesis can cause similar but overlapping clinical phenotypes. Notably, pathogenic variants in lysyl hydroxylase 3 (LH3) cause a multisystem connective tissue disorder that exhibits pathophysiological features of collagen-related disorders. LH3 is a multifunctional collagen-modifying enzyme; however, its precise role(s) and substrate specificity during collagen biosynthesis has not been defined. To address this critical gap in knowledge, we generated LH3 KO cells and performed detailed quantitative and molecular analyses of collagen substrates. We found that LH3 deficiency severely impaired secretion of collagen α1α1α2(IV) but not collagens α1α1α2(I) or α1α1α1(III). Amino acid analysis revealed that LH3 is a selective LH for collagen α1α1α2(IV) but a general glucosyltransferase for collagens α1α1α2(IV), α1α1α2(I), and α1α1α1(III). Importantly, we identified rare variants that are predicted to be pathogenic in the gene encoding LH3 in two of 113 fetuses with intracranial hemorrhage—a cardinal feature of Gould syndrome. Collectively, our findings highlight a critical role of LH3 in α1α1α2(IV) biosynthesis and suggest that LH3 pathogenic variants might contribute to Gould syndrome.

Collagens are the most abundant proteins in the human body and comprise 28 members encoded by 46 genes (1, 2). Pathogenic variants in genes encoding collagens cause a wide range of disorders affecting almost all tissues and organs (3, 4). Type I, type III, and type V collagens form fibrils contributing to structural frameworks and pathogenic variants in these collagens cause connective tissue disorders—osteogenesis imperfecta (OI) and Ehlers–Danlos syndrome (EDS)—that affect fibril-rich tissues, such as bone, skin, and vascular structures (5–10). In contrast, type IV collagens form flexible networks that are present in the basement membranes of all tissues. Pathogenic variants in type IV collagens cause Alport syndrome (*COL4A3*, *COL4A4*, and *COL4A5*), which primarily affects the glomerular basement membrane (11, 12) and Gould syndrome (*COL4A1* and *COL4A2*), a congenital multisystem disorder that includes highly variable cerebrovascular disease, pediatric epilepsy, ocular anterior segment dysgenesis, skeletal myopathy, and nephropathy (13–18).

Collagen biosynthesis is a complex and highly orchestrated process that occurs in the endoplasmic reticulum (ER) and is facilitated by a molecular ensemble (19) of over 20 proteins that participate in four major steps: post-translational modifications (PTMs) of individual collagen chains, trimer assembly, triple helix formation, and secretion (19–21). Although most components of the molecular ensemble are shared among collagens, the collagen-modifying enzymes differ between collagens (22–24), suggesting subtype specialization of the molecular ensemble (24, 25). Moreover, the molecular ensemble and PTMs for type I collagen, the most extensively characterized collagen, show that the PTM profile is also tissue specific (24, 26). PTMs on the collagen triple helix are classified into three types of primary modifications to prolyl or lysyl residues and two types of secondary modifications to hydroxylysines. Prolyl modifications include 4-hydroxyproline (4Hyp) and 3-hydroxyproline that are generated by prolyl 4 hydroxylases and prolyl 3 hydroxylases, respectively (27). Approximately 40% or more of the prolines are modified to 4Hyp (22, 24, 28), which provides thermal stability to the collagen triple helix (29–31). In contrast, less than 5% of the prolines are modified to 3-hydroxyproline (28, 32), which is proposed to mediate protein–protein interactions (32, 33). Primary modification of lysine to hydroxylysine (Hyl) is attributed to lysyl hydroxylases (LHs), (27) and secondary

* For correspondence: Yoshihiro Ishikawa, yoshihiro.ishikawa@uscf.edu; Douglas B. Gould, douglas.gould@ucsf.edu.

modifications of Hyl include two successive and coordinated *O*-glycosylation steps. Hyl is required for subsequent galactose attachment to make galactosyl Hyls, which are further modified by glucose attachment to make glucosyl galactosyl Hyls. This sequence of modifications from lysine, to Hyl, to galactosyl Hyl and glucosyl galactosyl Hyl occurs in the ER during collagen biosynthesis (34, 35).

In the case of OI and EDS, it is well documented that pathogenic variations in genes encoding collagens and genes encoding proteins required for collagen biosynthesis lead to similar clinical outcomes (5–10). A similar paradigm may also exist for Gould syndrome as pathogenic variants in glycosyltransferase 25 domain containing 1 (GLT25D1), encoded by collagen β (1-O) galactosyltransferase 1 (*COLGALT1*) cause musculoskeletal (36) and cerebrovascular (37, 38) manifestations similar to those observed in individuals with *COL4A1* and *COL4A2* pathogenic variants. Notably, because molecular ensemble proteins can be multifunctional and employed across diverse collagen types, pathogenic variants in encoding genes often cause phenotypes that blend across different clinical categories. An example being LH3, encoded by procollagen-lysine, 2-oxoglutarate 5-dioxygenase 3 (*PLOD3*), whereby pathogenic variants cause connective tissue phenotypes that overlap with Stickler-like syndrome and epidermolysis bullosa (39–42). Indeed, clinical features of patients with biallelic loss of function *PLOD3* variants include craniofacial dysmorphisms, skeletal and ocular disorders, sensorineural hearing loss, and variable skin manifestations. Interestingly, in addition to clinical manifestations overlapping Stickler-like syndrome and epidermolysis bullosa, cerebral hemorrhages have been reported in three cases with *PLOD3* variants (39, 43). LH3 is proposed to have substrate preference for heavily glycosylated collagens (44–46) such as type IV collagen (28, 32), raising the possibility that *PLOD3* pathogenic variants could also cause Gould syndrome. Supporting this, *Plod3* deficiency in the mouse lead to basement membrane disruptions caused by intracellular accumulation and impaired secretion of type IV collagen that had lower molecular mass suggesting incomplete PTMs (45). Importantly, according to the peptide-based *in vitro* enzymatic analyses, LH3 is striking in that it is a multifunctional enzyme that can execute all the enzymatic roles required for fully modifying lysine to glucosyl galactosyl Hyl—LH, galactosyltransferase, and glucosyltransferase (39, 44, 47–51). However, detailed quantitative PTM analysis has never been conducted for any type of collagen in the context of LH3 deficiency. Thus, in this study, we used CRISPR-mediated gene editing to knockout LH3 both in PFHR9 cells and mouse embryonic fibroblasts (MEFs) and performed a multifaceted investigation including cellular biochemistry, structural assessment, and quantitative PTM characterization for collagen α1α1α2(I), α1α1α1(III), and α1α1α2(IV) to elucidate the role of LH3 in the collagen molecular ensemble. Furthermore, since perinatal cerebral hemorrhage is a cardinal feature of Gould syndrome, we performed a genetic screen for *PLOD3* variants in a fetal stroke cohort that was negative for *COL4A1* and *COL4A2* pathogenic variants.

Results

LH3 is required for secretion of collagen α1α1α2(IV) but not collagens α1α1α2(I) or α1α1α1(III)

To explore the role of LH3 in the molecular ensemble for type IV collagen, we used CRISPR-mediated gene editing to knockout LH3 in PFHR9 cells (Figs. 1*A*, S1 and S2*A*). We selected PFHR9 cells because they are well characterized and secrete large amounts of collagen α1α1α2(IV) because of a duplication of the *Col4a1/Col4a2* locus (28, 32, 52). Western blot analysis revealed a significant reduction in COL4A1 and COL4A2 protein levels in the conditioned culture medium of LH3 KO PFHR9 cells (Figs. 1*B* and S2*B*) and a concomitant increase in intracellular COL4A1 protein levels (Figs. 1*C* and S2*C*), indicating that LH3 deficiency impairs collagen α1α1α2(IV) secretion. To quantitively evaluate the effect of LH3 KO on collagen α1α1α2(IV) biosynthesis, we tested the rate of secretion using pulse-chase labeling with the methionine analog L-azidohomoalanine (AHA) (53, 54). Although this approach is well established for collagen α1α1α2(I), precise measurements of collagen α1α1α2(IV) secretion have not been carefully determined (55–58). AHA pulse-chase analysis in WT PFHR9 cells showed that newly synthesized collagen α1α1α2(IV) takes approximately 7 to 8 h to be secreted (Fig. 1*D*, Table S1, and Fig. S3) compared with approximately 2 h for collagen α1α1α2(I) (59, 60). However, in LH3 KO PFHR9 cells, collagen α1α1α2(IV) secretion was significantly delayed, and only one-fourth of newly synthesized collagen α1α1α2(IV) was secreted even after 48 h in culture (Fig. 1*E* and Table S1).

Because PFHR9 cells produce very low levels of fibrillar collagens, we also used CRISPR-mediated gene editing to knockout LH3 in MEFs to test the effect of LH3 deficiency on collagen α1α1α2(I) and collagen α1α1α1(III) biosynthesis (Figs. 2*A*, S4 and S5*A*). Notably, Western blot analysis showed that the levels of extracellular collagen α1α1α2(I) were increased in LH3 KO MEFs compared with their WT controls (Figs. 2*B* and S5*B*). Since the signals of α2(I) chain were stronger than that of α1(I) chain in Western blot (Figs. 2*B* and S5*B*), we quantitively tested the effect of LH3 KO on collagen secretion rates in MEFs. AHA pulse-chase labeling showed that the secretion rates of fibrillar collagens α1α1α2(I) and α1α1α1(III) were comparable between WT and LH3 KO MEFs, although delays were observed at early time points for collagen α1α1α2(I) (Fig. 2, *C* and *D*, Table S2 and Fig. S6). Collectively, these data suggest that LH3 is critical for biosynthesis of collagen α1α1α2(IV) but not collagens α1α1α2(I) and α1α1α1(III).

LH3 is required for collagen α1α1α2(IV) to form triple helices and higher order oligomeric structures

To molecularly characterize the effects of LH3 deficiency on collagens, we purified collagens α1α1α2(I), α1α1α1(III), and collagen α1α1α2(IV) from the conditioned culture medium of MEFs and PFHR9 cells, respectively (Fig. S7, *A* and *B*). Consistent with the Western blot and AHA pulse-chase data, there was no difference between collagens α1α1α2(I) and

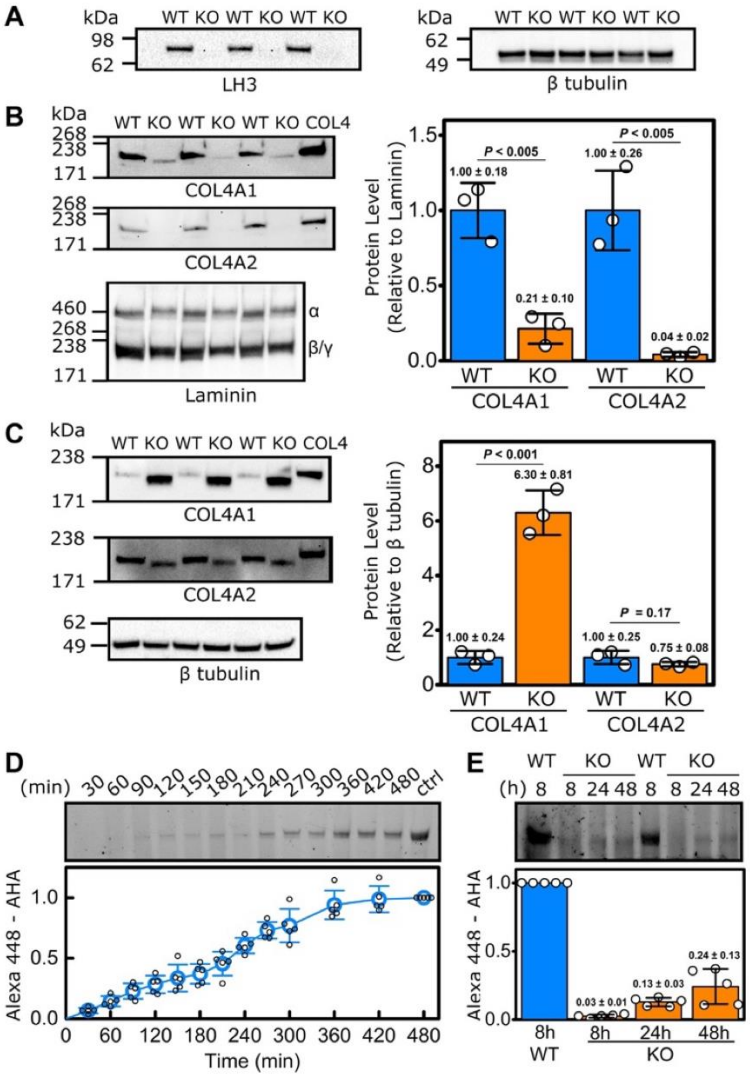


Figure 1. LH3 deficiency in PFHR9 cells impairs collagen α1α1α2(IV) secretion. *A*, semiquantitative Western blots showing successful ablation of LH3 in PFHR9 cells following CRISPR/Cas9-mediated LH3 inactivation (LH3 KO PFHR9 cells). β-tubulin was used as a loading control. *B*, semiquantitative Western blots and quantification of COL4A1 and COL4A2 levels in the conditioned cell culture medium showing reduced secretion of COL4A1 and COL4A2 in LH3 KO PFHR9 cells (KO) compared with their WT controls. Laminin α/β/γ was used as a loading control. *C*, semiquantitative Western blots and quantification of COL4A1 and COL4A2 levels in cell lysates showing intracellular COL4A1 accumulation in LH3 KO PFHR9 cells. β-tubulin was used as a loading control. COL4 indicates control-purified collagen α1α1α2(IV). The quantification data in *B* and *C* are shown as fold relative to WT levels and presented as means ± SD with individual data points representing independent preparations of culture medium and cell lysate and (n = 3). *D*, representative SDS-PAGE gel (*top*) and quantification (*bottom*) of time-dependent extracellular collagen α1α1α2(IV) levels in WT PFHR9 cells measured by AHA-Alexa Fluor pulse-chase labeling. "ctrl" indicates purified AHA-incorporated collagen α1α1α2(IV). The quantification of AHA-incorporated collagen α1α1α2(IV) was set to 1.0 at 480 min (n = 5). *E*, representative SDS-PAGE gel (*top*) and quantification (*bottom*) of time-dependent extracellular collagen α1α1α2(IV) levels for LH3 KO PFHR9 cells measured by AHA-Alexa Fluor pulse-chase labeling, demonstrating significantly impaired collagen α1α1α2(IV) secretion in LH3 KO PFHR9 cells compared with their WT AHA-incorporated collagen α1α1α2(IV) controls at 8 h (n = 5). Source data files for the figure: The uncropped images of Western blots used in the figure are presented in Fig. S2; generation of the purified AHA-incorporated collagen α1α1α2(IV) used as a control in the figure is presented in Fig. S3; and the value of AHA-incorporated collagen α1α1α2(IV) at each time points is shown in Table S1. AHA, L-azidohomoalanine; LH3, lysyl hydroxylase 3.

α1α1α1(III) produced by WT and LH3 KO MEFs (Fig. 3*A*). In contrast, collagen α1α1α2(IV) from LH3 KO PFHR9 cells did not show the higher order collagen α1α1α2(IV) structures that were observed in collagen α1α1α2(IV) from WT PFHR9 cells suggesting an absence of collagen α1α1α2(IV) disulfide crosslinking (Fig. S7*B*). We next used SDS-agarose gel electrophoresis to resolve the oligomeric state of collagen α1α1α2(IV). Consistent with previous reports (52, 56), the protein banding pattern observed in WT PFHR9 cells showed that most collagen α1α1α2(IV) was present as trimers or dodecamers unless denatured under reducing conditions when they were resolved as monomers (Fig. 3*B*). In contrast, most of

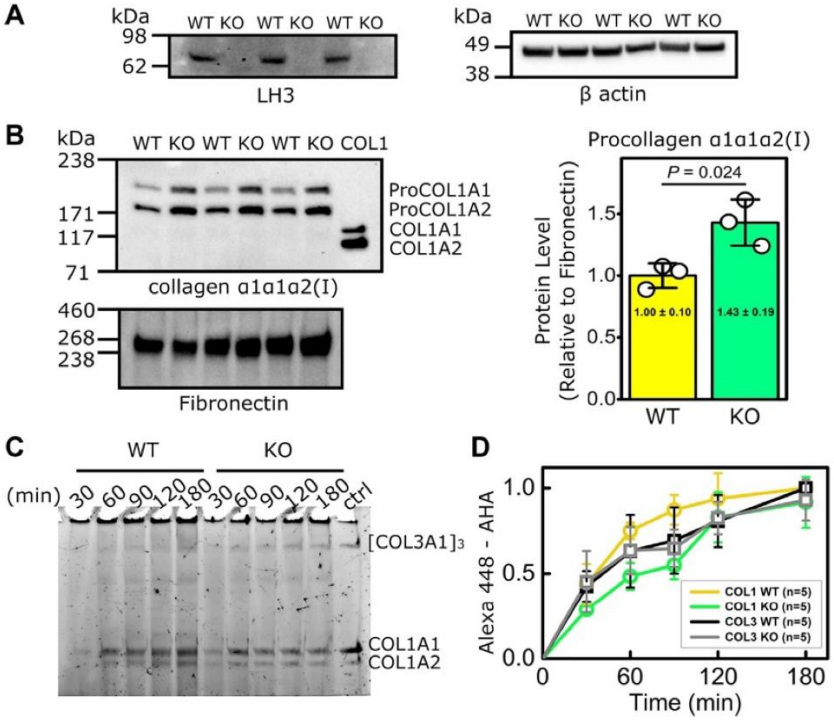


Figure 2. LH3 deficiency in MEFs does not impair biosynthesis of collagen α1α1α2(I) or collagen α1α1α1(III). *A*, semiquantitative Western blots showing successful ablation of LH3 in MEFs following CRISPR/Cas9-mediated LH3 inactivation (LH3 KO MEFs). β-actin was used as a loading control. *B*, semiquantitative Western blots and quantification of collagen α1α1α2(I) in conditioned cell culture medium showing increased secretion of ProCOL1A1 and ProCOL1A2 in LH3 KO MEFs (KO) compared with their WT controls. Fibronectin was used as a loading control. COL1 indicates purified pepsin-treated collagen α1α1α2(I). The quantification data in *B* are shown as fold relative to WT levels and presented as means ± SD with individual data points representing independent preparations of culture medium (n = 3). *C*, representative SDS-PAGE gel. *D*, quantification of time-dependent extracellular levels of collagens α1α1α2(I) and α1α1α1(III) measured by AHA-Alexa Fluor pulse-chase labeling, demonstrating comparable secretion rate between LH3 KO and WT MEFs. "ctrl" indicates purified AHA-incorporated collagens α1α1α2 (I) and α1α1α1(III). WT AHA incorporation signal was set to 1.0 at 180 min (n = 5). Source data files for the figure: The uncropped images of Western blots used in the figure are presented in Fig. S5; generation of the purified AHA-incorporated pepsin-treated collagens α1α1α2(I) and α1α1α1(III) mixture used as a control in *C* is presented in Fig. S6; the value of AHA-incorporated collagens α1α1α2(I) and α1α1α1(III) at each time points is shown in Table S2. AHA, L-azidohomoalanine; LH3, lysyl hydroxylase 3; MEF, mouse embryonic fibroblast.

the collagen α1α1α2(IV) isolated from LH3 KO PFHR9 cells was monomeric and failed to form higher order structures (Fig. 3*B*). Interestingly, we detected a lower molecular weight band in the samples isolated from LH3 KO PFHR9 cells on both SDS-agarose and SDS-PAGE gels (Figs. 3*B* and S7*C*), which was identified as truncated COL4A1 and COL4A2 proteins by LC–MS analysis (Fig. S7*D*). To better characterize collagen α1α1α2(IV) oligomeric structures and network formation, we performed rotary shadow electron microscopy on samples isolated from WT and LH3 KO PFHR9 cells. During type IV collagen network formation, two collagen α1α1α2(IV) heterotrimers associate with each other by way of head-to-head interactions of their carboxy terminal NC1 domains (hexamers) (61) and four collagen α1α1α2(IV) heterotrimers interact by way of lateral and antiparallel interactions of their amino-terminal 7S domains (dodecamers) (62). Electron micrographs from WT PFHR9 cells showed that over 80% of the collagen α1α1α2(IV) was trimeric, but we also found frequent examples of hexamers (~8.5%) and dodecamers (~9.0%; Fig. 3*C*). In contrast, we found reduced proportions of normal trimers and hexamers and did not observe dodecamers in samples from LH3 KO cells. Instead, we detected shortened versions of collagen α1α1α2(IV) hexamers (7.8%) and trimers (72.8%) that appeared truncated at the amino termini and corroborate the SDS-gel electrophoreses and LC–MS data (Figs. 3, *B*, *C*, and S7, *B–D*).

To evaluate the secondary structure of the collagen α1α1α2(IV) triple helix, we measured CD spectra of collagen purified from WT or LH3 KO PFHR9 cells (Fig. 3*D*). At 10 °C, collagen α1α1α2(IV) isolated from WT cells showed the characteristic triple helical spectra with negative and positive peaks at 200 and 220 nm, respectively (63), and these peaks were lost following heat denaturation at 65 °C. In contrast, even at 10 °C, collagen α1α1α2(IV) isolated from LH3 KO PFHR9 cells resembled denatured WT collagen α1α1α2(IV) indicating that collagen α1α1α2(IV) produced by LH3 KO cells did not form stable triple helical helices. The triple helical structure is required for binding of the collagen molecular chaperone heat shock protein 47 (HSP47) (64), which plays a critical role in collagen trafficking by anchoring collagens to TANGO1 at ER exit sites to continue along the secretory pathway to the Golgi (65). To test whether HSP47 could bind purified collagen α1α1α2(IV), we

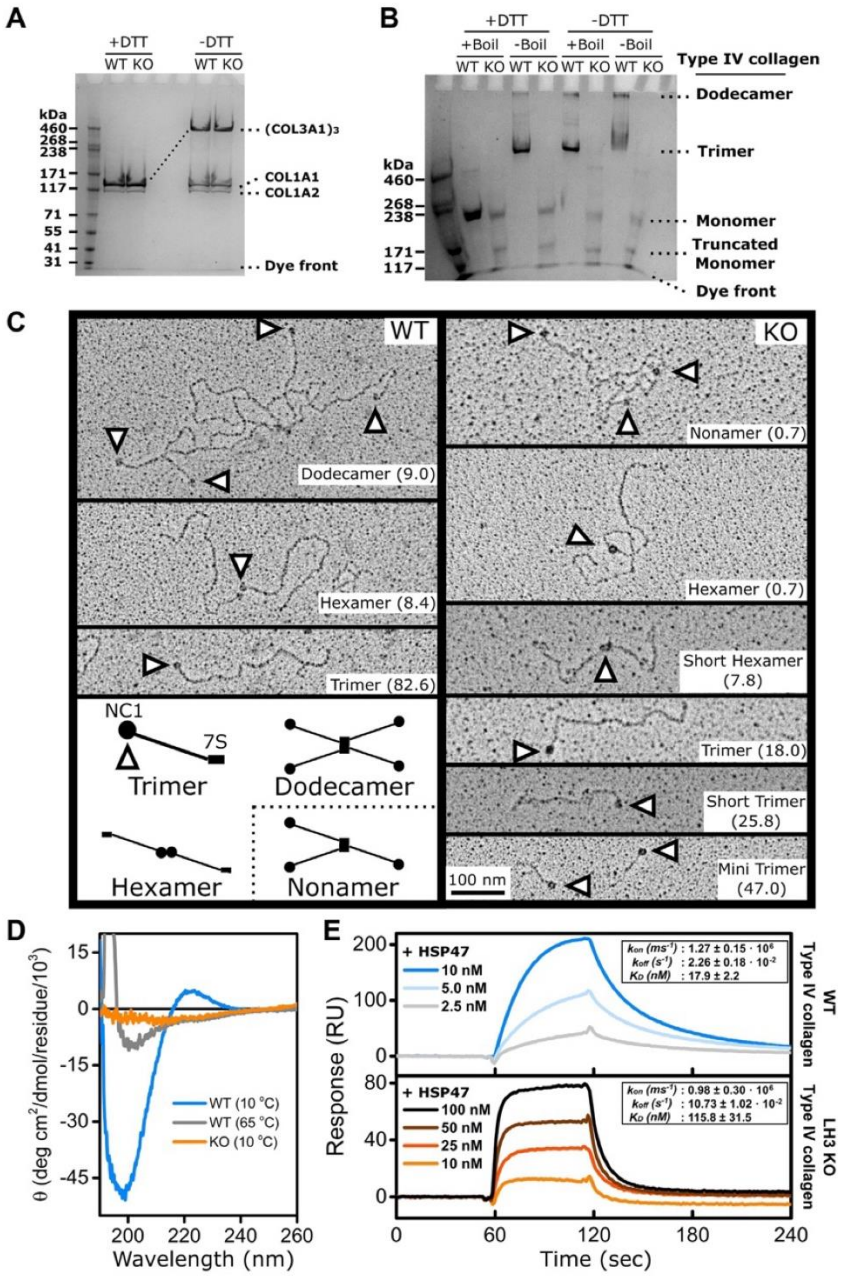


Figure 3. LH3 deficiency impairs collagen α1α1α2(IV) oligomerization and triple helix formation. *A*, representative Coomassie blue–stained 3 to 8% Tris–acetate gel showing pepsin-treated collagens α1α1α2(I) and α1α1α1(III) purified from conditioned culture medium of WT and LH3 KO MEFs. *B*, representative Coomassie blue–stained 3.5% SDS-agarose gel showing collagen α1α1α2(IV) purified from conditioned culture medium of WT and LH3 KO PFHR9 cells. *C*, representative transmission electron rotary shadow microscopy images showing structural alterations of collagen α1α1α2(IV) purified from conditioned culture medium of WT and LH3 KO PFHR9 cells. *Arrowheads* point to globular NC1 domains. Diagram at the *bottom left* illustrates normal collagen α1α1α2(IV) network formation. The frequency of each molecular species is indicated in parentheses (%). *D*, CD spectra of purified collagen α1α1α2(IV) from conditioned culture medium of WT and LH3 KO PFHR9 cells under native (10 °C) and denaturing (65 °C) conditions. *E*, surface plasmon resonance analysis showing concentration-dependent binding kinetics between HSP47 and collagen α1α1α2(IV) purified from conditioned culture medium of WT and LH3 KO PFHR9 cells. Each *curve* represents the average of a minimum of three measurements. The binding constants and affinities are shown as mean ± SD in the *inset*. LH3, lysyl hydroxylase 3; MEF, mouse embryonic fibroblast.

performed surface plasmon resonance (SPR) analysis. HSP47 interacted with immobilized collagen α1α1α2(IV) from WT PFRH9 cells in a concentration-dependent manner (Fig. 3*E*). Although HSP47 also interacted with collagen α1α1α2(IV) from LH3 KO PFRH9 cells, the binding kinetics were altered with lower binding affinity and increased dissociation constant.

LH3 is a selective LH but a general collagen glucosyltransferase

PTMs are important for the structural integrity of collagen triple helices. To understand how LH3 deficiency impacted collagen PTMs, we quantified prolyl hydroxylation, lysyl hydroxylation, and *O*-glycosylated Hyl using amino acid analyses (AAA) combined with LC–MS. First, we performed acid hydrolysis AAA (which removes sugar groups) of individual collagen α chains from WT or LH3 KO cells. We found that prolyl 3 hydroxylation, which occurs infrequently, was largely unchanged by LH3 deficiency across COL1A1, COL1A2, COL3A1, COL4A1, and COL4A2 proteins and that prolyl 4 hydroxylation was reduced for COL4A1 but not for the other α chains (Fig. 4*A* and Table S3). In contrast, we found more unmodified lysines (significantly fewer hydroxylysines), on COL4A1 and COL4A2 proteins isolated from LH3 KO PFHR9 cells compared with WT PFHR9 cells. The levels of lysine/Hyl for COL1A1, COL1A2, and COL3A1 isolated from MEFs were largely unchanged, suggesting that the LH function of LH3 is selective for collagen α1α1α2(IV) (Fig. 4*A* and Table S3). Next, we used alkaline hydrolysis AAA (which preserves sugar attachments) combined with LC–MS to compare the extent of galactosyl Hyl and glucosyl galactosyl Hyl on collagens between WT and LH3 KO cells. Type IV collagen is highly glycosylated, and Hyl residues from collagen α1α1α2(IV) produced by WT PFHR9 cells were almost fully modified to glucosyl galactosyl Hyl (Figs. 4*B* and S8). In contrast, collagen α1α1α2(IV) derived from LH3 KO PFHR9 cells had high occupancy of galactosyl Hyl but not glucosyl galactosyl Hyl suggesting that in addition to its role as an LH, LH3 also participates in glucosyl, but not galactosyl, modifications. As reported previously (44, 66), glucosyl galactosyl Hyl was also reduced in collagen α1α1α2(I) and collagen α1α1α1(III) suggesting that LH3 is also involved in glucosyl modifications of fibrillar collagens (Figs. 4*B* and S8). Taken together, these data suggest that LH3 has selective LH activity but general collagen glucosyltransferase activity.

Rare PLOD3 *variants in fetuses with intracerebral hemorrhage*

The role for LH3 in collagen α1α1α2(I) and collagen α1α1α1(III) glycosylation is consistent with the identification of *PLOD3* pathogenic variants in individuals with connective tissue disorders (39, 41). However, our data demonstrate that LH3 plays an even greater role in collagen α1α1α2(IV) biosynthesis. Consistent with this finding, *Plod3* deficiency in the mouse caused basement membrane defects (44, 45), suggesting that *PLOD3* pathogenic variants may also contribute to intracerebral hemorrhages (ICHs), which constitute a central pathophysiological hallmark of Gould syndrome caused by *COL4A1* and *COL4A2* pathogenic variants (13–15, 67–71). In addition, cerebrovascular hemorrhages have been reported in three patients with a connective tissue disorder caused by biallelic pathogenic variants of *PLOD3* (39, 43). To investigate the potential involvement of *PLOD3* in ICH, we specifically evaluated the presence of *PLOD3* variants in exome sequencing data from a fetal ICH cohort consisting of 113 probands who did not have *COL4A1* or *COL4A2* pathogenic variants. We did not find an enrichment of rare and predicted damaging *PLOD3* qualifying variants between this cohort and the gnomAD control database (Table 1). However, in two unrelated fetuses, we identified three rare (<1/1000) *PLOD3* missense variants—all three were predicted to be damaging by PolyPhen-2 (Fig. 5*A* and Table 2). Fetus F09 was heterozygous for p.L198P variant, and fetus F12 was compound heterozygous for p.R197W and p.P489L. Variants p.R197W and p.L198P are located in the catalytic glycosyltransferase domain close to a poly Asp sequence and predicted to affect the overall structure of active site in the transferase domain (50) (Fig. 5*B*). The p.P489L variant is located in the accessory glycosyltransferase domain and predicted to affect the substrate recognition (Fig. 5*B*). The two fetuses with predicted

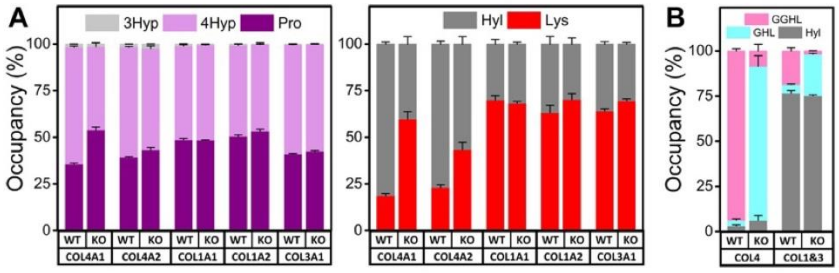


Figure 4. LH3 is a selective lysyl hydroxylase and general collagen glycosyltransferase. *A*, acid hydrolysis amino acid analysis. Histograms showing the occupancy of post-translational modifications in proline (3Hyp + 4Hyp + Pro = 100%) and lysine (Hyl + Lys = 100%) in individual alpha chains of collagens α1α1α2(IV), α1α1α2(I), and α1α1α1(III). *B*, alkaline hydrolysis amino acid analysis. Histogram showing the occupancy of *O*-glycosylation attached to Hyl (GGHL + GHL + Hyl = 100) in purified collagens α1α1α2(IV) and mixture of collagens α1α1α2(I) and α1α1α1(III). Source data files for the figure: The value of each amino acid, the number of biological replicates, and detailed statistical analyses for data shown in the figure are presented in Table S3 and Fig. S8, respectively. GGHL, glucosyl galactosyl Hyl; GHL, galactosyl Hyl; Hyl, hydroxylysine; 3Hyp, 3-hydroxyproline; 4Hyp, 4-hydroxyproline; LH3, lysyl hydroxylase 3; Lys, lysine; Pro, proline.

Table 1
Enrichment analysis of rare predicted damaging *PLOD3* variants in fetus cohort *versus* gnomAD cohorts

gnomAD	Number of loss-of-function and missense variants with an MAF ≤1/1000 in *PLOD3* gene	*p*
v2.1 (N = 141,456 individuals)	2160	0.69
v3.1 (N = 76,156 individuals)	1225	0.7

MAF, minor allele frequency.

pathogenic variants had ICH identified on ultrasound follow up during the third trimester of pregnancy. Upon cerebral imaging, they presented a bilateral ventriculomegaly and intraparenchymal lesions (Table 3). Because of the severity of these lesions, the parents opted for termination of pregnancy. The lesions were confirmed by postmortem examination, and the fetuses showed some phenotypic differences. The fetus F12 had intraparenchymal hemorrhagic and ischemic lesions, whereas the fetus F09 had a microcephaly, a very thin corpus callosum, and a white matter atrophy. Neither had bone abnormalities or signs of external visceral malformation.

Discussion

To clarify the role(s) for LH3 in collagen biosynthesis, we generated LH3 KO cells using CRIPSR-mediated gene editing to conduct detailed quantitative and molecular analyses of collagens α1α1α2(I), α1α1α1(III), and α1α1α2(IV) implicated in OI, EDS, and Gould syndrome, respectively. We discovered that LH3 deficiency severely impaired biosynthesis of collagen α1α1α2(IV) but not collagen α1α1α2(I) or collagen α1α1α1(III). Our findings revealed altered PTMs in collagen α1α1α2(IV) produced by LH3 KO PFHR9 cells characterized by reduced and impaired modification of lysine to Hyl and galactosyl Hyl to glucosyl galactosyl Hyl. These PTM alterations prevented proper heterotrimer formation and HSP47 binding, which impaired collagen α1α1α2(IV) secretion. Notably, collagen α1α1α2(IV) that was secreted from LH3 KO PFHR9 cells showed impaired collagen α1α1α2(IV) network formation and failed to make higher order oligomers. In contrast, LH3 deficiency did not impact primary modifications to prolyl and lysyl residues for collagens α1α1α2(I) and α1α1α1(III) but specifically reduced modification of galactosyl Hyl to glucosyl galactosyl Hyl. Thus, we suggest that LH3 may be a selective LH at least for collagen α1α1α2(IV) and that network-forming and fibrillar collagens are both important substrates for LH3 glucosyltransferase activity.

Our results indicate that LH3 plays a critical role for collagen α1α1α2(IV) biosynthesis in PFHR9 cells but not for collagens α1α1α2(I) and α1α1α1(III) in MEFs. Surprisingly, Western blot analysis for PFHR9 cells showed that protein levels of intracellular COL4A2 in KO cells were similar to WT control cells; however, the protein migrated slightly faster and its secretion was significantly impaired (Fig. 1, *B* and *C*), indicating there could be a different quality control mechanism between COL4A1 and COL4A2 proteins. Although the majority of synthesized collagen α1α1α2(IV) was accumulated

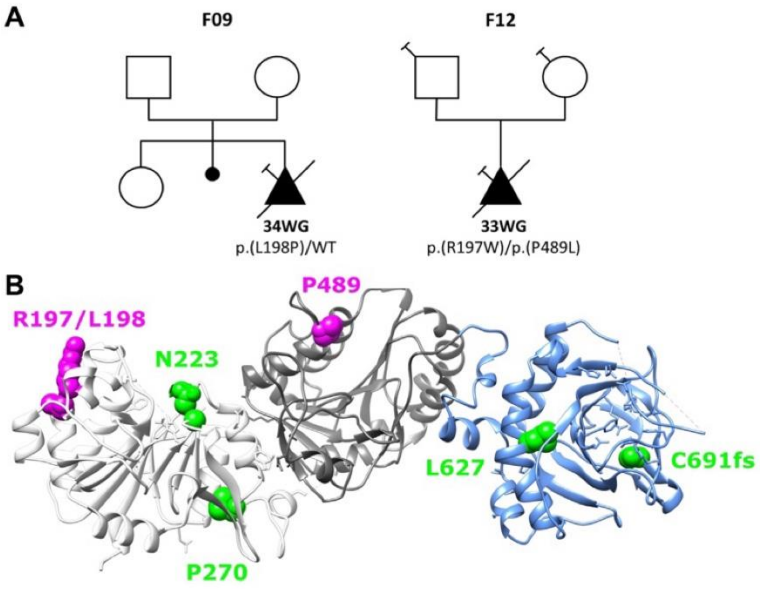


Figure 5. Identification of four rare *PLOD3* variants in fetal ICH cases. *A*, genealogical trees of the three families with *PLOD3* variants. *Square* = male, *circle* = female, *triangle* = pregnancy not carried to term, *black filled symbol* = affected individual, *empty symbol* = clinically healthy relative, *diagonal black line* = deceased fetuses, and *syringe symbol* = blood sampled individual. The age of termination of pregnancy is indicated under each case. *B*, the crystal structure of the LH3 shows three different domains: catalytic glycosyltransferase domain (*white*), accessory glycosyltransferase domain (*gray*), and lysine dioxygenase domain (*light blue*). Missense variants identified in humans are indicated in different colors on the LH3 crystal structure. *Green* indicates previously reported variants have connective disorders resembling Stickler syndrome-like and epidermolysis bullosa. *Magenta* indicates variants identified in this study. ICH, intracerebral hemorrhage; LH3, lysyl hydroxylase 3.

Table 2
Rare variants identified in *PLOD3* gene in the cohort of 113 ICH fetuses

Fetus	HGVSc position (NM_001084.5)	Protein change	Variant class	Status	gnomAD, v.2.1 AF (%)	PolyPhen-2 prediction	ACMG class
F09	c.593T>C	p.(L198P)	Missense	HTZ	0.0004	Probably damaging (0.994)	3
F12[a]	c.589C>T	p.(R197W)	Missense	HTZ	0.0012	Possibly damaging (0.735)	3
F12[a]	c.1466C>T	p.(P489L)	Missense	HTZ	0.051	Possibly damaging (0.769)	3

Abbreviations: ACMG, American College of Medical Genetics and Genomics; AF, allele frequency; HTZ, heterozygous.
ClinVar accession numbers: SCV002558784 (p.L198P), SCV002558785 (p.R197W), and SCV002558786 (p.P489L).
[a] Biallelic variants (inherited from each parent).

intracellularly in LH3 KO PFHR9 cells, some collagen α1α1α2(IV) was successfully secreted but did not have stable triple helical structures. Based on the current understanding of fibrillar collagen biosynthesis, this finding is intriguing for two reasons. First, the amount of 4Hyp is essential to the formation of triple helix rigidity (29). Type II collagen containing approximately 30% 4Hyp from *P4ha1*$^{+/-}$, *P4ha2*$^{-/-}$ mice (*P4ha1* and *P4ha2* encoded prolyl 4-hydroxylase 1 and 2, respectively) decreased its thermal stability compared with a WT control; however, this type II collagen remained capable of forming a triple helical structure (72). Notably, the COL4A1 produced by LH3 KO PFHR cells showed a reduction of 4Hyp (Fig. 4*A*) but still contained around total 45% 4Hyp in the collagen α1α1α2(IV) heterotrimer isolated from LH3 KO PFHR9 cells, which is calculated by the amount of 4Hyp in COL4A1 and COL4A2 (Table S3). Second, the amounts of Hyl do not affect triple helix formation. Collagen α1α1α2(I) containing only 1.5% Hyl can form fully rigid triple helices and secrete into the extracellular matrix properly (24). Also, increased hydroxylysines in collagen α1α1α2(I) do not improve collagen thermal stability (73). Based on these observations, it is tempting to speculate that lysyl hydroxylation and/or the subsequent *O*-glycosylation critically contributes triple helix stability in collagen α1α1α2(IV).

Interestingly, the AHA pulse-chase experiments showed that the secretion rate of collagen α1α1α2(IV) was remarkably slower than that of collagens α1α1α2(I) and α1α1α1(III) (Figs. 1*D* and 2*D*). A significant difference between collagen α1α1α2(IV) and fibrillar collagens (type I, II, III, and V) is the presence of multiple interruptions of the collagenous repeats, which occur more frequently nearer the amino terminus (74). Repeat interruptions, which would be pathogenic if they occurred in fibrillar collagens, may interfere with propagation of triple helix formation, requiring renucleation of the three alpha chains, and likely contribute to the much longer secretion time for collagen α1α1α2(IV) compared with collagens α1α1α2(I) and α1α1α1(III). We hypothesize that the complement of PTMs near repeat interruptions may provide thermal or structural stability required for renucleation and propagation of the collagen α1α1α2(IV) triple helix. Supporting this notion, collagen α1α1α2(IV) from LH3 KO cells appeared shorted on rotary shadow images, which may suggest that the amino termini are disordered and cannot be resolved by electron microscopy. Alternatively, they may be more susceptible to cleavage, leading to a truncated protein of lower molecular weight as demonstrated by SDS gel electrophoreses and LC–MS analysis (Figs. 3, *B*, *C* and S7, *B–D*). Structural alterations of the 7S domain at the amino termini could be responsible for the inability to form higher order dodecamer structures in collagen α1α1α2(IV) isolated from LH3 KO PFHR9 cells (Figs. 3, *B*, *C* and S7, *B–D*). While the CD spectra showed that the overall secondary structure of collagen α1α1α2(IV) produced by LH3 KO was not fully triple helical (Fig. 3*D*), HSP47 was still able to bind with this collagen α1α1α2(IV). We speculate that this binding occurs on the triple helical carboxyl terminus region and weakly associate with the immature amino-terminal gelatin-like structure (Fig. 3*E*). Interestingly, small amounts of collagen α1α1α2(IV) were secreted from LH3 KO cells without a robust triple helical structure, which raises the question: what are the essential chemical or structural factors for collagen α1α1α2(IV) triple helix formation and for passing through quality control checkpoints in the ER and secretory pathway? Further experiments are required to provide additional mechanistic insights

Table 3
Main clinical and pathological features of ICH-affected fetuses with *PLOD3* variants

Fetus	F09	F12
Sex	Female	Female
Gestation at diagnosis (weeks)	31	30
Brian imaging findings	Microcephaly, ventriculomegaly, interhemispheric cyst, porencephalic cavities, thin corpus callosum	Asymmetric bilateral ventriculomegaly, ischemic–hemorrhagic lesions in the right hemisphere
Outcome	TOP	TOP
Fetal brain examination	Important ventriculomegaly, white matter atrophy, severe microcephaly with diffuse bilateral lesions, absent or very thin corpus callosum	Asymmetric bilateral ventriculomegaly, intraparenchymal hemorrhages with cortical and subcortical ischemic foci
Skeletal X-ray examination	No sign of bone abnormality except for the absence of the 12th right rib	No sign of bone abnormality
Other autopsy anomalies	No sign of external or visceral malformation	No sign of external or visceral malformation

Abbreviation: TOP, termination of pregnancy.

into the role of PTMs in collagen α1α1α2(IV) biosynthesis, secretion, and function(s).

Since our findings establish a critical role for LH3 in collagen α1α1α2(IV) biosynthesis, we predicted that *PLOD3* might lead to clinical manifestations similar to those observed in individuals with *COL4A1* and *COL4A2* pathogenic variants causing Gould syndrome, a multisystem disorder frequently manifesting as fetal and perinatal ICH (13–15, 67–70). Previously identified pathogenic *PLOD3* missense variants cause phenotypically overlapping connective disorders including Stickler-like syndrome and epidermolysis bullosa (39–42). In addition to the clinical features encountered in these two conditions, spontaneous ICH has been reported by Salo *et al.* (39) in two siblings. One of them was a female child with a connective tissue disorder, consisting of craniofacial dysmorphism, sensorineural deafness, and skeletal features, who presented a spontaneous large brain hematoma. Her male sibling was stillborn at 28 weeks of gestation; in addition to skeletal anomalies, autopsy of this fetus showed porencephaly. DNA analysis showed biallelic *PLOD3* variants (p.N223S/p.C691AfsX9) located in a conserved region of the LH3 amino acid sequence responsible for collagen glycosyltransferase and LH activities. Recently, a homozygous *PLOD3* small inframe deletion was shown to cause a cerebral small vessel disease characterized by multiple white matter hypersignals and bleeding foci in an infant with a few dysmorphic features (43). To evaluate the contribution of *PLOD3* pathogenic variants to cerebrovascular manifestations, we analyzed a series of 113 cases of fetal ICH that were negative for *COL4A1* and *COL4A2* pathogenic variants. We identified a rare and predicted damaging heterozygous variant in one fetus (F09) and two distinct biallelic candidate variants in another fetus (F12) (Fig. 5 and Table 2). When comparing the whole cohort of fetuses to control databases, we did not detect an enrichment in rare and predicted damaging *PLOD3* variants; however, this lack of enrichment could be attributable to the high genetic heterogeneity of fetal ICH (71). Additional work is required to establish the causality of the heterozygous *PLOD3* variant present in fetus F09. Indeed, connective tissue disorders associated with *PLOD3* are typically autosomal recessive, and heterozygous parents of affected cases are reported to be asymptomatic. We cannot exclude however that some heterozygous *PLOD3* variants could have dominant negative effects or that they could act as risk factors acting in concert with other triggering events to cause fetal ICH. For fetus F12, who is a compound heterozygote with both variants being predicted to be possibly pathogenic, no tissue was available to evaluate the consequences of the two variants at the protein level. Autopsy data including skeletal X-ray data were normal. Additional *in vitro* experiments are required to determine the causality of these variants, and larger cohorts of fetuses and children with ICH should be analyzed to identify additional cases.

In conclusion, our results suggest that LH3 plays a critical and multifunctional role in collagen α1α1α2(IV) biosynthesis, and that *PLOD3* variants could lead to clinical manifestations associated with *COL4A1* and *COL4A2* pathogenic variants.

Experimental procedures

Cell culture

PFHR9 cells (CRL-2423) and MEFs (CRL-1503) were purchased from American Type Culture Collection. These cells were cultured in Dulbecco's modified Eagle's medium (DMEM)/high glucose/pyruvate (Gibco; catalog no.: 11995065) supplemented with 10% (v/v) fetal bovine serum (Atlanta Biologicals), pencillin streptomycin glutamine 100× (Gibco; catalog no.: 10378016), and 5 mM Hepes in the presence of ascorbic acid phosphate (100 μg/ml; Wako Chemicals). After this section, standard DMEM indicates the aforementioned DMEM.

CRISPR-mediated KO of LH3

LH3 was knocked out in PFHR9 cells and MEFs using two CRISPR/Cas9 All-in-One plasmid of pSpCas9(BB)-2A-Puro (PX459) v2.0 and pSpCas9(BB)-2A-GFP (PX458) including the LH3 guide RNA (gRNA) sequence for *Plod3* exon 5 (5′-TCCACTGGCGGACAATCTGATGG-3′) and exon 7 (5′-TCG TGTGCGCATCCGGAATGTGG-3′), which were designed and purchased from GenScript, respectively. Two plasmids were transfected into the cells using Lipofectamine 3000 (Life Technologies) according to the manufacturer's instructions. After 24 h, transfected cells were treated with 10 μg/ml puromycin (Gibco) to select cells containing the PX459 plasmid. After puromycin selection, the GFP-positive cells containing the PX458 plasmid were isolated using a corning ring and cultured them with standard DMEM.

DNA sequencing

To confirm LH3 gRNA edited *Plod3* exon 5 and 7, PCR products including the region targeted with gRNA were sequenced. DNA was prepared from WT and LH3 KO cells. Cells were plated onto 6 cm dish and grown to 100% confluency. Cells were collected by centrifugation after treating with trypsin (Promega). This cell pellet was digested by 3 mg/ml proteinase K in 20 μl of 50 mM Tris–HCl buffer containing 20 mM NaCl, 1% SDS, and 1 mM EDTA, pH 8.0, at 95 °C for 5 min after preincubating at 60 °C for 60 min. After the digestion, 200 μl of water was added to this solution as a template for PCR. PCR was performed with primer set A (forward: 5′-AAGCGCTTCCTCAACTCTG-3′ and reverse: 5′-TCCCTCATTACCCACCTCTAA-3′) and B (forward: 5′-ACTCCTGAGCATGTGTGT TAG-3′ and reverse 5′-ACAGGGTTGAGTGGCTAAAG-3′) for exon 5 and 7, respectively. Sequencing samples including the forward primer were sent to Quintara Biosciences. Sequencing results were analyzed by GeneStudio Professional Edition, version 2.2.0.0 (GeneStudio, Inc).

Protein analyses in cell lysate and conditioned culture medium

WT and LH3 KO cells were plated and grown to 80 to 90% confluency. Following stimulation of procollagen biosynthesis with ascorbic acid for 1 day, the medium was replaced to the fresh standard DMEM, and the cells were cultured for 2 days. The medium containing secreted proteins was dissolved in Bolt LDS sample buffer 4× (Life

Technologies) with reducing agents. After the cells were washed with PBS once, cell lysates were extracted using M-PER (Thermo Fisher Scientific) containing Halt Protease Inhibitor Cocktail, EDTA-Free (Thermo Fisher Scientific) at 4 °C according to the manufacturer's instructions. After centrifugation, soluble proteins in the supernatant were mixed with the same sample buffer with reducing agents. These protein solutions were separated on SDS-PAGE gels and then transferred to polyvinylidene difluoride (PVDF) membranes, and Western blots were performed. Gels, buffers, transfer condition, and antibodies are listed in Table S4. Blots were developed with horseradish peroxidase–enhanced Super-Signal West Pico Chemiluminescent Substrate (Thermo Fisher Scientific) and detected by ChemiDoc MP imaging system (Bio-Rad) using the software Image Lab, version 4.0.1 (Bio-Rad). The intensities of protein signals were measured by ImageJ (National Institutes of Health).

Secretion rate assay for PFHR9 cells

WT and LH3 KO cells were cultured in a 16 cm dish grown to 100% confluency and equally plated to three 6-well tissue culture plates after the cells were trypsinized. Following stimulation of procollagen biosynthesis with ascorbic acid overnight, the cells were preincubated in pulse medium (Met- and Cys-free DMEM with 10% fetal bovine serum and 100 μg/ml ascorbic acid 2-phosphate) without AHA (Click Chemistry Tools) at 37 °C for 30 min to deplete methionine. Pulse labeling was performed with 1 mM AHA in pulse medium at 37 °C for 15 min, and cells were incubated in 1.5 ml chase medium (standard DMEM) at 37 °C for different time intervals (*i.e.*, every 30 min by 360 min, 420 min 480 min, 24 h and 48 h). The conditioned culture medium was collected at each time point and frozen at −20 °C. After thawing and centrifuging at high speed to remove residual cells and large aggregates, 1.4 ml of medium was mixed with 1.4 ml of 0.1 M Tris–HCl containing 1.0 M urea, pH 7.5. This solution was incubated with 0.4 ml of Q Sepharose Fast Flow (GE Healthcare) to remove AHA-incorporated laminins at room temperature for an hour. After centrifugation at high speed for 10 min, 2.8 ml supernatant was mixed with 0.2 ml 10 M acetic acid and 1 ml of 5 M NaCl at 4 °C overnight. Collagen α1α1α2(IV) was precipitated at the highest speed by a tabletop centrifuge. The pellet containing collagen α1α1α2(IV) was dissolved in 27 μl of Tris-buffered saline and labeled with 3 μl of Click-iT Alexa Fluor 488 sDIBO Alkyne (Thermo Scientific) at room temperature for an hour. In LH3 KO secretion rate assay, the pellets from two wells in the 6-well plate were combined to detect adequate signals. This labeled solution was mixed with 10 μl of 4× Bolt LDS sample buffer with reducing agent and separated on precast 6% Tris–glycine gel (Invitrogen). The signals from AHA-Alexa Fluor 488 were detected by the ChemiDoc MP imaging system using the software Image Lab, version 4.0.1. The intensities of fluorescent signals were measured by ImageJ. The AHA-incorporated collagen α1α1α2(IV) as a loading control was purified from conditioned culture medium after an hour pulse and the overnight chase. The purification method is described in the section "Collagen α1α1α2(IV) purification" later. This assay was performed five times, and representative images are shown.

Secretion rate assay for MEFs

WT and LH3 KO cells were cultured in a 16 cm dish grown to 100% confluency and equally plated to five 10 cm dishes plate after the cells were trypsinized. Following stimulation of procollagen biosynthesis with ascorbic acid overnight, cells were preincubated in pulse medium without AHA at 37 °C for 30 min to deplete methionine. Pulse labeling was performed with 1 mM AHA in pulse medium at 37 °C for 15 min, and cells were incubated in 3.7 ml chase medium (standard DMEM) at 37 °C for different time intervals (*i.e.*, every 30, 60, 90, 120, and 180 min). The conditioned culture medium was collected at each time point and frozen at −20 °C. After thawing and centrifuging at high speed to remove residual cells and large aggregates, 3.5 ml of medium was mixed with 0.5 ml 10 M acetic acid and 20 μl of 100 mg/ml pepsin at 4 °C overnight. After adding 1.0 ml of 5 M NaCl to the pepsin-treated solution and incubating at 4 °C overnight, collagens α1α1α2(I) and α1α1α1(III) were precipitated at the highest speed by a tabletop centrifuge. The pellet containing collagens α1α1α2(I) and α1α1α1(III) was dissolved in 27 μl of Tris-buffered saline and labeled with 3 μl of Click-iT Alexa Fluor 488 sDIBO Alkyne at room temperature for an hour. This labeled solution was mixed with 10 μl of 4× Bolt LDS sample buffer without reducing agent and separated on precast 6% Tris–glycine gel. The signals from AHA-Alexa Fluor 488 were detected by the ChemiDoc MP imaging system using the software Image Lab, version 4.0.1. The intensities of fluorescent signals were measured by ImageJ. The AHA-incorporated type I and type III collagen as a loading control was purified from conditioned culture medium after an hour pulse and the overnight chase. The purification method is described in the section"Collagens α1α1α2(I) and α1α1α1(III) purification" later. This assay was performed five times, and representative images are shown.

Collagen α1α1α2(IV) purification

Extraction of collagen α1α1α2(IV) was performed using the standard DMEM with PFHR9 cells. DMEM was prepared by the same procedures as described in the aforementioned "Protein level analyses" section, except using serum-free DMEM instead of the standard DMEM. The DMEM was dialyzed with 0.2 M acetic acid at 4 °C overnight. After spinning down to remove residual cells or aggregates, the DMEM was dialyzed with 0.2 M acetic acid containing 1.2 M NaCl at 4 °C overnight. The precipitates including collagen α1α1α2(IV) were collected by centrifugation at 13,000 rpm for 20 min using JA-14 rotor (Beckman) and resuspended with 50 mM Tris–HCl buffer containing 25 mM NaCl and 0.5 M urea, pH 7.5. After dialysis to the same buffer at 4 °C overnight and centrifugation at 13,000 rpm for 10 min using JA-20 rotor (Beckman), the solution was applied to HiTrap Q FF column (cytiva), and the flow through fraction was collected. The flow through was dialyzed to 50 mM Mes buffer containing 25 mM

NaCl and 0.5 M urea, pH 6.0, at 4 °C overnight and applied to HiTrap SP HP column (cytiva). After washing with the same Mes buffer at least five column volumes, collagen α1α1α2(IV) was eluted by 50 mM Mes buffer containing 250 mM NaCl and 0.5 M urea, pH 6.0. The fractions containing collagen α1α1α2(IV) were concentrated using Amicon Ultra centrifugal filters Ultracel—100 K (Millipore) at the speed × 100*g* up to 1.5 ml as a total volume. The concentrated fraction was dialyzed to 0.2 M acetic acid at 4 °C overnight, and 0.5 ml of 5 M NaCl was added (final concentration of 1.25 M NaCl) to the dialyzed solution. After mixing at 4 °C overnight, collagen α1α1α2(IV) was precipitated by centrifugation at 14,000 rpm for 20 min using a tabletop centrifuge. The pellet was resuspended with 0.2 M acetic acid dialyzed to 0.2 M acetic acid at 4 °C overnight to remove residual NaCl for experiments.

Collagens α1α1α2(I) and α1α1α1(III) purification

Extraction of collagens α1α1α2(I) and α1α1α1(III) was performed using the cultured standard DMEM with MEFs. The DMEM was prepared by the same procedures as described in aforementioned "Protein level analyses" section. The DMEM was dialyzed with 0.5 M acetic acid at 4 °C overnight. After spinning down to remove residual cells or aggregates, the DMEM was incubated with pepsin added to a final concentration of 0.25 mg/ml at 4 °C overnight. The pepsin-treated DMEM was dialyzed with 0.5 M acetic acid containing 0.7 M NaCl at 4 °C overnight. The precipitates including collagens α1α1α2(I) and α1α1α1(III) were collected by centrifugation at 13,000 rpm for 20 min using JA-14 rotor. The pellet was resuspended in 0.2 M acetic acid, and, this solution contained enriched collagens. α1α1α2(I) and α1α1α1(III) were dialyzed to 0.2 M acetic acid at 4 °C overnight to remove residual NaCl for experiments.

SDS-gel analyses

To check purity, the isolated collagens were run on a Bolt 4 to 12% Bis–Tris Plus gel (Invitrogen) in the presence or the absence of DTT with Mes running buffer (Novex) after boiling denaturation. To compare differences in migration, the purified collagens α1α1α2(I) and α1α1α1(III) mixture were run on a 3 to 8% Tris–acetate gel (Invitrogen) in the presence or the absence of DTT with boiling denaturation. To demonstrate higher order collagen α1α1α2(IV) assemblies, SDS-agarose gel electrophoresis was performed for purified WT and LH3 KO collagen α1α1α2(IV). About 3.5% agarose was dissolved in 0.2 M Tris–HCl, pH 8.8. After measuring the total weight of the beaker and agarose solution, the beaker was heated in a microwave oven and gently swirled to thoroughly mix the agarose solution. After adding 20% SDS up to 0.25% and sufficient hot distilled water to obtain the initial weight, the SDS-agarose solution was mixed thoroughly again. Bio-Rad Mini-PROTEAN Tetra Handcast Systems was adapted to cast the gel, and hot SDS-agarose solution was loaded between preheated glass plates of a vertical electrophoresis apparatus. Four different sample conditions were prepared using purified WT and LH3 KO collagen α1α1α2(IV) in the presence or the absence of DTT with and without boiling denaturation. The samples were separated on a SDS—3.5% agarose gels with Tris–glycine SDS running buffer using 75 V constant voltage at 4 °C for about 2 h. All gels were stained with GelCode Blue Stain Reagent (Thermo Scientific). For PTM analysis, we used 6% Tris–glycine SDS-PAGE and 3.0% SDS-agarose gel to separate individual collagen chains of collagens α1α1α2(I) and α1α1α1(III) and collagen α1α1α2(IV) under reducing condition, respectively. After gels were electrophoretically transferred to PVDF membranes with 0.05% SDS, the membranes were stained with Ponceau S.

Protein identification by LC–MS

Gel bands were reduced with 10 mM DTT at 56 °C for 30 min followed by alkylation with 50 mM iodoacetamide at room temperature for 30 min. The samples were digested with trypsin at 37 °C for 16 h, and the generated tryptic peptides were sequentially extracted from the gels with 5% formic acid, 5% formic acid/50% acetonitrile, and 5% formic acid/95% acetonitrile. The extracted solutions were concentrated by a centrifugal evaporator CVE-3100 (EYELA) and then analyzed by LC–MS on a maXis II quadrupole time-of-flight mass spectrometer (Bruker Daltonics) coupled to a Shimadzu Prominence UFLC-XR system (Shimadzu) with chromatographic separation using an Ascentis Express Peptide ES-C18 column (2.7 μm particle size, L × I.D. 150 mm × 2.1 mm; Supelco) (75). The mass spectrometry (MS) scan and MS/MS acquisition were performed over the *m/z* ranges of 50 to 2500 with a frequency of 5 Hz. The acquired MS/MS spectra were searched against the UniProtKB/Swiss-Prot database (release 2018_05) for *Mus musculus* species using ProteinPilot software 4.5 (AB Sciex), as described previously (24).

Rotary shadow electron microscopy

Rotary shadowing was performed by methods described previously (76). In brief, samples were prepared to mix WT and LH3 KO collagen α1α1α2(IV) (approximate concentrations of 100 μg/ml) in 0.2 M acetic acid with glycerol to a final concentration of 70% glycerol. Each protein was nebulized onto freshly cleaved mica chips using an airbrush. The samples were then rotary shadowed with carbon/platinum using an electron beam gun within a vacuum evaporator. Images were acquired using an FEI G20 transmission electron microscope. Proportions of different heterotrimer structures were obtained from images of WT (n = 690) and LH3 KO (n = 295) molecules, and representative images were used in Figure 4*C*.

CD

CD spectra were recorded on an AVIV 202 spectropolarimeter (AVIV Biomedical, Inc) using a Peltier thermostat–controlled cell holder and a 1 mm path length rectangular quartz cell (Starna Cells, Inc). Protein concentrations were determined by acid hydrolysis amino acid analysis described in the section "PTM analysis by LC–MS" and using collagen solution instead of PVDF membrane. Both WT and LH3 KO collagen α1α1α2(IV) were measured in 0.2 M acetic

acid at 10 °C. The denatured WT collagen α1α1α2(IV) was measured using the cuvette after the measurements at 10 °C with heating up to 65 °C for 1 h. The spectra represent the average of at least three scans recorded at a wavelength resolution of 0.1 nm.

SPR analysis

SPR experiments were carried out using a BIAcore X instrument (GE Healthcare). Purified WT and LH3 KO collagen α1α1α2(IV) were immobilized on a CM5 sensor chip by amide coupling. The approximate coupled protein concentration was 1.2 ng/mm^2 (1200 response units) and 1.4 ng/mm^2 (1400 response units) of WT and LH3 KO collagen α1α1α2(IV), respectively. The experiments were performed at 20 °C in 10 mM Hepes, pH 7.4, containing 150 mM NaCl and 0.005% Surfactant P20 using a flow rate of 10 μl/min. All curves are the average of at least three replicates, and three independent measurements were performed. For the analysis of the binding affinity, the curves were fitted with the Langmuir binding model (BIAevaluation software; GE Healthcare). Hsp47 were purified from 17-day-old chicken embryos using methods described previously (77).

PTM analysis by LC–MS

The hydroxylation rate of Lys (Lys + Hyl = 100%) and Pro (Pro + 4-Hyp + 3-Hyp = 100%) in each α chain of collagens α1α1α2(IV), α1α1α2(I), and α1α1α1(III) was evaluated by LC–MS after acid hydrolysis, as described previously (78). In brief, after respective α chains were transferred to PVDF membranes as described in aforementioned SDS-gel analysis, the membrane bands were excised and subjected to acid hydrolysis (6 N HCl/1% phenol, 110 °C for 20 h in the gas phase under N_2) after adding stable isotope–labeled collagen (SI-collagen) (79) as an internal standard. The acid hydrolysates were analyzed by LC–MS in multiple reaction monitoring mode on a QTRAP 5500+ hybrid triple quadrupole/linear ion trap mass spectrometer (AB Sciex) coupled to an ExionLC AD HPLC system (AB Sciex) with a ZIC-HILIC column (3.5 μm particle size, L × I.D. 150 mm × 2.1 mm; Merck Millipore). The content of Pro, 3-Hyp, 4-Hyp, Lys, and Hyl was quantitated by the peak area ratio of the analytes relative to the corresponding stable isotopically heavy analytes derived from SI-collagen (79). The occupancy of *O*-glycosylations attached to Hyl (free Hyl + galactosyl Hyl + glucosyl galactosyl Hyl = 100) in the collagens α1α1α2(I) and α1α1α1(III) mixture and purified collagen α1α1α2(IV) was evaluated by LC–MS after alkaline hydrolysis, as described previously (79). In brief, the purified collagen samples were subjected to alkaline hydrolysis (2 N NaOH, 110 °C for 20 h under N_2) after adding SI-collagen as the internal standard. The content of free Hyl, galactosyl Hyl, and glucosyl galactosyl Hyl was quantitated by LC–MS analysis of the alkaline hydrolysates as described previously.

Human material and ethics statement

About 113 unrelated fetuses affected by an ICH grade III or IV were included following termination of pregnancy or intrauterine fetal death. They were referred for *COL4A1/COL4A2* screening to the French national molecular genetics reference center for inherited cerebrovascular disorders (Saint Louis Hospital, Paris) and shown to be negative using a targeted high-throughput sequencing (conditions available upon request). Fetal ICH was in most cases detected at systematic second and third trimester ultrasound examinations and in some cases confirmed by fetal MRI and/or pathological examination. Systematic review of medical charts was performed in order to exclude fetuses with an identifiable cause or known risk factor for ICH, including evidence of maternal trauma during pregnancy, cocaine or maternal drug use, maternal or neonatal infections, and fetal alloimmune thrombocytopenia. Written informed consent for genetic investigation and research was provided by parents in accordance with the declaration of Helsinki and the French law. This study has been approved by the Inserm Ethics Committee (INSERM IRB00003888). Genomic DNA was isolated from postmortem fetal tissue and peripheral blood leukocytes of both parents and relatives when available. Prior to whole exome sequencing, we excluded the presence of a mutation in *COL4A1* and *COL4A2* genes by performing a targeted high-throughput sequencing (conditions available upon request).

Exome sequencing

Exon capture was performed at the IntegraGen platform (Evry) for fetus probands and relatives using the SureSelect Human All Exon V5-UTR (Agilent Technologies) or the Twist Human Core Exome Enrichment System (Twist Bioscience). Followed by 100 base pair paired-end sequencing using an Illumina NovaSeq platform. Data analysis was performed with the IntegraGen in house bioinformatic pipeline. Sequence reads were aligned to the human genome reference GRCh38/hg38 using BWA. Variant calling for the identification of single nucleotide variations, and small insertions/deletions were performed *via* the Broad Institute's GATK Haplotype Caller GVCF tool (GATK 3.8.1). Ensembl's VEP (Variant Effect Predictor, release VEP 95.1) program was used to process variants for further annotations. Allele frequency annotations was based on gnomAD (version 2.1.1) and 1000Genomes datasets. PolyPhen-2 algorithm was used to predict deleteriousness of missense variants. Finally, we used CANOES for the detection of copy number variation in whole exome sequencing data (80).

Qualifying variants in the PLOD3 gene

Several criteria were used for the selection of qualifying variants in the *PLOD3* gene: (1) quality control: a minimum of 10 reads was required for depth coverage and a minimum of 25% for allelic balance, (2) variants' nature: all coding and splice site variants were retained, (3) minor allele frequency ≤1/1000 in external databases (gnomAD, version 2.1.1 and 1000Genomes phase 3), and (4) for missense variants, predicted possibly or probably damaging with PolyPhen-2. Investigation for variant enrichment in the *PLOD3* gene was performed by comparing our fetus cohort (N = 113) to cohorts

of control individuals in gnomAD v2.1 and v3.1 with the same filter criteria. Burden tests/p values were used for statistical analyses.

Mapping of the amino acid residues in protein structure

The human LH3 crystal structure was retrieved from the Research Collaboratory for Structural Bioinformatics Protein Data Bank (www.rcsb.org, accession number: 6FXK) and used for variant modeling using the UCSF Chimera software (version 1.14; build 42094).

Statistical analyses

For comparisons between two groups, one-way ANOVA was performed to determine whether differences between groups are significant using OriginPro, version 9.1 (OriginLab Corp). A p value of less than 0.05 was considered statistically significant.

Data availability

All data are contained within the article and supporting information. All source data are available from the corresponding author upon reasonable request. The MS datasets for protein identification of collagen α1α1α2(IV) samples have been deposited to the ProteomeXchange consortium *via* the jPOST partner repository with the dataset identifier PXD035051 (http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD035051).

(References: https://www.nature.com/articles/nbt.2839 and https://academic.oup.com/nar/article/45/D1/D1107/2605695?login=false).

Supporting information—This article contains supporting information.

Acknowledgments—We gratefully acknowledge the research department of Shriners Hospitals for Children in Portland for using CD instrument. We thank Hans Peter Bächinger and Cassandre Labelle-Dumais for useful discussions and suggestions on this work. Research reported in this publication was supported in part by the UCSF Vision Core shared resource of the National Institutes of Health/National Eye Institute P30 EY002162 and by an unrestricted grant from Research to Prevent Blindness, New York, NY.

Author contributions—Y. I., K. M., E. T.-L., and D. B. G. conceptualization; Y. I., Y. T., S. F. T., and D. R. K. methodology; Y. I., Y. T., T. C., S. F. T., and D. R. K. validation; Y. I., Y. T., T. C., S. F. T., and D. R. K. formal analysis; Y. I., Y. T., T. C., S. F. T., and D. R. K. investigation; Y. I. and T. C. resources; Y. I., Y. T., T. C., S. F. T., D. R. K., K. M., E. T.-L., and D. B. G. data curation; Y. I. and D. B. G. writing–original draft; Y. I., Y. T., T. C., S. F. T., D. R. K., K. M., E. T.-L., and D. B. G. writing–review & editing; Y. I. visualization; Y. I., E. T.-L., and D. B. G. supervision; Y. I., E. T.-L., and D. B. G. project administration; Y. I. and D. B. G. funding acquisition.

Funding and additional information—Research reported in this publication was supported by All May See Foundation award 7031182 (to Y. I.) and the National Institutes of Health under award number R01NS096173 (to D. B. G.). The content of this publication is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or other funding agencies.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: AAA, amino acid analyses; AHA, L-azidohomoalanine; DMEM, Dulbecco's modified Eagle's medium; EDS, Ehlers–Danlos syndrome; ER, endoplasmic reticulum; gRNA, guide RNA; HSP47, heat shock protein 47; Hyl, hydroxylysine; 4Hyp, 4-hydroxyproline; ICH, intracerebral hemorrhage; LH3, lysyl hydroxylase 3; MEF, mouse embryonic fibroblast; MS, mass spectrometry; OI, osteogenesis imperfecta; P4H, prolyl 4-hydroxylase; PTM, post-translational modification; PVDF, polyvinylidene difluoride; SI-collagen, stable isotope–labeled collagen; SPR, surface plasmon resonance.

References

1. Bella, J., and Hulmes, D. J. (2017) Fibrillar collagens. *Subcell Biochem.* **82**, 457–490
2. Ricard-Blum, S., and Ruggiero, F. (2005) The collagen superfamily: from the extracellular matrix to the cell membrane. *Pathol. Biol.* **53**, 430–442
3. Lamande, S. R., and Bateman, J. F. (2020) Genetic disorders of the extracellular matrix. *Anat. Rec. (Hoboken)* **303**, 1527–1542
4. Bateman, J. F., Boot-Handford, R. P., and Lamande, S. R. (2009) Genetic diseases of connective tissues: cellular and extracellular effects of ECM mutations. *Nat. Rev. Genet.* **10**, 173–183
5. Marini, J. C., Forlino, A., Bachinger, H. P., Bishop, N. J., Byers, P. H., Paepe, A., *et al.* (2017) Osteogenesis imperfecta. *Nat. Rev. Dis. Primers* **3**, 17052
6. Malfait, F., Castori, M., Francomano, C. A., Giunta, C., Kosho, T., and Byers, P. H. (2020) The Ehlers-Danlos syndromes. *Nat. Rev. Dis. Primers* **6**, 64
7. Besio, R., Chow, C. W., Tonelli, F., Marini, J. C., and Forlino, A. (2019) Bone biology: insights from osteogenesis imperfecta and related rare fragility syndromes. *FEBS J.* **286**, 3033–3056
8. Claeys, L., Storoni, S., Eekhoff, M., Elting, M., Wisse, L., Pals, G., *et al.* (2021) Collagen transport and related pathways in Osteogenesis Imperfecta. *Hum. Genet.* **140**, 1121–1141
9. Omar, R., Malfait, F., and Van Agtmael, T. (2021) Four decades in the making: collagen III and mechanisms of vascular Ehlers danlos syndrome. *Matrix Biol. Plus* **12**, 100090
10. Van Damme, T., Colman, M., Syx, D., and Malfait, F. (2022) The ehlers-danlos syndromes against the backdrop of inborn errors of metabolism. *Genes (Basel)* **13**, 265
11. Hudson, B. G., Tryggvason, K., Sundaramoorthy, M., and Neilson, E. G. (2003) Alport's syndrome, Goodpasture's syndrome, and type IV collagen. *N. Engl. J. Med.* **348**, 2543–2556
12. Boudko, S. P., Pokidysheva, E., and Hudson, B. G. (2022) Prospective collagen IValpha345 therapies for Alport syndrome. *Curr. Opin. Nephrol. Hypertens.* **31**, 213–220
13. Jeanne, M., and Gould, D. B. (2017) Genotype-phenotype correlations in pathology caused by collagen type IV alpha 1 and 2 mutations. *Matrix Biol.* **57-58**, 29–44
14. Mao, M., Alavi, M. V., Labelle-Dumais, C., and Gould, D. B. (2015) Type IV collagens and basement membrane diseases: cell biology and pathogenic mechanisms. *Curr. Top Membr.* **76**, 61–116
15. Kuo, D. S., Labelle-Dumais, C., and Gould, D. B. (2012) COL4A1 and COL4A2 mutations and disease: Insights into pathogenic mechanisms and potential therapeutic targets. *Hum. Mol. Genet.* **21**, R97–110
16. Boyce, D., McGee, S., Shank, L., Pathak, S., and Gould, D. (2021) Epilepsy and related challenges in children with COL4A1 and COL4A2 mutations: a Gould syndrome patient registry. *Epilepsy Behav.* **125**, 108365

17. Mao, M., Kiss, M., Ou, Y., and Gould, D. B. (2017) Genetic dissection of anterior segment dysgenesis caused by a Col4a1 mutation in mouse. *Dis. Model Mech.* **10**, 475–485
18. Mao, M., Smith, R. S., Alavi, M. V., Marchant, J. K., Cosma, M., Libby, R. T., *et al.* (2015) Strain-dependent anterior segment dysgenesis and progression to glaucoma in Col4a1 mutant mice. *Invest. Ophthalmol. Vis. Sci.* **56**, 6823–6831
19. Ishikawa, Y., and Bachinger, H. P. (2013) A molecular ensemble in the rER for procollagen maturation. *Biochim. Biophys. Acta* **1833**, 2479–2491
20. Ito, S., and Nagata, K. (2021) Quality control of procollagen in cells. *Annu. Rev. Biochem.* **90**, 631–658
21. Onursal, C., Dick, E., Angelidis, I., Schiller, H. B., and Staab-Weijnitz, C. A. (2021) Collagen biosynthesis, processing, and maturation in lung ageing. *Front. Med. (Lausanne)* **8**, 593874
22. Ihme, A., Krieg, T., Nerlich, A., Feldmann, U., Rauterberg, J., Glanville, R. W., *et al.* (1984) Ehlers-danlos syndrome type VI: collagen type specificity of defective lysyl hydroxylation in various tissues. *J. Invest. Dermatol.* **83**, 161–165
23. Weis, M. A., Hudson, D. M., Kim, L., Scott, M., Wu, J. J., and Eyre, D. R. (2010) Location of 3-hydroxyproline residues in collagen types I, II, III, and V/XI implies a role in fibril supramolecular assembly. *J. Biol. Chem.* **285**, 2580–2590
24. Ishikawa, Y., Taga, Y., Zientek, K., Mizuno, N., Salo, A. M., Semenova, O., *et al.* (2021) Type I and type V procollagen triple helix uses different subsets of the molecular ensemble for lysine posttranslational modifications in the rER. *J. Biol. Chem.* **296**, 100453
25. Ishikawa, Y., and Bächinger, H. P. (2014) A substrate preference for the rough endoplasmic reticulum resident protein FKBP22 during collagen biosynthesis. *J. Biol. Chem.* **289**, 18189–18201
26. Taga, Y., Kusubata, M., Ogawa-Goto, K., and Hattori, S. (2016) Developmental stage-dependent regulation of prolyl 3-hydroxylation in tendon type I collagen. *J. Biol. Chem.* **291**, 837–847
27. Salo, A. M., and Myllyharju, J. (2021) Prolyl and lysyl hydroxylases in collagen synthesis. *Exp. Dermatol.* **30**, 38–49
28. Basak, T., Vega-Montoto, L., Zimmerman, L. J., Tabb, D. L., Hudson, B. G., and Vanacore, R. M. (2016) Comprehensive characterization of glycosylation and hydroxylation of basement membrane collagen IV by high-resolution mass spectrometry. *J. Proteome Res.* **15**, 245–258
29. Berg, R. A., and Prockop, D. J. (1973) The thermal transition of a non-hydroxylated form of collagen. Evidence for a role for hydroxyproline in stabilizing the triple-helix of collagen. *Biochem. Biophys. Res. Commun.* **52**, 115–120
30. Mizuno, K., Hayashi, T., Peyton, D. H., and Bachinger, H. P. (2004) Hydroxylation-induced stabilization of the collagen triple helix. Acetyl-(glycyl-4(R)-hydroxyprolyl-4(R)-hydroxyprolyl)(10)-NH(2) forms a highly stable triple helix. *J. Biol. Chem.* **279**, 38072–38078
31. Taga, Y., Tanaka, K., Hattori, S., and Mizuno, K. (2021) In-depth correlation analysis demonstrates that 4-hydroxyproline at the Yaa position of Gly-Xaa-Yaa repeats dominantly stabilizes collagen triple helix. *Matrix Biol. Plus* **10**, 100067
32. Montgomery, N. T., Zientek, K. D., Pokidysheva, E. N., and Bachinger, H. P. (2018) Post-translational modification of type IV collagen with 3-hydroxyproline affects its interactions with glycoprotein VI and nidogens 1 and 2. *J. Biol. Chem.* **293**, 5987–5999
33. Grafe, I., Yang, T., Alexander, S., Homan, E. P., Lietman, C., Jiang, M. M., *et al.* (2014) Excessive transforming growth factor-beta signaling is a common mechanism in osteogenesis imperfecta. *Nat. Med.* **20**, 670–675
34. Hennet, T. (2019) Collagen glycosylation. *Curr. Opin. Struct. Biol.* **56**, 131–138
35. De Giorgi, F., Fumagalli, M., Scietti, L., and Forneris, F. (2021) Collagen hydroxylysine glycosylation: non-conventional substrates for atypical glycosyltransferase enzymes. *Biochem. Soc. Trans.* **49**, 855–866
36. Geister, K. A., Lopez-Jimenez, A. J., Houghtaling, S., Ho, T. H., Vanacore, R., and Beier, D. R. (2019) Loss of function of Colgalt1 disrupts collagen post-translational modification and causes musculoskeletal defects. *Dis. Model Mech.* **12**, dmm037176.
37. Miyatake, S., Schneeberger, S., Koyama, N., Yokochi, K., Ohmura, K., Shiina, M., *et al.* (2018) Biallelic COLGALT1 variants are associated with cerebral small vessel disease. *Ann. Neurol.* **84**, 843–853
38. Teunissen, M. W. A., Kamsteeg, E. J., Sallevelt, S., Pennings, M., Bauer, N. J. C., Vermeulen, R. J., *et al.* (2021) Biallelic variants in the COLGALT1 gene causes severe congenital porencephaly: a case report. *Neurol. Genet.* **7**, e564
39. Salo, A. M., Cox, H., Farndon, P., Moss, C., Grindulis, H., Risteli, M., *et al.* (2008) A connective tissue disorder caused by mutations of the lysyl hydroxylase 3 gene. *Am. J. Hum. Genet.* **83**, 495–503
40. Ewans, L. J., Colley, A., Gaston-Massuet, C., Gualtieri, A., Cowley, M. J., McCabe, M. J., *et al.* (2019) Pathogenic variants in PLOD3 result in a Stickler syndrome-like connective tissue disorder with vascular complications. *J. Med. Genet.* **56**, 629–638
41. Vahidnezhad, H., Youssefian, L., Saeidian, A. H., Touati, A., Pajouhanfar, S., Baghdadi, T., *et al.* (2019) Mutations in PLOD3, encoding lysyl hydroxylase 3, cause a complex connective tissue disorder including recessive dystrophic epidermolysis bullosa-like blistering phenotype with abnormal anchoring fibrils and type VII collagen deficiency. *Matrix Biol.* **81**, 91–106
42. Maddirevula, S., Alzahrani, F., Al-Owain, M., Al Muhaizea, M. A., Kayyali, H. R., AlHashem, A., *et al.* (2019) Autozygome and high throughput confirmation of disease genes candidacy. *Genet. Med.* **21**, 736–742
43. Zhou, J., Feng, W., Zhuo, X., Lu, W., Wang, J., Fang, F., *et al.* (2022) Cerebral small vessel disease caused by PLOD3 mutation: expanding the phenotypic spectrum of lysyl hydroxylase-3 deficiency. *Pediatric Investig.* **6**, 219–223
44. Ruotsalainen, H., Sipila, L., Vapola, M., Sormunen, R., Salo, A. M., Uitto, L., *et al.* (2006) Glycosylation catalyzed by lysyl hydroxylase 3 is essential for basement membranes. *J. Cell Sci.* **119**, 625–635
45. Rautavuoma, K., Takaluoma, K., Sormunen, R., Myllyharju, J., Kivirikko, K. I., and Soininen, R. (2004) Premature aggregation of type IV collagen and early lethality in lysyl hydroxylase 3 null mice. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 14120–14125
46. Sipila, L., Ruotsalainen, H., Sormunen, R., Baker, N. L., Lamande, S. R., Vapola, M., *et al.* (2007) Secretion and assembly of type IV and VI collagens depend on glycosylation of hydroxylysines. *J. Biol. Chem.* **282**, 33381–33388
47. Heikkinen, J., Risteli, M., Wang, C., Latvala, J., Rossi, M., Valtavaara, M., *et al.* (2000) Lysyl hydroxylase 3 is a multifunctional protein possessing collagen glucosyltransferase activity. *J. Biol. Chem.* **275**, 36158–36163
48. Rautavuoma, K., Takaluoma, K., Passoja, K., Pirskanen, A., Kvist, A. P., Kivirikko, K. I., *et al.* (2002) Characterization of three fragments that constitute the monomers of the human lysyl hydroxylase isoenzymes 1-3. The 30-kDa N-terminal fragment is not required for lysyl hydroxylase activity. *J. Biol. Chem.* **277**, 23084–23091
49. Wang, C., Luosujarvi, H., Heikkinen, J., Risteli, M., Uitto, L., and Myllyla, R. (2002) The third activity for lysyl hydroxylase 3: Galactosylation of hydroxylysyl residues in collagens *in vitro*. *Matrix Biol.* **21**, 559–566
50. Scietti, L., Chiapparino, A., De Giorgi, F., Fumagalli, M., Khoriauli, L., Nergadze, S., *et al.* (2018) Molecular architecture of the multifunctional collagen lysyl hydroxylase and glycosyltransferase LH3. *Nat. Commun.* **9**, 3163
51. Guo, H. F., Bota-Rabassedas, N., Terajima, M., Leticia Rodriguez, B., Gibbons, D. L., Chen, Y., *et al.* (2021) A collagen glucosyltransferase drives lung adenocarcinoma progression in mice. *Commun. Biol.* **4**, 482
52. Duncan, K. G., Fessler, L. I., Bachinger, H. P., and Fessler, J. H. (1983) Procollagen IV. Association to tetramers. *J. Biol. Chem.* **258**, 5869–5877
53. Mirigian, L. S., Makareeva, E., and Leikin, S. (2014) Pulse-chase analysis of procollagen biosynthesis by azidohomoalanine labeling. *Connect Tissue Res.* **55**, 403–410
54. Syx, D., Ishikawa, Y., Gebauer, J., Boudko, S. P., Guillemyn, B., Van Damme, T., *et al.* (2021) Aberrant binding of mutant HSP47 affects posttranslational modification of type I collagen and leads to osteogenesis imperfecta. *PLoS Genet.* **17**, e1009339
55. Fessler, L. I., and Fessler, J. H. (1982) Identification of the carboxyl peptides of mouse procollagen IV and its implications for the assembly and structure of basement membrane procollagen. *J. Biol. Chem.* **257**, 9804–9810

56. Lunstrum, G. P., Bachinger, H. P., Fessler, L. I., Duncan, K. G., Nelson, R. E., and Fessler, J. H. (1988) Drosophila basement membrane procollagen IV. I. Protein characterization and distribution. *J. Biol. Chem.* **263**, 18318–18327
57. Toth, M., Sado, Y., Ninomiya, Y., and Fridman, R. (1999) Biosynthesis of alpha2(IV) and alpha1(IV) chains of collagen IV and interactions with matrix metalloproteinase-9. *J. Cell Physiol.* **180**, 131–139
58. Matsuoka, Y., Kubota, H., Adachi, E., Nagai, N., Marutani, T., Hosokawa, N., *et al.* (2004) Insufficient folding of type IV collagen and formation of abnormal basement membrane-like structure in embryoid bodies derived from Hsp47-null embryonic stem cells. *Mol. Biol. Cell* **15**, 4467–4475
59. Ishida, Y., Kubota, H., Yamamoto, A., Kitamura, A., Bachinger, H. P., and Nagata, K. (2006) Type I collagen in Hsp47-null cells is aggregated in endoplasmic reticulum and deficient in N-propeptide processing and fibrillogenesis. *Mol. Biol. Cell* **17**, 2346–2355
60. Vranka, J. A., Pokidysheva, E., Hayashi, L., Zientek, K., Mizuno, K., Ishikawa, Y., *et al.* (2010) Prolyl 3-hydroxylase 1 null mice display abnormalities in fibrillar collagen-rich tissues such as tendons, skin, and bones. *J. Biol. Chem.* **285**, 17253–17262
61. Pedchenko, V., Bauer, R., Pokidysheva, E. N., Al-Shaer, A., Forde, N. R., Fidler, A. L., *et al.* (2019) A chloride ring is an ancient evolutionary innovation mediating the assembly of the collagen IV scaffold of basement membranes. *J. Biol. Chem.* **294**, 7968–7981
62. Ivanov, S. V., Bauer, R., Pokidysheva, E. N., and Boudko, S. P. (2021) Collagen IV exploits a Cl- step gradient for scaffold assembly. *Adv. Exp. Med. Biol.* **21**, 129–141
63. Bhatnagar, R. S., and Gough, C. A. (1996) Circular dichroism of collagen and related polypeptides. In: Fasman, G. D., ed. *Circular Dichroism and the Conformational Analysis of Biomolecules*, Springer US, Boston, MA: 183–199
64. Ono, T., Miyazaki, T., Ishida, Y., Uehata, M., and Nagata, K. (2012) Direct *in vitro* and *in vivo* evidence for interaction between Hsp47 protein and collagen triple helix. *J. Biol. Chem.* **287**, 6810–6818
65. Ishikawa, Y., Ito, S., Nagata, K., Sakai, L. Y., and Bachinger, H. P. (2016) Intracellular mechanisms of molecular recognition and sorting for transport of large extracellular matrix molecules. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E6036–E6044
66. Sricholpech, M., Perdivara, I., Yokoyama, M., Nagaoka, H., Terajima, M., Tomer, K. B., *et al.* (2012) Lysyl hydroxylase 3-mediated glucosylation in type I collagen: molecular loci and biological significance. *J. Biol. Chem.* **287**, 22998–23009
67. Gould, D. B., Phalan, F. C., Breedveld, G. J., van Mil, S. E., Smith, R. S., Schimenti, J. C., *et al.* (2005) Mutations in Col4a1 cause perinatal cerebral hemorrhage and porencephaly. *Science* **308**, 1167–1171
68. Gould, D. B., Phalan, F. C., van Mil, S. E., Sundberg, J. P., Vahedi, K., Massin, P., *et al.* (2006) Role of COL4A1 in small-vessel disease and hemorrhagic stroke. *N. Engl. J. Med.* **354**, 1489–1496
69. Breedveld, G., de Coo, I. F., Lequin, M. H., Arts, W. F., Heutink, P., Gould, D. B., *et al.* (2006) Novel mutations in three families confirm a major role of COL4A1 in hereditary porencephaly. *J. Med. Genet.* **43**, 490–495
70. Vahedi, K., Boukobza, M., Massin, P., Gould, D. B., Tournier-Lasserve, E., and Bousser, M. G. (2007) Clinical and brain MRI follow-up study of a family with COL4A1 mutation. *Neurology* **69**, 1564–1568
71. Coste, T., Vincent-Delorme, C., Stichelbout, M., Devisme, L., Gelot, A., Deryabin, I., *et al.* (2022) COL4A1/COL4A2 and inherited platelet disorder gene variants in fetuses showing intracranial hemorrhage. *Prenat Diagn.* **42**, 601–610
72. Aro, E., Salo, A. M., Khatri, R., Finnila, M., Miinalainen, I., Sormunen, R., *et al.* (2015) Severe extracellular matrix abnormalities and chondrodysplasia in mice lacking collagen prolyl 4-hydroxylase isoenzyme II in combination with a reduced amount of isoenzyme I. *J. Biol. Chem.* **290**, 16964–16978
73. Zou, Y., Donkervoort, S., Salo, A. M., Foley, A. R., Barnes, A. M., Hu, Y., *et al.* (2017) P4HA1 mutations cause a unique congenital disorder of connective tissue involving tendon, bone, muscle and the eye. *Hum. Mol. Genet.* **26**, 2207–2217
74. Al-Shaer, A., Lyons, A., Ishikawa, Y., Hudson, B. G., Boudko, S. P., and Forde, N. R. (2021) Sequence-dependent mechanics of collagen reflect its structural and functional organization. *Biophys. J.* **120**, 4013–4028
75. Taga, Y., Kusubata, M., Ogawa-Goto, K., and Hattori, S. (2012) Development of a novel method for analyzing collagen O-glycosylations by hydrazide chemistry. *Mol. Cell Proteomics* **11**. https://doi.org/10.1074/mcp.M111.010397. M111 010397
76. Morris, N. P., Keene, D. R., Glanville, R. W., Bentz, H., and Burgeson, R. E. (1986) The tissue form of type VII collagen is an antiparallel dimer. *J. Biol. Chem.* **261**, 5638–5644
77. Macdonald, J. R., and Bachinger, H. P. (2001) HSP47 binds cooperatively to triple helical type I collagen but has little effect on the thermal stability or rate of refolding. *J. Biol. Chem.* **276**, 25399–25403
78. Fujii, K. K., Taga, Y., Sakai, T., Ito, S., Hattori, S., Nagata, K., *et al.* (2019) Lowering the culture temperature corrects collagen abnormalities caused by HSP47 gene knockout. *Sci. Rep.* **9**, 17433
79. Taga, Y., Kusubata, M., Ogawa-Goto, K., and Hattori, S. (2014) Stable isotope-labeled collagen: A novel and versatile tool for quantitative collagen analyses using mass spectrometry. *J. Proteome Res.* **13**, 3671–3678
80. Backenroth, D., Homsy, J., Murillo, L. R., Glessner, J., Lin, E., Brueckner, M., *et al.* (2014) CANOES: detecting rare copy number variants from whole exome sequencing data. *Nucleic Acids Res.* **42**, e97

Prenatal Diagnosis of *COL4A1* Mutations in Eight Cases: Further Delineation of the Neurohistopathological Phenotype

Pediatric and Developmental Pathology
2022, Vol. 0(0) 1–12

DOI: 10.1177/10935266221080134
journals.sagepub.com/home/pdp
SAGE

Francesca Gubana[1,2], Christo Christov[3], Thibault Coste[4], Elisabeth Tournier-Lasserve[4], Alexandra Benachi[2], Catherine Fallet-Bianco[5], Ferechte Encha-Razavi[1], and Jelena Martinovic[1]

Abstract

Background: Increasing number of mutations responsible for vascular lesions, leading to ischemic or hemorrhagic stroke in young adults, has been identified in the recent years. It has been demonstrated in both mice and humans, that mutations in *COL4A1* gene promote cerebral hemorrhages. In humans, both adults and children may be affected, and the spectrum has been broadened recently to neonates and fetuses.

Methods: We present a cohort of eight *COL4A1* mutated fetuses in which cerebral hemorrhages were detected by ultrasound leading to elective terminations of pregnancy.

Results: Our neuropathological studies demonstrated a strikingly similar pathological pattern, dominated by supra- and infratentorial multifocal hemorrhagic lesions of various abundance and age in the vicinity of enlarged small vessels having a discontinuous wall. This was constantly associated with a spectrum of supratentorial post-ischemic damages of the grey and white matters. Morphometric studies of brain vessels confirmed vascular dilation and hypervascularization in both grey and white matters and severe attenuation of the smooth-muscle actin staining in the white matter.

Conclusion: These observations add to the rare human neuropathological phenotype of *COL4A1* mutations. Its recognition is mandatory to enhance the number of tested patients in the future, as well as the genetic counseling of parents.

Keywords
COL4A1, prenatal diagnosis, ultrasonography, fetus, cerebral hemorrhage, porencephaly, autosomal dominant inheritance, genetic counseling

Introduction

The *COL4A1* gene encodes for the α1 type IV collagen chain. Type IV collagen molecules attach to each other to form a complex protein network that is the main component of basement membranes in the human body; particularly in the basement membranes of blood vessels. The *COL4A1* protein plays an important role in the inhibition of endothelial cell proliferation, migration, tube formation, and neovascularization.[1]

Mutations in the *COL4A1* gene cause vascular vulnerability that is responsible for different manifestations associated with small vessel disease and hemorrhagic stroke, including retinal arterial tortuosity, cataract, Raynaud phenomenon, diffuse periventricular leukoencephalopathy, and hereditary angiopathy with nephropathy, aneurysm, hemolytic anemia, and muscle cramps (HANAC syndrome, OMIM #611773). Interestingly, *COL4A1* homozygous mutations are lethal in mice, while heterozygotes show brain malformations, myopathy, and ocular dysgenesis.[2–4] The physiopathological

[1]Unit of Embryo-Fetal Pathology, AP-HP, Antoine Béclère Hospital, Clamart, France
[2]Department of Obstetrics and Gynecology, AP-HP, Antoine Béclère Hospital, Paris Saclay University, Clamart, France
[3]Department of Histology, CHRU, INSERM U1256, NGERE, Nancy, France
[4]Department of Neurovascular Genetics, AP-HP, St Louis Hospital, Paris, France
[5]Department of Pediatric Pathology, Saint Justine Hospital, Montreal, Canada

Corresponding Author:
Jelena Martinovic, Unit of Embryo-Fetal Pathology, AP-HP, Antoine Beclere University Hospital, 157 rue de la Porte de Trivaux, Clamart 92141, France.
Email: jelena.martinovic@aphp.fr

process underlying this phenotypic variability begins to be understood.[3-8]

Consistent with the large distribution of α1 (IV) and α2 (IV) heterotrimers in nearly all basement membranes, *COL4A1* and *COL4A2* mutations cause multisystem disorders postnatally.[7-9] Cerebrovascular disease is one of the most prominent consequences of *COL4A1* and *COL4A2* mutations and comprises a constellation of clinical manifestations.

In accordance with the medical literature, the most frequent clinical signs of *COL4A1* mutation are the following:

1. Prenatal and neonatal intracranial hemorrhage and porencephaly. Porencephaly is often caused by germinal layer hemorrhage leading to deep venous infarction with subsequent tissue necrosis and cavitation.[10-13]
2. Periventricular leukomalacia (PVL) with intracranial calcifications that is due to a hypoxic-ischemic insult. PVL is characterized by focal periventricular necrosis and gliosis in the surrounding white matter. Intracranial calcifications have also been reported in this context of hypoxic damage.[13-15]
3. Hydranencephaly is the result of massive hemisphere necrosis and huge ventricular dilatation, the hemispheres are replaced by cerebrospinal fluid, but the posterior fossa and the diencephalic brain structures are preserved.[16]
4. Schizencephaly is a congenital anomaly characterized by an abnormal gray matter-lined defect extending from the pial surface to the lateral ventricles.[2,11,13]
5. Extracranial manifestations like nephropathy, ocular dysgenesis, muscle abnormalities, and hemolytic anemia have also been documented.[3,12,13,15,17]

Literature evidence about *COL4A1* mutation manifestations before birth is growing.[9,18-23] In 2012, Lichtenbelt et al. published the first case report of a prenatal *COL4A1* mutation in a fetus of 21,6 weeks with ultrasound findings of cerebral hemorrhage.[21] Maurice et al[22] published a retrospective study aiming to evaluate the prevalence of *COL4A1* and *COL4A2* mutations in fetuses presenting a characteristic pattern of cerebral injury including multifocal supra- and infratentorial and ischemic-hemorrhagic lesions associated with schizencephaly or porencephaly. Recently, Schannon et al[23] studied a series of 4 *COL4A1* mutated fetuses and their placentas correlating the vascular placental malperfusion lesions with cerebral involvement.

The estimated general incidence of fetal intracranial hemorrhage is about 1 in 10 000 pregnancies.[8] It is associated with postnatal seizure disorders, mental retardation, psychomotor delay, and cerebral palsy.[9,10]

We report herein a series of 8 fetuses affected by a severe cerebral hemorrhage related to a *COL4A1* mutation. Detailed prenatal and neuropathological data are provided. In addition, we present, for the first time, morphometric estimations of cerebral vessels in *COL4A1* mutated and wild-type fetal brains.

Patients and Methods

The eight pregnancies were referred due to cerebral anomalies found on ultrasound (US) scan. Fetal karyotypes and array-CGH were normal in all cases. Maternal risk factors for fetal brain hemorrhage such as platelet or coagulation disorders, seizures, trauma, viral or bacterial infections, medications, or drugs were excluded. US examinations were performed abdominally, and vaginally when needed, using a multiplanar approach. Three fetuses had in addition a prenatal brain magnetic resonance imaging (MRI) (Table 1).

After medical terminations at various ages, all fetuses underwent unrestricted visceral and neuropathological postmortem examination, with parental informed consent, as required by the French law. Fetal examinations were conducted in the Fetal Pathology Unit of Antoine Beclere Hospital (AP-HP). Brain in situ pre-fixation was performed by transfontanellar injection of zinc-formalin solution (4% formaldehyde), 24h prior to fetal autopsy, under our standardized protocol.[24] Macroscopical evaluation was done after 4 weeks immersion in 4% zinc-formalin solution. Cerebral hemispheres were cut coronally. Whole mount and small sections were taken from anterior to posterior, at the frontoparietal, thalamus–hippocampal, and occipital levels and at salient pathological areas. The cerebellum was first observed on sagittal and then on axial sections. Specimens were taken from mid sagittal and from axial sections from anterior to posterior. They were processed for histological study following routine processes, cut at 7 microns, and stained with Hematoxylin-Eosin (H&E). Immunohistochemical staining on selected blocs was assessed by CD34 (Monoclonal Mouse Anti-Human CD34 Class II antibody, Dako, Glostrup, Danemark) and SMA (Monoclonal Mouse Anti-Actin, α-Smooth Muscle antibody, Sigma-Aldrich, St Louis, USA). The detailed feto- and neuropathological data for each fetus are presented in Table 2.

All eight fetuses were tested heterozygous for pathogenic *COL4A1* gene mutations after fetopathological examination. Molecular analyses were performed in the Molecular Genetics Laboratories of Lariboisière and Saint Louis Hospitals (AP-HP) in Paris. Sanger sequencing of *COL4A1* and *COL4A2* genes was performed using standardized procedures on DNA extracted from fetal tissues. To fully establish *de novo* nature of sequence variations, the DNAs of the fetus and these of both parents were tested using nine sets of microsatellites using the AmpFlSTR Profiler PCR Amplification Kit (Applied Biosystems). The reaction was carried out according to the manufacturer's instructions, loaded on an ABI 3130 genetic analyzer (Applied Biosystems) and the fragments were analyzed using GeneMapper v4.0 software (Applied Biosystems). Variants were classified according to the American College of Medical Geneticists and Genomics (Table 3).

Table 1. Clinical and ultrasound findings in the cohort.

Case	MA	Family history	G/P	GA at Diagnosis	US findings
1	32	Aneurysm in a brother	2/2	22,5	Bilateral hyperechogenic foci in frontal cerebral parenchyma, Enlarged and irregular Sylvian fissure
2	26	UR	3/2	18,3	Bilateral ventriculomegaly (20 mm) Non-visualized posterior fossa CHM with ventricular septal defect Hands' malposition
3	37	UR	3/1	24,5	Cerebral biometry <5th p Multiple calcifications in the thalami, frontal, and lateral ventricular walls suspecting hemorrhagic event Hepatosplenomegaly
4	29	UR	3/1	22,3	Thick corpus callosum, increased thickness of the cerebral ventricular walls, multiple hyperechogenic millimetric foci in the parenchyma Evolution toward shizencephaly
5	37	UR	2/0 (1 IUFD at 19w)	22	Schizencephaly in right hemisphere Large hyperechogenic mass in left ventricule
6	38	UR	3/0	18	Large ischemic- hemorrhagic lesion in right frontal area of cerebral parenchyma
7	40	UR	4/0	22	Tetra-ventricular post-hemorrhagic hydrocephaly (LV: 15mm), increased thickness of ventricular walls
8	31	UR	1/0	39	IUGR<5th p from 32 w Bilateral hydrocephaly (15 mm) Hyperechogenic periventricular foci within the brain parenchyma

Abbreviations: *MA* maternal age, *G/P* gestity/parity, *GA* gestational age in weeks, *US* ultrasound, *IUFD* in-utero fetal demise, *LV* lateral ventricle, *IUGR* in-utero growth restriction, UR unremarkable.

Brain vessels from 3 *COL4A1 mutated* (GA: 23–38 weeks) and 4 control (GA: 23–38 weeks) cases with normal brain examination were studied by morphometry. Images from tissue sections immunostained for CD34 were acquired on a BX63 microscope equipped with a DP70 color camera (both from Olympus, France). Four to six x200 images of microscopically normal (devoid of hemorrhage) brain tissue were acquired in grey matter (GM) and white matter (WM) corresponding to 248 ± 86 (mean ± SD) vessels per case (*COL4A1* GM), 93 ± 58 (control, GM), 137 ± 13 (*COL4A1*, WM), and 95 ± 36 (control, WM). Vessel profiles were segmented by texture, color, and edge-detection algorithms in Ilastik 1.3.2, and[25] measurements were carried out in KS400.3 (Zeiss, France). Measured parameters included the following: (i) number of vessels per microscopic field (V/MF); (ii) percentage of microscopic field area occupied by vessel area (VA/MFA%), and (iii) vessel area per unit of vessel length (VA/VL); this latter parameter assessing vessel dilatation was derived by drawing a 1 µm thick centerline through the already segmented vessel profiles using Angiotool[26] and then dividing the measured vessel profile area by its length (Supplemental figure 1). Brain specimens (GM and WM) from 2 fetuses (28 and 40 weeks) *Col4A1* and their age-matched controls were examined by morphometry after Smooth-Muscle-Actin (α-SMA) immunostaining. Some 10–15 vessels per case at each of the two locations were examined by two quantitative parameters: (i) α-SMA-positive area (media) as percentage of whole vessel area; (ii) length of "gaps" in media as percentage of whole vessel perimeter (Supplemental figure 2). In all cases, vessels with the most abundant α-SMA immunostaining signal were selected for measurement.

Statistical comparisons were performed by Student's t-test or Mann and Whitney's test in SAS 9.2. For each parameter, first the two groups were compared and then 3 pairwise comparisons between a *COL4A1* mutated case and available age-matched control cases were carried out.

Results

The first trimester US and routine prenatal blood results of all pregnancies were unremarkable.

In the majority (7/8) of cases, the second trimester US was abnormal, while in one case only the 3rd trimester US performed at 32 weeks revealed in-utero growth restriction (IUGR), and in a control case, the 35 weeks scan showed bilateral hydrocephaly with hyperechogenic periventricular foci. Clinical data and prenatal US findings of all 8 cases are presented in Table 1.

Brain Imaging

The most frequent brain US findings during routine and referral examination were poor cephalic biometry, mild or severe ventriculomegaly with increased echogenicity and irregularity of the ventricular walls (Figure 1), hyperechogenic areas or calcifications in the brain parenchyma, and cavitated lesions. Among the 8 cases, there are two with additional extra-cerebral abnormalities (Cases 2 and 7).

Table 2. Post-mortem examination findings.

Case	GA	Sex	External, visceral, or placental abnormal findings	Neuropathology gross and microscopic findings
1	27,5	F	No	Cerebral biometry at 5th percentile Left schizencephaly Diffuse ependymal and subependymal hypoxic/ischemic lesions
2	19,1	M	Facial dysmorphy: Macrocrania, hypertelorism; Generalized edema/bilateral pleural effusions with pulmonary hypoplasia; Anal imperforation/Low rectal atresia Hepatomegaly	Tetraventricular hydrocephaly Porencephaly of right insula Diffuse petechial cortical hemorrhagies Dandy Walker-like malformation
3	25,4	F	No	Biometry <5th percentile Bi-ventricular post-hemorrhagic hydrocephaly with massive periventricular calcifications Several ischemic and hemorrhagic lesions in the frontal lobes
4	26	F	No	Biometry <5th percentile Diffuse petechial hemorrhagic cortical lesions with frontal schizencephaly Foci of calcified leukomalacia
5	24	F	Focal thrombosis of allanto-chorial vessels, and FVM	Bilateral ischemic/hemorrhagic periventricular and germinal layer lesions Large right schizencephaly
6	23,1	M	MPVFD	Biometry <5th percentile Multiple ischemic and hemorrhagic bilateral cerebral lesions, including cerebellum and cerebral stem. Bilateral schizencephaly
7	25	M	Right kidney hypo/dysplasia; Placental FVM	Biometry <5th percentile Bilateral porencephaly with microgyric cortical pattern Widespread parenchymal petechial hemorrhagic lesions
8	40	M	Mild facial dysmorphy:anteverted nares, long philtrum, small chin, protruded tongue Severe IUGR, Placental FVM	Biometry <5th percentile Cerebral atrophy Diffuse petechial cortical hemorrhagies Necrotic lesions of subcortical and periventricular white matter

Table 3. Molecular results of the eight fetuses carrying a *COL4A1* mutation.

Case	*COL4A1* (NM_001845.4) cDNA position	Protein change	Inheritance
1	c.2132G>T	p.(Gly711Val)	De novo
2	c.324+1G>A.	Exon 5 skipping	De novo
3	c.1990+1G>A	Exon 27 skipping	De novo
4	c.2753 G>A	p.(Gly918Glu)	De novo
5			
6	c.2476G>A	p.(Gly826Arg)	Maternal mosaic
7			
8	c.3041G>A	p.(Gly1014Asp)	De novo

In two cases, fetal MRI only confirmed the US findings. However, in one case (Case 8), MRI proved to be more informative than US, probably due to the late gestational age at diagnosis.

Neuropathological Findings

Our neurohistopathological studies performed on specimens from supra- and infratentorial levels demonstrated a strikingly similar pathological pattern summarized in Table 2. This was dominated by diffuse, multifocal hemorrhagic lesions of various abundance and ages, associated at the supratentorial level to a range of post-ischemic damages involving both the grey and the white matters. Multifocal hemorrhages were scattered over wide areas of the cerebral hemispheres, from sub-arachnoïdal to periventricular zones with ventricular flooding. Of note, parenchymal extravasations were found in the vicinity of dilated straight veins with discontinuous wall, devoid of branching and running with a "fan-like" pattern from the ventricular region toward

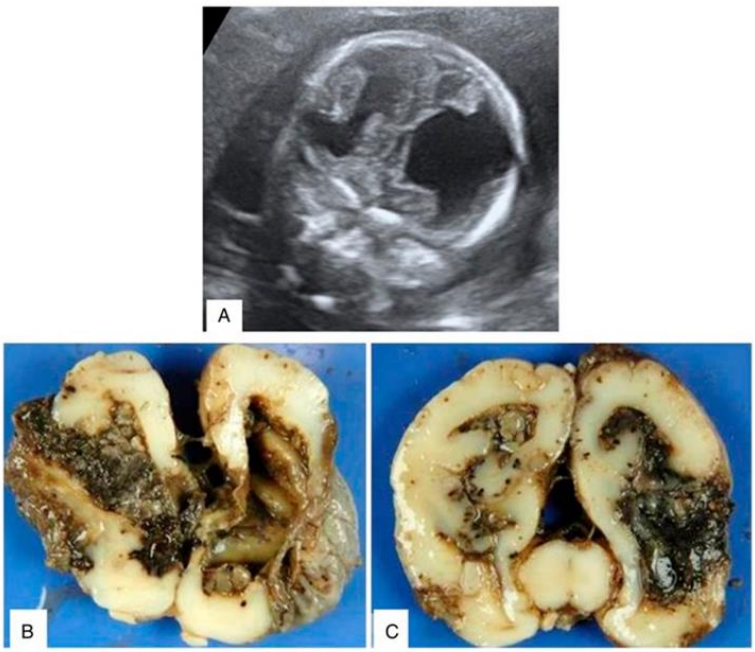


Figure 1. Ultrasound and neuropathology gross findings in 23 weeks fetal patient (Case N°6). Bilateral ventricular dilatation with schizencephaly on coronal cranial view (a); corresponding coronal sections showing: numerous cortical petechiae, large cleft extending from the ventricle to the cortical surface in the right hemisphere, and periventricular calcifications; midline structures are not identified, (b, c).

the sub-arachnoïdal zone. Intraventricular hemorrhage was associated with ependymal post-hemorrhagic responses marked by excess of iron-laden macrophages, while superficial parenchymal extravasates were devoid of inflammatory reaction. In all cases, parenchymal disruptive necrotic and hemorrhagic lesions were noticed at the supratentorial levels; associating capillary proliferations, endothelial cells swelling, mineralization, astrogliosis, and inflammatory reaction with iron-laden macrophages, but without axonal swelling or apoptotic figures. In 6 over 8 cases, cavitary lesions leading to uni/bilateral porencephaly and/or schizencephaly were found. Porencephalic lesions showed a large defect (porus) in the full thickness of one or both cerebral hemispheres, delimited by polymicrogyria and heterotopias. Schizencephaly presented with clefts in the cortical mantle and was also associated with polymicrogyria and heterotopias. In both, cavitary lesions were covered by thickened and hypervascularized leptomeninges in continuity with a rim of cortical scar; both filled with nuclear residues and inflammatory cells, made of active astrocytes and foamy and iron-laden macrophages.

Table 2 summarizes various macro- and microscopic findings on post-mortem examination.

Morphometric Study

Group comparisons indicated a strong albeit not always significant tendency of hypervascularization (Figure 2 A, B; D, E) and of vessel dilatation: (Figure 2 C, F); all three pairwise comparisons (Figures 3–5) reaffirmed these findings: increased numbers of vessels and vessel dilatation were patent. In WM, data showed significant differences between the two *COL4A1* fetus cases and their respective age-matched control, suggesting an important attenuation and disruption of the smooth muscle/media layers of small vessels, most likely small arteries and arterioles (Supplemental figure 2). These observations were further substantiated by our morphometric results. Figure 6 and Supplemental figure 2 illustrate the 40 weeks cases; values for the 28 weeks cases followed the same pattern (not shown). Interestingly, GM seemed to be devoid of smooth muscle attenuation lesions at both gestational ages.

Molecular Screening

All fetuses had a molecular screening for *COL4A1* and *COL4A2* genes. Identified mutations and their segregation are summarized in Table 3.

Discussion

Cerebrovascular disease is one of the most prominent consequences of *COL4A1* and *COL4A2* mutations. To establish a characteristic pattern of the disease, most studies rely on brain imaging data.[21,22] Our neuropathological study performed in

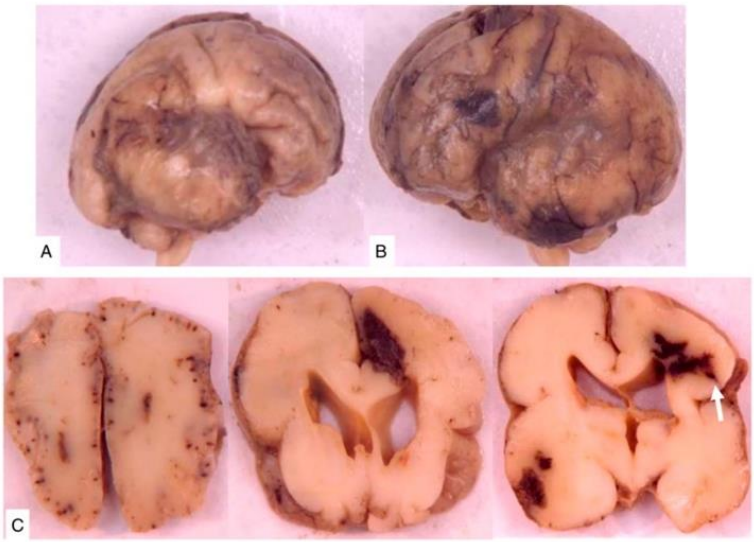


Figure 2. Macroscopic brain findings in a 26-week fetal patient (Case N°4) showing the signs of meningeal hemorrhages on lateral views (a, b) and cortical petechiae, several large parenchymal hemorrhagic foci, a hemorrhagic schizencephaly in the left frontal lobe (white arrow), and a ventricular enlargement (c).

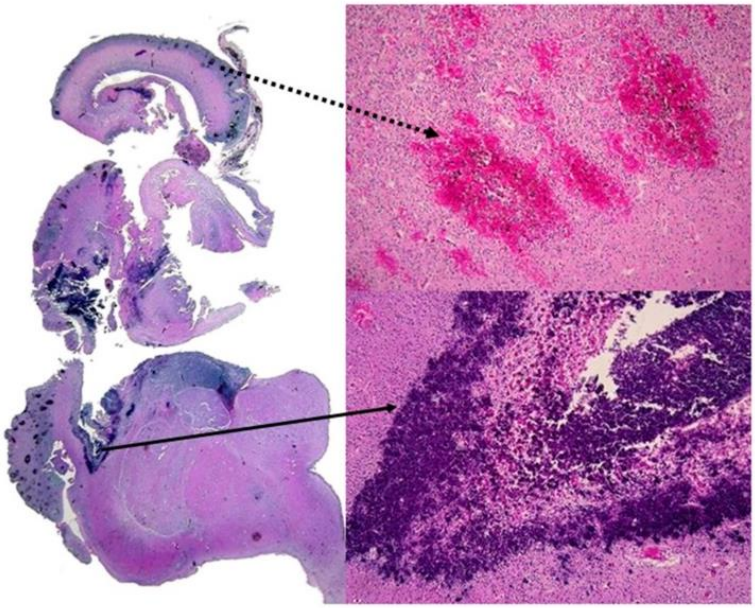


Figure 3. Histological features were identical in all cases: the presence of multiple, recent, petechial hemorrhages (dashed arrow) and thick calcified border surrounding the schizencephalies (solid arrow) (H&E).

a series of 8 fetuses after pregnancy termination for severe brain anomalies permits to go further in the evaluation of the natural history of the complex lesional spectrum of *COL4A1* mutations.

Adequacy of Brain Imaging

This study shows an overall good correlation between US scan and post-mortem gross neuropathological findings. In

one case (case 8), this correlation is weaker probably due to the late gestational age at US diagnosis. MRI might be indicated to refine prenatal US diagnosis in the third trimester. These findings are in accordance with the data reported in the literature on isolated cases.[19,20,26-30] On US examination, the principal brain manifestations were poor cranial biometry, schizencephaly, porencephaly, and ventricular dilatation. The mean age at diagnosis of cerebral hemorrhage was 23.5 weeks, in accordance with the cases hitherto reported.[8,9,29-37] In this series, seven cases had brain anomalies detected on US during the second trimester screening. In one fetus, severe cerebral lesions were of late onset, revealed at 39 WG, while a previous US scan performed at 32 WG was normal (case 8).

Lesional Spectrum of fetal COL4A1 *Disease*

Publications on *COL4A1* mutation manifestations before birth are sparse.[9,18-20]In 2012, Lichtenbelt et al. published the first case of a prenatal *COL4A1* mutation in a fetus of 21,6 weeks with US findings of cerebral hemorrhage.[21] All fetuses in our study had severe brain manifestations of the *COL4A1* disease (Figures 7 and 8). Our neuropathological studies demonstrated a strikingly similar pattern dominated by hemorrhagic encephalopathy, constantly associated with post-ischemic damage in both grey and white matters. No major thrombotic phenomena were noted, in contrast with a recent study,[23] suggesting that endothelial damage, a major thrombosis initiating event, was less pronounced in our cases. Indeed, as pointed out by these authors, *COL4A1* mutations may show different expressivity[23] and, consequently, different spectrum of pathology.

Hemorrhagic Encephalopathy

Hemorrhagic Encephalopathy (HE) is characterized by disseminated multifocal petechial hemorrhages extending into the brain stem and the cerebellum. This peculiar form of cerebral hemorrhage is classically reported in the context of abnormal clotting, mainly during materno-fetal alloimmunization and anemia. Fetal venous immaturity and/or stasis are considered aggravating factors. The prevailing view is that rupture of veins and capillaries occurs first and results in a spectrum of hemorrhagic lesions leading to secondary hemodynamic disturbances. In 4 fetuses carrying *COL4A1* mutations, Schannon et al.[23] also observed placental vascular malperfusion as in our series and postulated that may contribute to fetal brain lesions. Interestingly, low to high grade fetal vascular malperfusion (FVM) and massive perivillous fibrin deposition (MPVFD) have been observed as well in 4/8 cases in our cohort (Table 2).

Maurice et al.[22] published a retrospective study aiming to establish the prevalence of *COL4A1* and *COL4A2* mutations in fetuses presenting a characteristic pattern of cerebral injury. They proposed a fetal phenotype associated with the disease that they characterized by severe and/or multifocal hemorrhagic lesions that can be supratentorial or infratentorial with multifocal ischemic and hemorrhagic lesions of different ages, associated with schizencephaly or porencephaly. The clinical picture can worsen throughout pregnancy.

In our series, disseminated petechial hemorrhages were of various abundance and of different ages. Interestingly, extravasates were found in the vicinity of dilated capillaries and small vessels, with discontinuous wall. Morphometric studies of brain micro vessels confirmed this and showed an important attenuation and disruption of the smooth muscle/media layers. These findings are consistent with small vessel walls fragility, the main component of *COL4A1* disease,[1] while hypervascularity may be linked to augmented angiogenesis due to hypoxic-ischemic insult. Of note, while iron-laden macrophages were abundantly observed in the ventricular zone and deep white matter, they were lacking in the cortical and sub-cortical extravasations. This permits to oppose in the same brain early rising periventricular hemorrhages to late onset superficial ones that have possibly occurred during medical terminations by vaginal delivery.

Disruptive Lesions

On US, the principal brain manifestations in this series were poor cranial biometry, schizencephaly, porencephaly, and ventricular dilatation. The definitions of porencephaly and schizencephaly remain challenging due to the great variability in their descriptions in the medical literature. Both are typically considered as a destructive lesion of the full thickness of the cerebral mantle extending from the ventricle to the cortical surface, often bilateral and occurring in the first half of gestation. Porencephaly (OMIM #175780) describes a large defect of the cerebral hemispheres classically associated with heterotopia and polymicrogyria. However, the term is often used to indicate a wide range of encephaloclastic cavitary lesions. Schizencephaly (OMIM#269160) indicates a narrow cleft in the cortical mantle, lined by heterotopic grey matter, often associated with polymicrogyria too. Both porencephaly and schizencephaly clinically cause hemiplegia, tetraplegia, epilepsy, and intellectual disability.[33,34] While these lesions are commonly seen in acquired ischemic events, there is evidence for a genetic cause of vascular malformation in some cases.[5,35] Indeed, both lesions are found in association with *COL4A1* and *COL4A2* gene mutations.[2,3,5,22,32,36]

Morphometric Data

The most important morphometric findings in this study are that *COL4A1* patients showed abnormally augmented angiogenesis, vessel dilatation, and focal loss and attenuation of their arterial smooth muscle coats. These findings concord with studies in mutant *COL4A1* mice, a good model to study brain arteries by non-invasive techniques.[37,38] In these

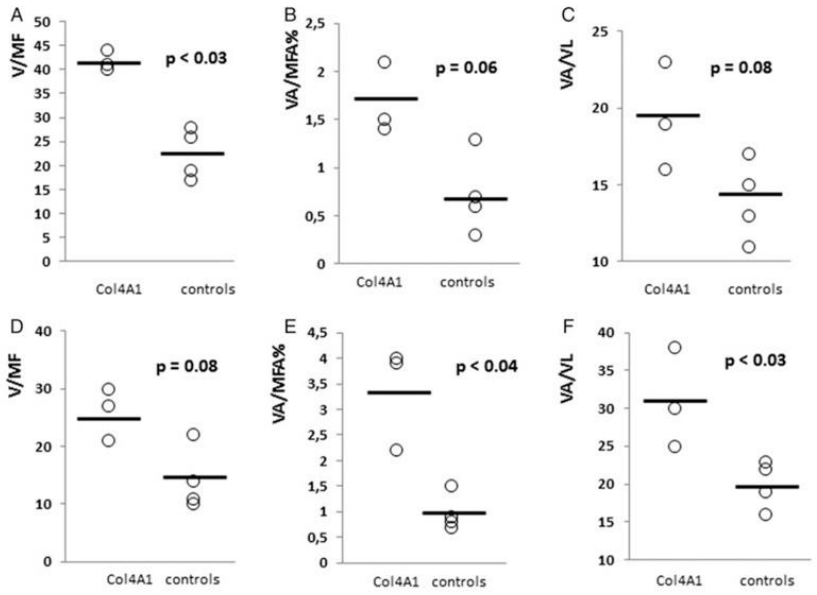


Figure 4. Brain vessels, comparisons between groups: upper row: GM; lower row: WM; *COL4A1* values are invariably higher than control values. A,D: V/MF: number of vessels per microscopic field; B,E: V/MF%: percentage area occupied by vessels per microscopic field; C,F: VA/VL: vessel area to vessel length ratio (see text for definitions). Points represent mean values per case and lines are medians. In this figure, A–C and D–F indicate, respectively, GM and WM.

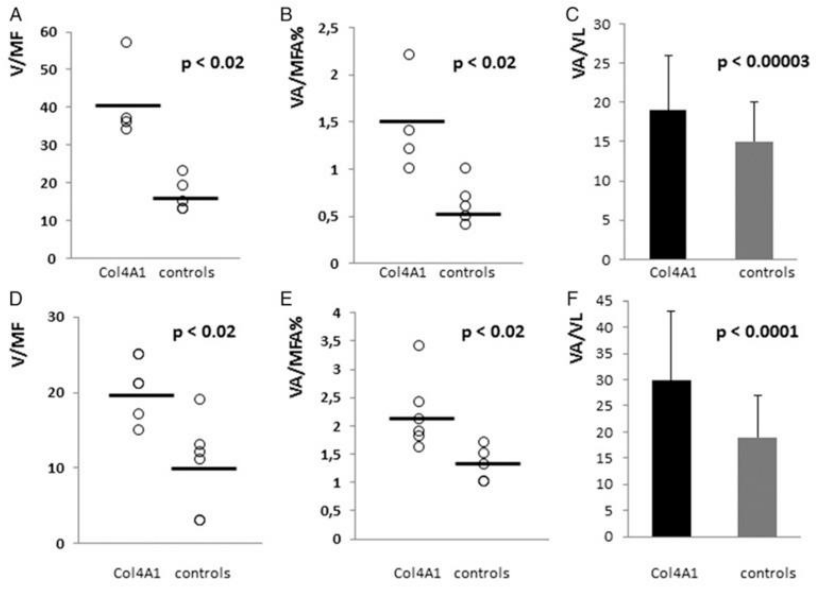


Figure 5. Brain vessels, 1st pairwise comparison, 23 weeks: upper row: grey matter; lower row: white matter. A,D: V/MF: number of vessels per microscopic field; B,E: V/MF%: percentage area occupied by vessels per microscopic field; C,F: VA/VL: vessel area to vessel length (see text for definitions). Points represent mean values per microscopic field, lines are medians, and individual vessel measures are compared in 5F.

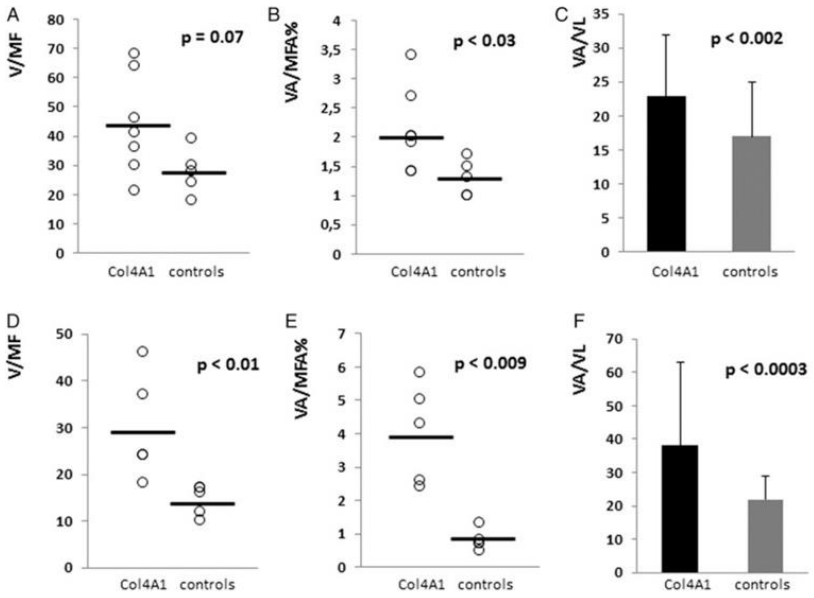


Figure 6. Brain vessels, 2nd pairwise comparison, 26 weeks: upper row: grey matter; lower row: white matter. A,D: V/MF: number of vessels per microscopic field; B,E: V/MF%: percentage area occupied by vessels per microscopic field; C,F: VA/VL: vessel area to vessel length (see text for definitions). Points represent mean values per microscopic field, lines are medians, and individual vessel measures are compared in 6F.

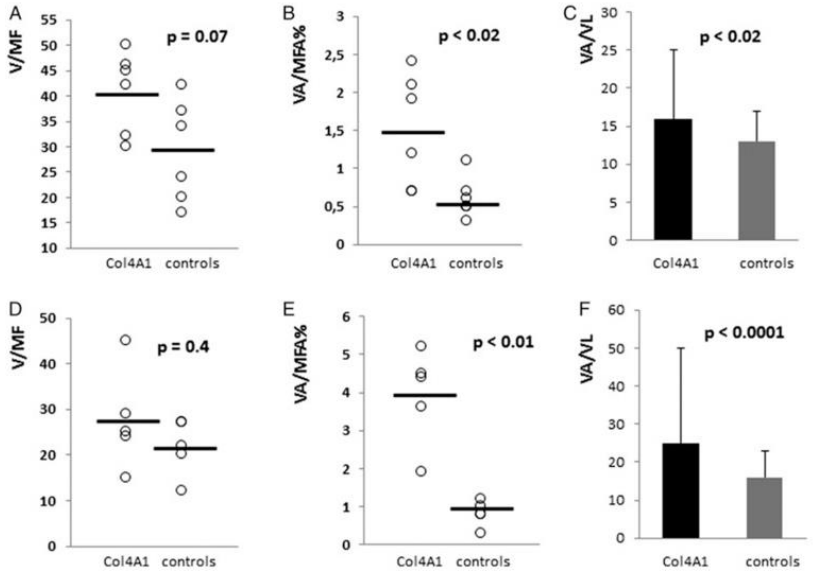


Figure 7. Brain vessels, 3rd pairwise comparison, 38 weeks: upper row: grey matter; lower row: white matter. A,D: V/MF: number of vessels per microscopic field; B,E: V/MF%: percentage area occupied by vessels per microscopic field; C,F: VA/VL: vessel area to vessel length (see text for definitions). Points represent mean values per microscopic field, lines are medians, and individual vessel measures are compared in 7F.

studies, increased vessel density, abnormal excessive branching and increase of vessel diameter,[38] or abnormal tortuosity in retinal vessels has been attributed to accumulation of abnormal *COL4A1*/*COL4A2* heterodimers in endothelial cells, resulting in cytotoxic endoplasmic reticulum stress. Hypervascularization was likely attributable to the known angiogenic properties of mutated *COL4A1*, but we cannot exclude the causative effect of ischemia.

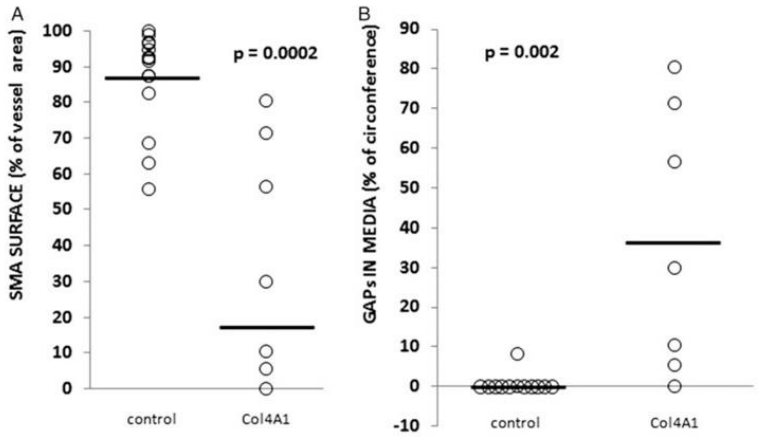


Figure 8. Pairwise comparison for the two parameters: SMA surface and GAPs in media of the *COL4A1* case and its age-matched control at 40 weeks after SMA immunostaining: both parameters (SMA surface and GAPs in media) are significantly different. Points represent individual vessels and lines are medians.

Furthermore, a recent study using *Col4A1*$^{1+/G498V}$ and *Col4A1*$^{G498V/G498V}$ mutant mice[4] reported smooth muscle cells abnormalities of variable severity in brain arteries. In this model, authors report strongly attenuated SMA immunostaining of vessel walls, as well as more subtle ultrastructural abnormalities such as thinning of smooth muscle cells and blurring of their edges. They postulated that, the initially well-constituted vessel media degenerated through progressive apoptosis of smooth muscle cells purportedly attributable to pericellular basal membrane dysfunction. In turn, local and segmental loss of smooth muscle cells leaves endothelial cells in direct contact with brain parenchyma. Interestingly, our morphometric results in human brains based on α-SMA immunohistochemistry are in accordance with this experimental data and a recent study in human fetuses describing gaps and vacuolation of smooth muscle cells in arterial media.[23]

Our quantitative results (Figures 3–6) support the following mechanistic sequence: weakening of the arterial walls due to basal membrane abnormalities and media thinning is followed by progressive dilatation and ultimately by vessel rupture and hemorrhage. It may be speculated further, that brain blood supply may be more widely affected since it is thought that arteriolar smooth muscle cell contractility functionally regulates regional blood flow in the brain.[39]

COL4A1 *Disease and Placental Lesions*

Interestingly, Schannon et al.[23] correlated placental finding in 4 fetuses carrying *COL4A1* mutations and postulated that observed FVM (3/4 cases) may be causally related to the mutation. Although low to high grade fetal vascular malperfusion (FVM, 2/8 cases) and massive perivillous fibrin deposition (MPVFD, 1/8 case) were observed in our cohort (Table 2), we agree with the authors that a larger and a more systematic study focused on placentas is needed to explore the potential role of the mutation in the FVM lesions.

COL4A1 *Disease and Genetic Counseling*

All fetuses of this study had severe brain manifestations of the *COL4A1* disease with small vessels encephalopathy. Small vessel anomalies seem to be a recurrent finding in *COL4A1* mutations. Shannon et al[23] report as well thrombosis and arteriopathy as a part of the fetal spectrum of pathology. Interestingly, in our cohort, 2 fetuses presented additional extracranial malformations: fetal hydrops with pulmonary hypoplasia, rectal atresia with imperforate anus, and hepatosplenomegaly in one (case 2); hypo-dysplastic kidneys in the other fetal patient (case 7).

In agreement with the physiopathological mechanism underlying *COL4A1* mutation, hypoplastic kidneys could belong to the spectrum of *COL4A1* disease. However, other genetic variants might be questioned in our second polymalformed fetus with a possible dual genetic diagnosis.

The spectrum of neurological symptoms in *COL4A1/A2* patients varies in severity and age of onset with wide intrafamilial heterogeneity. Patients typically present infantile hemiparesis, seizures, intellectual disability, and stroke. Importantly, manifestations may occur in previously asymptomatic adults and brain imaging can be normal.[26,31,37,38]

In our familial cases (cases 5, 6, ant 7) with a maternal inheritance of the mutation, the mother is a mosaic carrier with a predominance of the healthy allele and remains asymptomatic even after detailed clinical and imaging examination according to protocols for *COL4A1/A2* mutated patients. In addition, although one might expect penetrance and expression variability in her different *COL4A1* mutated offsprings, the neuropathological phenotype remained severe in all cases with an early-onset cerebral hemorrhagic lesion.

Molecular studies guided by detailed neuropathological examination are mandatory if the parents opt for terminations, in order to confirm the US diagnosis and exclude other hemorrhagic encephalopathies bearing distinct molecular backgrounds. This multidisciplinary prenatal care approach offers an optimal genetic counseling to the parents, crucial for both the follow-ups of their future pregnancies, and, if carriers, for their own.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Jelena Martinovic https://orcid.org/0000-0003-4595-7271

Supplemental Material

Supplemental material for this article is available online.

References

1. YonedaY HK, Kato M, et al. Phenotypic spectrum of COL4A1 mutations: Porencephaly to schizencephaly. *Ann Neurol*. 2013; 73(73):48-57. https:/.ghr.nlm.nih.gov/gene/COL4A1
2. Gould DB, Phalan FC, Van Mil SE, et al. Role of *COL4A1* in small-vessel disease and hemorrhagic stroke. *N Engl J Med*. 2006;354(14):1489-1496.
3. Ratelade J, Mezouar N, et al. Severity of arterial defects in the retina correlates with the burden of intracerebral haemorrhage in C*OL4A1*-related stroke. *J Pathol*. 2018;244(4):408-420.
4. Govaert P. Prenatal stroke. *Semin Fetal Neonatal Med*. 2009; 14:250-266.
5. Jeanne M, Gould DB. Genotype-phenotype correlations in pathology caused by collagen type IV alpha 1 and 2 mutations. *Matrix Biol*. 2017;57-58:29-44.
6. Gould DB, Phalan FC, Breedveld GJ, et al. Mutations in *COL4A1* cause perinatal cerebral hemorrhage and porencephaly. *Science*. 2005;308:1167-1171.
7. Mao M, Smith RS, Alavi MV, et al. Strain-dependent anterior segment dysgenesis and progression to glaucoma in *COL4A1* mutant mice. *Invest Ophthalmol Vis Sci*. 2015;56:6823-6831.
8. Elchalal U, Yagel S, Gomori JM, et al. Fetal intracranial hemorrhage (fetal stroke): Does grade matter? *Ultrasound Obstet Gynecol*. 2005;26:233-243.
9. Lustig-Gillman I, Young BK, Silverman F, et al. Fetal intraventricular hemorrhage: Sonographic diagnosis and clinical implications. *J Clin Ultrasound*. 1983;11:277-280.
10. Tonduti D, Pichiecchio A, La Piana R, et al. *COL4A1*-related disease: Raised creatine kinase and cerebral calcification as useful pointers. *Neuropediatrics*. 2012;43:283-288.
11. Vahedi K, Alamowitch S. Clinical spectrum of type IV collagen (*COL4A1*) mutations: A novel genetic multisystem disease. *Curr Opin Neurol*. 2011;24(1):63-68.
12. Cavallin M, Mine M, Philbert M, et al. Further refinement of COL4A1 and COL4A2 related cortical malformations. *Eur J Med Genet*. 2018;61(12):765-772
13. Livingston J, Doherty D, Orcesi S, et al. *COL4A1* mutations associated with a characteristic pattern of intracranial calcification. *Neuropediatrics*. 2011;42:227-233.
14. Sibon I, Coupry I, Menegon P, et al. *COL4A1* mutation in axenfeld-rieger anomaly with leukoencephalopathy and stroke. *Ann Neurol*. 2007;62(2):177-184.
15. Cecchetto G, Milanese L, Giordano R, et al. Looking at the missing brain: hydranencephaly case series and literature review. *Pediatr Neurol*. 2013;48:152-158.
16. Coupry I, Sibon I, Mortemousque B, et al. Ophthalmological features associated with *COL4A1* mutations. *Arch Ophthalmol*. 2010;128(4):483-489.
17. De Vries LS, Koopman C, Groenendaal F, et al. *COL4A1* mutation in two preterm siblings with antenatal onset of parenchymal hemorrhage. *Ann Neurol*. 2009;65:12-18.
18. Meuwissen ME, de Vries LS, Verbeek HA, et al. Sporadic *COL4A1* mutations with extensive prenatal porencephaly resembling hydranencephaly. *Neurology*. 2011;76: 844-846.
19. Vermeulen RJ, Peeters-Scholte C, Van Vught JJ, et al. Fetal origin of brain damage in 2 infants with a *COL4A1* mutation: Fetal and neonatal MRI. *Neuropediatrics*. 2011;42:1-3.
20. Lichtenbelt KD, Pistorius LR, de Tollenaer SM, et al. Prenatal genetic confirmation of a *COL4A1* mutation presenting with sonographic fetal intracranial hemorrhage. *Ultrasound Obstet Gynecol*. 2012;39:726-727
21. Maurice P, Guilbaud L, et al. COL4A1 and *COL4A2* mutations: when to test a fetus? *Ultrasound Obstet Gynecol*. 2021;57(5): 783-789
22. Shannon P, Hum C, Parks T, et al. Brain and placental pathology in fetal *COL4A1* related disease. *Pediatr Dev Pathol*. 2021;24(3):175-186.
23. Martinovic J, Sebire NJ. Modern Fetal Autopsy Protocole. In: *Martinovic J. Practical Manual of Fetal Pathology*. Springer Nature; 2021:87-109.
24. Sommer C, Straehle C, Koethe U, Hamprecht FA. Ilastik: Interactive learning and segmentation toolkit. IEEE International Symposium on Biomedical imaging: From nano to macro; 30 March-2 April 2011; Chicago, IL
25. Zudaire E, Gambardella L, Kurcz C, Vermeren S. A Computational tool for quantitative analysis of vascular networks. *PLoS One*. 2011;6(11):e27385.
26. Meuwissen MEC, Halley DJJ, Smit LC, et al. The expanding phenotype of *COL4A1* and *COL4A2* mutations: Clinical data on 13 newly identified families and a review of the literature. *Genet Med*. 2015;17:843-853.
27. Sato Y, Shibasaki J, Aida N, et al. Novel *COL4A1* mutation in fetus with early prenatal onset of schizencephaly. *Human Genome Variation*. 2018;5(5):4

28. Matsumoto T, Miyakoshi K, Fukutake M, Ochiai D, Minegishi K, Tanaka M. Intracranial sonographic features demonstrating in utero development of hemorrhagic brain damage leading to schizencephaly-associated *COL4A1* mutation. *J Med Ultrason*. 2001;42(3):445-446.
29. Colin E, Sentilhes L, Sarfati A, et al. Fetal intracerebral hemorrhage and cataract: Think *COL4A1*. *J Perinatol Off J Calif Perinat Assoc*. 2014;34(1):75-77
30. Giorgio E, Vaula G, Bosco G, et al. Two families with novel mutations in *COL4A1*: When diagnosis can be missed. *J Neurol Sci*. 2015;352(1-2):99-104.
31. Garel C, Rosenblatt J, Moutard ML, et al. Fetal intracerebral hemorrhage and *COL4A1* mutation: promise and uncertainty. *Ultrasound Obstet Gynecol Off J IntSoc Ultrasound Obstet GynecolFeb*. 2013;41(2):228-230
32. Durrani-Kolarik S, Manickam K, Chen B. *COL4A1* mutation in a neonate with intrauterine stroke and anterior segment dysgenesis. *Pediatr Neurol*. 2017;66:100-103.
33. Smith C, Squier MV. Acquired diseases of the nervous system. In: Keeling JW, Khong TY, eds. *Fetal and Neonatal Pathology*. 4th ed. London, UK: Springer; 2007:719-746.
34. Mancini GM, De Coo IF, Lequin MH, etal. Hereditary porencephaly: Clinical and MRI findings in two Dutch families. *Eur J Paediatr Neurol*. 2004;8:45-54.
35. Yakovlev PI and Wadsworth RC. Shyzencephalies, a study of the congenital clefts in cerebral mantle; clefts with fused lips. *J Neurpathol Exp Neurol*. 1946;5:116-130.
36. Hunter A. In: Re S, Jg H, eds. *Porencephaly in Human Malformations and Related Anomalies*. New York, NY: Oxford University Press; 2006:645-654.
37. Jeanne M, Jorgensen J, Gould DB. Molecular and genetic analyses of collagen type IV mutant mouse models of spontaneous intracerebral hemorrhage identify mechanisms for stroke prevention. *Circulation*. 2015;131:1555-1565.
38. Alavi MV, Mao M, Pawlikowski BT, Kvezereli M, Duncan JL, Libby RT, et al. *COL4A1* mutations cause progressive retinal neovascular defects and retinopathy. *Sci Rep*. 2016;6: 18602.
39. Hill RA, Tong L, Yuan P, Murikinati S, Gupta S, Grutzendler J. Regional blood flow in the normal and ischemic brain is controlled by arteriolar smooth muscle cell contractility and not by capillary pericytes. *Neuron*. 2015;87:95-110.

Ultrasound Obstet Gynecol 2021; **57**: 783–789
Published online in Wiley Online Library (wileyonlinelibrary.com). **DOI:** 10.1002/uog.22106

Prevalence of *COL4A1* and *COL4A2* mutations in severe fetal multifocal hemorrhagic and/or ischemic cerebral lesions

P. MAURICE[1], L. GUILBAUD[1], J. GAREL[2], M. MINE[3], A. DUGAS[1], S. FRISZER[1], E. MAISONNEUVE[1], M.-L. MOUTARD[4], T. COSTE[3], D. HÉRON[5], E. TOURNIER-LASSERVE[3], C. GAREL[2] and J.-M. JOUANNIC[1]

[1]*Service de Médecine Fœtale, Hôpital Armand Trousseau APHP, Sorbonne Université, Paris, France;* [2]*Service de Radiopédiatrie, Hôpital Armand Trousseau APHP, Sorbonne Université, Paris, France;* [3]*Service de Génétique Moléculaire Neurovasculaire, Hôpital Saint-Louis, APHP, Paris, France;* [4]*Service de Neuropédiatrie, Hôpital Armand Trousseau APHP, Sorbonne Université, Paris, France;* [5]*Service de Génétique, Groupe Hospitalier Pitié-Salpêtrière, Hôpital Armand Trousseau APHP, Paris, France*

KEYWORDS: *COL4A1*; *COL4A2*; fetal cerebral injury; hemorrhagic lesion; ischemic-hemorrhagic lesion; porencephaly; prenatal diagnosis; schizencephaly

CONTRIBUTION

What are the novel findings of this work?
This is the first prenatal series of cases establishing the prevalence of *COL4A1* and *COL4A2* gene mutations among fetuses presenting with a phenotype suggestive of cerebral injury, specifically severe and/or multifocal hemorrhagic or ischemic-hemorrhagic cerebral lesions. These anomalies could be of different ages and associated with schizencephaly or porencephaly.

What are the clinical implications of this work?
COL4A1 and *COL4A2* gene mutations should be sought systematically in cases of severe and/or multifocal hemorrhagic or ischemic-hemorrhagic cerebral lesions.

ABSTRACT

Objective *To establish the prevalence of* COL4A1 *and* COL4A2 *gene mutations in fetuses presenting with a phenotype suggestive of cerebral injury.*

Methods *This was a single-center retrospective analysis of all cases of fetal cerebral anomalies suggestive of* COL4A1 *or* COL4A2 *gene mutation over the period 2009–2018. Inclusion criteria were: (1) severe and/or multifocal hemorrhagic cerebral lesions; (2) multifocal ischemic-hemorrhagic cerebral lesions. These anomalies could be of different ages and associated with schizencephaly or porencephaly. Between fetuses with and those without a mutation, we compared gestational age at the time of diagnosis, parity and fetal gender.*

Results *Among the 956 cases of cerebral anomaly diagnosed in our center during the 10-year study period, 18 fetuses were identified for inclusion. A pathogenic* COL4A1 *gene mutation was found in five of these cases, among which four were* de-novo *mutations. A variant of unknown significance was found in four fetuses: in the* COL4A1 *gene in one case and in the* COL4A2 *gene in three cases. No* COL4A1 *or* COL4A2 *mutation was found in the remaining nine fetuses. The median (interquartile range) gestational age at diagnosis was significantly lower in cases with a mutation (24 (22–26) weeks) than in cases without a mutation (32 (29.5–34.5) weeks) (*P = 0.03*).*

Conclusions *A phenotype suggestive of cerebral injury was found in 18 of the 956 (1.9%) cases in our population, in 28% of which there was an associated* COL4A1 *or* COL4A2 *mutation.* COL4A1 *and* COL4A2 *gene mutations should be sought systematically in cases of severe and/or multifocal hemorrhagic or ischemic-hemorrhagic cerebral lesions, with or without schizencephaly or porencephaly.* © *2020 International Society of Ultrasound in Obstetrics and Gynecology.*

INTRODUCTION

Type-IV collagens are proteins of the basal membrane and are expressed in all tissues, including vessel walls. *COL4A1* and *COL4A2* genes code for the most abundant type-IV collagen[1]. Mutations of these genes, particularly mutations in the *COL4A1* gene, have been associated with a wide spectrum of anomalies, in particular of the brain (cerebral aneurysms, small-vessel brain disease, stroke) and the eye (retinal arterial tortuosity, cataract, microphthalmia, Axenfeld–Rieger syndrome), as well as multiple systemic findings, including kidney

Correspondence to: Prof. J.-M. Jouannic, Service de Médecine Fœtale, Hôpital Armand Trousseau, 26 rue du Docteur Arnold Netter, 75012 Paris, France (e-mail: jean-marie.jouannic@trs.aphp.fr)

Accepted: 27 May 2020

involvement, muscular cramps, Raynaud phenomenon, cardiac arrhythmia and hemolytic anemia[2]. Some of these anomalies have been described in children and, more recently, in a fetus, in the context of intracerebral hemorrhage[3,4]. Several isolated cases of prenatal cerebral damage related to *COL4A1* mutations have been reported in fetuses and neonates[5–9]. No *COL4A2* mutations have been reported so far in the fetus, although *COL4A1* and *COL4A2* mutations present with similar phenotypes in children and adults[10–12]. Based on analysis of these reported cases of cerebral anomalies, we have defined a possible distinctive fetal phenotype for *COL4A1* gene mutations, represented by: (1) severe and/or multifocal hemorrhagic cerebral lesions; and/or (2) multifocal ischemic-hemorrhagic cerebral lesions. These anomalies could be of different ages and associated with schizencephaly or porencephaly.

In this study, we aimed to establish the prevalence of *COL4A1* and *COL4A2* gene mutations in fetuses carrying these cerebral anomalies.

METHODS

We conducted a single-center retrospective analysis of all cases of fetal cerebral anomalies suggestive of *COL4A1* or *COL4A2* gene mutation identified over a 10-year period (2009–2018) at Trousseau Hospital. Inclusion criteria were: (i) severe and/or multifocal hemorrhagic cerebral lesions; (ii) multifocal ischemic-hemorrhagic cerebral lesions. These anomalies could be of different ages and associated with schizencephaly or porencephaly.

Informed consent had been obtained from the parents and this retrospective study was approved by the local institutional review board (DHU MAMUTH). All methods were performed in accordance with relevant guidelines and regulations.

All cases were referred after routine second- or third-trimester ultrasound (US) examination had revealed fetal cerebral anomaly. All second-line US examinations were performed within 1 week of referral using a Voluson E8 or E10 (GE Healthcare, Zipf, Austria) or an Aplio 400 (Toshiba, Tokyo, Japan) US system. The indication for magnetic resonance imaging (MRI) was discussed individually at multidisciplinary meetings. Fetal MRI examinations were performed using an Achieva 1.5-Tesla (T) unit (Philips Medical Systems, Best, The Netherlands) or, since 2016, an Optima 1.5-T MR450W GEM unit (GE Healthcare, Milwaukee, WI, USA), 30–40 min following fetal sedation, achieved by maternal oral administration of flunitrazepam (1 mg) or zolpidem (10 mg). Fetal brain examination was performed using a phased-array abdominal coil. For each patient, the MRI protocol included T1, T2, gradient-echo T2 and diffusion-weighted sequences.

Amniocentesis was offered in all cases and the first-line genetic analysis consisted of fetal karyotyping and array comparative genomic hybridization analysis.

According to French law, termination of pregnancy (TOP) was performed at the mother's request once two experts had certified the severity of the fetal anomaly. Fetal blood sampling for platelet count was performed in all cases of TOP. Postmortem (PM) examination was offered in all cases of TOP and *in-utero* death (IUD).

Molecular screening for *COL4A1* and *COL4A2* genes was performed using targeted next-generation sequencing (SureSelect QXT, Agilent Technologies, Santa Clara, CA, USA) and Sanger sequencing on DNA extracted from PM fetal tissues, with informed consent from the parents. Reference National Center for Biotechnology Information (NCBI) transcripts used for analysis were NM_001845.5 for the *COL4A1* gene and NM_001846.3 for the *COL4A2* gene. Pathogenic or probably pathogenic sequence variants detected in fetuses were investigated in the parents' DNA when available. To fully establish the *de-novo* nature of a variant, the fetus's DNA and DNA from both non-carrier parents were tested using nine sets of microsatellites, using the AmpFlSTR Profiler PCR Amplification Kit (Applied Biosystems, Thermo Fisher Scientific, Villebon sur Yvette, France). The reaction was carried out according to the manufacturer's instructions and loaded on an ABI 3130 genetic analyzer (Applied Biosystems) and the fragments were analyzed using GeneMapper v4.0 software (Applied Biosystems). Variants were classified according to the American College of Medical Genetics (ACMG)[13].

Following retrieval of clinical and genetic information from the hospital database, we used a Wilcoxon rank sum test to compare gestational age at the time of diagnosis between fetuses with and those without a mutation. Parity (nulliparous or parous) and fetal gender were also compared, using Fisher's exact test. Statistical analysis was performed using R, version 3.3.1 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Among the 956 cases of cerebral anomaly diagnosed in our center during the study period, 18 fetuses from 17 different families were identified for inclusion. There were nine cases with severe and/or multifocal hemorrhagic cerebral lesions, two with multifocal ischemic-hemorrhagic cerebral lesions and seven presenting with both. In addition, taking into account the PM findings, porencephaly was observed in six cases, schizencephaly in another four cases and both porencephaly and schizencephaly in one case (Case 5) (Figure 1). Cerebral injuries were supratentorial in 11 cases, infratentorial in two cases and both infra- and supratentorial in five cases. Lesions of different ages were found in seven cases.

The median gestational age at diagnosis was 30.5 weeks. Fetal brain anomalies were suspected at the second-trimester scan in seven cases and the third-trimester scan in 11 cases.

MRI was performed in 13 cases, allowing the diagnosis of additional cerebral hemorrhagic and/or ischemic-hemorrhagic cerebral lesions in seven cases, and proving helpful in characterizing lesions of different ages. In the other five cases, MRI was not performed, either because IUD occurred shortly after the US examination or

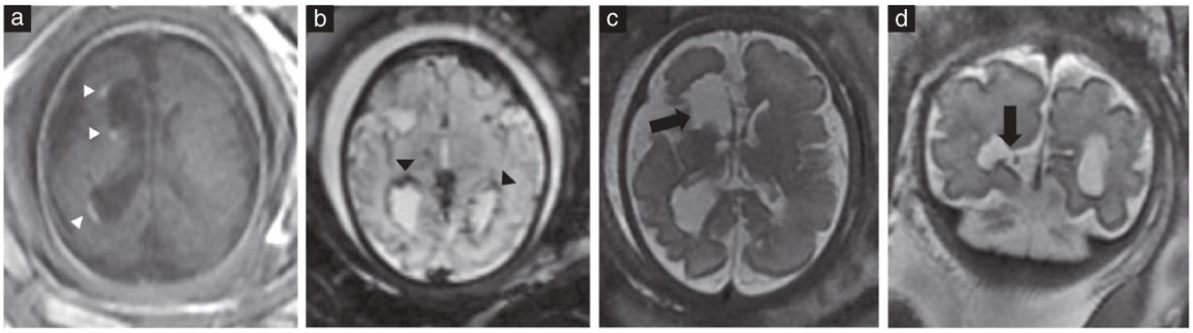


Figure 1 Cerebral magnetic resonance images in Case 5 at 33 weeks. (a) T1-weighted axial image demonstrating three subependymal hemorrhagic lesions, appearing as focal hyperintensities (arrowheads) lining the right lateral ventricle. (b) T2*-weighted axial image showing hypointensities (arrowheads) due to subependymal bleeding. (c) T2-weighted axial image, showing porencephaly in the anterior part of the right frontal lobe (arrow). (d) T2-weighted coronal image, showing schizencephaly, consisting of a communication between the occipital horn of the right lateral ventricle and the interhemispheric space (arrow), due to parenchymal destruction.

because the US examination was considered sufficient to diagnose severe cerebral lesions associated with a poor prognosis. In one case (Case 16), intraventricular hemorrhage associated with supratentorial extensive parenchymal hemorrhage was diagnosed at 34 weeks. The parents decided to continue the pregnancy. Delivery occurred at 36 weeks and postnatal imaging performed on day 5 confirmed the prenatal findings with associated porencephalic lesions. There are no data available on this surviving neonate, who was lost to follow-up.

The parents opted for TOP in 14 of the 18 cases in the series, with a normal fetal platelet count on blood sampling obtained before TOP in each case. There were three IUD. PM examination was performed in 15 of these 17 cases and all confirmed the prenatal findings. In 8/15 cases, PM examination allowed characterization of additional cerebral hemorrhagic and/or ischemic-hemorrhagic cerebral lesions and/or porencephalic lesions and/or signs of lesions of different ages. Four of these eight cases (Cases 9, 10, 14, 18) had undergone fetal MRI 1–3 weeks before TOP. MRI examination was not performed in five cases (Tables 1–3). In one case (Case 13), IUD was diagnosed shortly after the US examination. In this case, PM examination confirmed severe supra- and infratentorial hemorrhagic lesions and revealed additional ischemic-hemorrhagic lesions. In three other cases (Cases 2, 15 and 17), the US examination was considered sufficient to diagnose severe cerebral lesions and the parents opted for TOP, which was performed 2–4 weeks after the US examination. In all three cases, the PM examination revealed additional ischemic-hemorrhagic lesions, and an additional porencephalic lesion was revealed in Case 17. In the last case (Case 16), MRI was not performed because of late diagnosis (34 weeks). In this case, postnatal MRI confirmed the prenatal US findings, and additionally identified porencephalic lesions.

There were no additional extracerebral fetal anomalies identified, except for one case (Case 1) with extensive multivisceral hemorrhage found at histological examination. In this case, no extracerebral lesions were observed on prenatal imaging. There were no cases with associated cataract. No chromosomal anomalies were found except for one case (Case 3) with an incidental karyotype of 47,XXX.

A pathogenic *COL4A1* gene mutation was found in five of the 18 cases (Table 1), including one case reported previously (Case 1)[5]. Three of these cases (Cases 1, 3, 5) carried a typical glycine mutation within one of the Gly-X-Y motifs of the *COL4A1* triple helix (Class-5 mutations, according to ACMG)[12]. In Case 1, it was inherited from the mother, who did not suffer from stroke but has glaucoma and had a few punctate hypersignals on cerebral T2-weighted MRI (without any microbleed). In Cases 3 and 5, the mutation occurred *de novo*. One other case showed a *de-novo* mutation within the NC1 domain of *COL4A1* (Case 2), a domain essential for the wiring of the COL4A1/COL4A2 heterotrimer, and the fifth case had a *COL4A1* splice-site mutation (Case 4); this mutation, also *de novo*, has been reported previously in an affected child[11]. Both mutations are probably pathogenic (Class 4), according to ACMG criteria, and their *de-novo* occurrence is a strong argument for pathogenicity.

A variant of unknown significance (VOUS) was found in four fetuses (Class-3 mutations, according to ACMG) (Table 2): in *COL4A1* in one case and in *COL4A2* in three cases. Among the cases with *COL4A2* VOUS, two fetuses were siblings (from two different pregnancies, Cases 8 and 9) and carried an identical variant inherited from their clinically unaffected mother. The parents' status was unknown for the other two cases.

No *COL4A1* or *COL4A2* mutation was detected in the remaining nine fetuses (Table 3).

We did not observe significant phenotypical differences between the five fetuses carrying a pathogenic *COL4A1* mutation and the 13 fetuses with no pathogenic mutation. However, the median gestational age at diagnosis was significantly different between these two groups: 24 (interquartile range (IQR), 22–26) weeks for the fetuses with a pathogenic mutation *vs* 32 (IQR, 29.5–34.5) weeks for the fetuses without ($P = 0.03$).

The imaging of Case 3 is worth discussing in more detail (Figure 2). In this case, the routine second-trimester US examination at 22 weeks suggested abnormal corpus callosum. The second-line US examination performed at 25 weeks revealed a hyperechoic lesion of the cerebral parenchyma within the cingulate gyrus, above the genu of the corpus callosum, which was otherwise normal. No other fetal anomaly was found. Follow-up US and MRI examinations were performed at 27 weeks. An additional lesion, a large hyperechoic lesion of the left frontal lobe, was found at this US examination. MRI examination performed on the same day confirmed the ischemic-hemorrhagic nature of these lesions and also showed an additional hemorrhagic lesion in the right cerebellar lobe, which had not been seen during the US examination.

DISCUSSION

In this study we report the first consecutive series of cases of fetal brain anomalies associated with a phenotype typical of *COL4A1* mutation. We describe five cases associated with a pathogenic *COL4A1* mutation. Approximately 20 cases of perinatal cerebral anomalies associated with *COL4A1* mutations have been reported, mostly as isolated cases. In these cases,

Table 1 Description of five cases with fetal phenotype suggestive of cerebral injury, and *COL4A1* pathogenic mutation

Case	*GA*	*US*	*MRI*	*Outcome*	*Fetal mutation*	*Parents' status*	*Postmortem*
1	26	Subependymal and intraventricular hemorrhage with parenchymal extension	Same findings	TOP at 30 wks	*COL4A1*: c.563G>A, p.(G188E) (ACMG Class 5)	Inherited (mother)	Same findings, with: polyvisceral hemorrhagic suffusion on histological examination
2	24	Supra- and infratentorial parenchymal hemorrhage, schizencephaly	NP	TOP at 26 wks	*COL4A1*: c.4756C>A, p.(H1586N) (ACMG Class 4)	*De novo*	Same findings, with: ischemic-hemorrhagic lesions
3	22	Ischemic-hemorrhagic supratentorial lesions	Same findings, with: infratentorial hemorrhage, lesions of different ages, mild schizencephaly	TOP at 29 wks	*COL4A1*: c.1946G>A, p.(G655E) (ACMG Class 5) (karyo 47, XXX)	*De novo*	Same findings
4	22	Ischemic-hemorrhagic supratentorial lesions, polymicrogyria	Same findings, with: lesions of different ages, schizencephaly	TOP at 31 wks	*COL4A1*: c.4150+1G>A (ACMG Class 4)	*De novo*	Declined
5	32	Hemorrhagic supratentorial lesions, lesions of different ages, porencephaly	Same findings, with: schizencephaly	TOP at 35 wks	*COL4A1*: c.2008G>A, p.(G670R) (ACMG Class 5)	*De novo*	Same findings

ACMG, American College of Medical Genetics; GA, gestational age at diagnosis in weeks; karyo, karyotype; MRI, magnetic resonance imaging; NP, not performed; TOP, termination of pregnancy; US, ultrasound; wks, weeks.

Table 2 Description of four cases with fetal phenotype suggestive of cerebral injury, and variant of unknown significance in *COL4A1* or *COL4A2* (ACMG Class 3)

Case	*GA*	*US*	*MRI*	*Outcome*	*Fetal mutation*	*Parents' status*	*Postmortem*
6	38	Supratentorial porencephalic lesions	Same findings, with: hemorrhagic lesions	TOP at 39 wks	*COL4A2*: c.3258G>A, p.(A1086=)	Unknown	Same findings
7	22	Supratentorial schizencephaly	Same findings, with: hemorrhagic lesions	TOP at 25 wks	*COL4A1*: c.2642T>C, p.(M881T)	Unknown	Same findings
8*	27	Massive hemorrhage of posterior fossa	Same findings	IUD at 28 wks	*COL4A2*: c.1011G>A, p.(K337=)	Inherited (mother)	Declined
9*	32	Infratentorial hemorrhage	Same findings	IUD at 34 wks	*COL4A2*: c.1011G>A, p.(K337=)	Inherited (mother)	Same findings, with: supratentorial hemorrhagic lesions

*Siblings from two different pregnancies (both parents the same). ACMG, American College of Medical Genetics; GA, gestational age at diagnosis in weeks; IUD, intrauterine death; MRI, magnetic resonance imaging; TOP, termination of pregnancy; US, ultrasound; wks, weeks.

Table 3 Description of nine cases with fetal phenotype suggestive of cerebral injury, and with no *COL4A1* or *COL4A2* mutation

Case	*GA*	*US*	*MRI*	*Outcome*	*Postmortem/ postnatal imaging*
10	33	Subependymal and intraventricular hemorrhage with parenchymal extension	Same findings	TOP at 35 weeks	Same findings, with: signs of bleeding of different ages
11	32	Intraventricular hemorrhage with ventricular dilatation	Same findings, with: extensive supratentorial parenchymal hemorrhage, ischemic-hemorrhagic lesion of left caudate nucleus, porencephaly	TOP at 33 weeks	Same findings
12	32	Numerous porencephalic cavitations	Same findings, with: numerous ischemic-hemorrhagic lesions	TOP at 35 weeks	Same findings
13	28	Supra- and infratentorial parenchymal hemorrhagic lesions, intraventricular hemorrhage	NP	IUD at 28 weeks	Same findings, with: ischemic-hemorrhagic lesions
14	25	Ischemic-hemorrhagic lesions in left cerebellar hemisphere	Same findings	TOP at 36 weeks	Same findings, with: hemorrhagic lesions of different ages
15	29	Supratentorial intraventricular hemorrhage with parenchymal extension, porencephalic lesions	NP	TOP at 33 weeks	Same findings, with: supra- and infratentorial ischemic-hemorrhagic lesions of different ages
16	34	Intraventricular hemorrhage, supratentorial extensive parenchymal hemorrhage	NP	Live birth at 36 weeks	Same findings, with: porencephalic cavities (MRI on day 5)
17	33	Intraventricular hemorrhage with parenchymal extension	NP	TOP at 35 weeks	Same findings, with: ischemic-hemorrhagic lesions, porencephaly in basal ganglia
18	35	Large right temporal hematoma	Same findings	TOP at 38 weeks	Same findings, with: signs of bleeding of different ages

GA, gestational age at diagnosis in weeks; IUD, intrauterine death; MRI, magnetic resonance imaging; NP, not performed; TOP, termination of pregnancy; US, ultrasound.

the cerebral anomalies were frequently hemorrhagic features[6,8,12,14–16]. An association between *COL4A1* gene mutations and a prenatal diagnosis of porencephaly or schizencephaly, which are consequences of intracerebral bleeding, has also been reported[7,17,18].

Based on our experience (Case 1)[5] and data on isolated cases reported subsequently, we defined a fetal phenotype suggestive of *COL4A1* mutation. This phenotype comprises: (i) severe and/or multifocal hemorrhagic lesions; (ii) multifocal ischemic-hemorrhagic lesions. These lesions can be of different ages and associated with schizencephaly or porencephaly. Among 18 cases presenting with this type of cerebral anomaly, five cases with a pathogenic mutation in *COLA41* were identified. We did not observe any phenotypical difference between fetuses with and those without a mutation. However, cerebral anomalies were observed significantly earlier in cases with a mutation (24 *vs* 32 weeks, $P = 0.02$). This could indicate a more severe phenotype associated with an earlier cerebral insult in cases with a mutation.

We did not find any pathogenic *COL4A2* mutation in our series. *COL4A2* mutations have been reported in surviving children and adults; however, the frequency is lower than that of *COL4A1* mutations[11]. The limited size of our sample might explain why no *COL4A2* mutation was identified. Another hypothesis is that *COL4A2* mutations could be less severe. The analysis of a larger cohort of fetuses is needed to elucidate this, and testing for both genes is recommended until this point is clarified.

One feature suggestive of *COL4A1* mutation in the fetus is severe brain hemorrhagic lesions. These lesions are specifically either extensive or multifocal and can be supratentorial (intraparenchymal or intraventricular) and/or infratentorial. They can also worsen throughout pregnancy, as described in Case 3. These lesions may reflect the fragility of cerebral vessels associated with *COL4A1* mutation[19]. Another feature suggestive of *COL4A1* mutation is multifocal ischemic-hemorrhagic parenchymal lesions in supra- and/or infratentorial regions. These lesions are isolated and their incidence varies from 1% for prenatal series to 3–5% for PM series[20]. In

these cases, no etiology is found. Ischemic-hemorrhagic lesions associated with *COL4A1* mutations are typically multifocal and often extend to several cerebral territories.

Interestingly, another brain lesion that might aid in diagnosis was the observation of porencephaly or schizencephaly associated with ischemic-hemorrhagic or hemorrhagic lesions. Schizencephaly is a cortical dysplasia characterized by clefts in the cerebral mantle. Since these clefts result from a destructive process prior to neuronal migration, they are lined with dysplastic gray matter[17,21,22]. Porencephaly is due to a post-migration insult within the cerebral mantle, resulting in a fluid-filled cavity with no gray-matter lining[23]. It has been suggested that these two types of lesion result from a clastic process secondary to a vascular injury[17]. A chronological description of the cerebral insult in the context of *COL4A1* mutation has been reported in the case of a fetus in which a hyperechoic lesion in one parietal lobe, considered to be an ischemic lesion, had been noted at 19 weeks, followed by observation of another hyperechoic area at the frontotemporal level, and finally the appearance at 35 weeks of a parenchymal cleft extending to the ventricular margin[24]. In two other cases of prenatal schizencephaly, MRI showed hemosiderosis in the clefts due to hemorrhage[17]. Interestingly, in 11 of our cases, multifocal ischemic-hemorrhagic lesions and/or hemorrhagic lesions were associated with porencephaly or schizencephaly. This strongly suggests that the phenotype associated with *COL4A1* gene mutation includes schizencephaly and/or porencephaly. Thus, the performance of MRI is critical in revealing associated lesions in other cerebral territories, or features of previous hemorrhages.

In our series, second-line US examination was suggestive of a compatible *COL4A1* phenotype in all cases. MRI, however, allowed additional characterization in seven of 13 cases, in particular indicating the hemorrhagic nature of parenchymal lesions, suggesting the existence of lesions of different ages or revealing lesions missed by US.

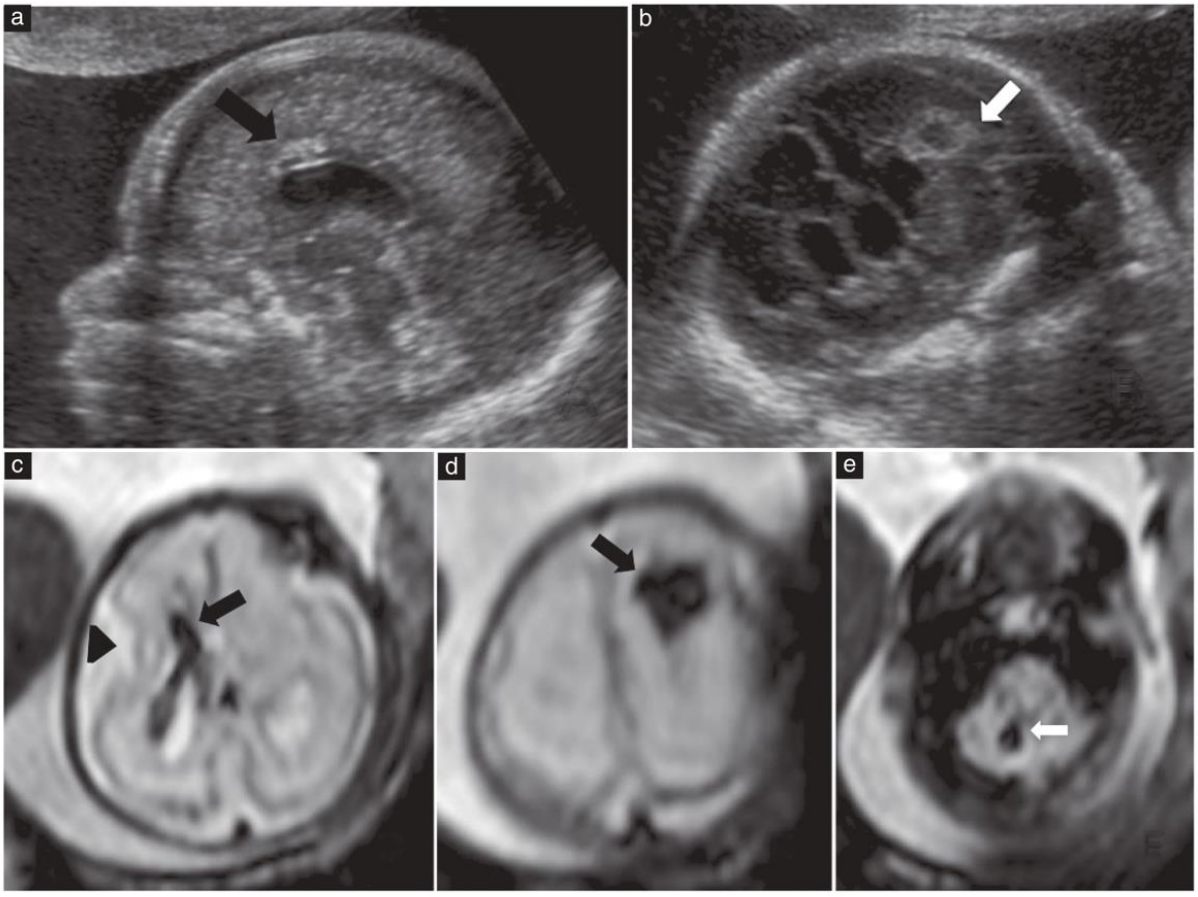


Figure 2 Cerebral imaging findings in Case 3. (a) Ultrasound image (midsagittal view) at 25 weeks, showing hyperechoic parenchymal lesion in the cingulate gyrus, above the genu of the corpus callosum (arrow). (b) Ultrasound image (coronal view) at 27 weeks, showing large hyperechoic lesion with hypoechoic center in the left frontal lobe (arrow). (c–e) T2*-weighted magnetic resonance images (axial views) at 27 weeks, demonstrating ischemic-hemorrhagic lesions, showing: (c) hypointensity next to the body of the right lateral ventricle (arrow) and enlarged Sylvian fissure (arrowhead); (d) hypointensity in the left frontal lobe (arrow); and (e) hypointensity in the inner part of the right cerebellar hemisphere (arrow).

The neurodevelopmental prognosis in our series was considered severe in all 18 cases and the parents opted for TOP in 14. The neurological outcome of children with comparable lesions is reported to be poor[12,25]. PM examination provided additional extracerebral findings in only one of 15 cases (Case 1) and additional cerebral findings in eight. Thus, similar to MRI, PM examination may provide important data suggestive of a *COL4A1* phenotype which may have been overlooked in cases with no extensive prenatal imaging. Moreover, in our series, additional findings were revealed at PM examination even in cases which had undergone prenatal MRI, which could suggest worsening of the cerebral insult during the pregnancy.

The phenotype that we defined seems appropriate, since five of the 18 fetuses carried a pathogenic mutation in *COL4A1*. In four cases, these were *de-novo* mutations which have not been reported, to our knowledge, in adult patients. This suggests that *COL4A1* mutations leading to fetal-onset manifestations might be more severe. However, additional data from a larger series of fetuses are needed to establish this hypothesis.

This series included four fetuses with a VOUS, affecting *COL4A2* in three cases and *COL4A1* in one case. Cases 8 and 9 were siblings from two different pregnancies and carried the same *COL4A2* variant, inherited from their healthy mother. In both cases, similar hemorrhagic cerebral lesions were observed. These pregnancies were both complicated by IUD. We have not been able to investigate the possible consequences of this variant on splicing in the absence of available RNA and therefore cannot conclude on its causality. The phenotypes of these fetuses could also be associated with a mutation in another gene whose identification might be possible using exome sequencing.

Our study has some limitations. First, this was a retrospective analysis in a referral center for fetal cerebral anomalies, with a limited number of cases. Second, we could not test the parents in two of four cases with VOUS, which limits further genetic characterization of these variants.

In conclusion, *COL4A1* and *COL4A2* gene mutations should be sought systematically in cases of severe and/or multifocal hemorrhagic or ischemic-hemorrhagic cerebral lesions, in particular when these anomalies are of different ages and associated with schizencephaly or porencephaly. It is likely that exome analysis would be useful in the characterization of cases carrying this phenotype.

REFERENCES

1. Mao M, Alavi MV, Labelle-Dumais C, Gould DB. Type IV Collagens and Basement Membrane Diseases: Cell Biology and Pathogenic Mechanisms. *Curr Top Membr* 2015; **76**: 61–116.
2. Plaisier E, Ronco P. COL4A1-Related Disorders. In *GeneReviews® [Internet]*, Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJ, Stephens K, Amemiya A (eds). University of Washington, Seattle, WA, 1993. http://www.ncbi.nlm.nih.gov/books/NBK7046/.
3. de Vries LS, Mancini GMS. Intracerebral hemorrhage and COL4A1 and COL4A2 mutations, from fetal life into adulthood. *Ann Neurol* 2012; **71**: 439–441.
4. Lichtenbelt KD, Pistorius LR, De Tollenaer SM, Mancini GM, De Vries LS. Prenatal genetic confirmation of a COL4A1 mutation presenting with sonographic fetal intracranial hemorrhage. *Ultrasound Obstet Gynecol* 2012; **39**: 726–727.
5. Garel C, Rosenblatt J, Moutard ML, Heron D, Gelot A, Gonzales M, Miné E, Jouannic JM. Fetal intracerebral hemorrhage and COL4A1 mutation: promise and uncertainty. *Ultrasound Obstet Gynecol* 2013; **41**: 228–230.
6. Colin E, Sentilhes L, Sarfati A, Mine M, Guichet A, Ploton C, Boussion F, Delorme B, Tournier-Lasserve E, Bonneau D. Fetal intracerebral hemorrhage and cataract: think COL4A1. *J Perinatol* 2014; **34**: 75–77.
7. Takenouchi T, Ohyagi M, Torii C, Kosaki R, Takahashi T, Kosaki K. Porencephaly in a fetus and HANAC in her father: variable expression of COL4A1 mutation. *Am J Med Genet A* 2015; **167A**: 156–158.
8. Matsumoto T, Miyakoshi K, Fukutake M, Ochiai D, Minegishi K, Tanaka M. Intracranial sonographic features demonstrating in utero development of hemorrhagic brain damage leading to schizencephaly-associated COL4A1 mutation. *J Med Ultrason (2001)* 2015; **42**: 445–446.
9. Durrani-Kolarik S, Manickam K, Chen B. COL4A1 Mutation in a Neonate With Intrauterine Stroke and Anterior Segment Dysgenesis. *Pediatr Neurol* 2017; **66**:100–103.
10. Kuo DS, Labelle-Dumais C, Gould DB. COL4A1 and COL4A2 mutations and disease: insights into pathogenic mechanisms and potential therapeutic targets. *Hum Mol Genet* 2012; **21**: R97–110.
11. Zagaglia S, Selch C, Nisevic JR, Mei D, Michalak Z, Hernandez-Hernandez L, Krithika S, Vezyroglou K, Varadkar SM, Pepler A, Biskup S, Leão M, Gärtner J, Merkenschlager A, Jaksch M, Møller RS, Gardella E, Kristiansen BS, Hansen LK, Vari MS, Helbig KL, Desai S, Smith-Hicks CL, Hino-Fukuyo N, Talvik T, Laugesaar R, Ilves P, Õunap K, Körber I, Hartlieb T, Kudernatsch M, Winkler P, Schimmel M, Hasse A, Knuf M, Heinemeyer J, Makowski C, Ghedia S, Subramanian GM, Striano P, Thomas RH, Micallef C, Thom M, Werring DJ, Kluger GJ, Cross JH, Guerrini R, Balestrini S, Sisodiya SM. Neurologic phenotypes associated with COL4A1/2 mutations: Expanding the spectrum of disease. *Neurology* 2018; **91**: e2078–88.
12. Meuwissen MEC, Halley DJJ, Smit LS, Lequin MH, Cobben JM, de Coo R, van Harssel J, Sallevelt S, Woldringh G, van der Knaap MS, de Vries LS, Mancini GMS. The expanding phenotype of COL4A1 and COL4A2 mutations: clinical data on 13 newly identified families and a review of the literature. *Genet Med* 2015; **17**: 843–53.
13. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL, ACMG Laboratory Quality Assurance Committee. 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* 2015; **17**: 405–424.
14. Giorgio E, Vaula G, Bosco G, Giacone S, Mancini C, Calcia A, Cavalieri S, Di Gregorio E, Rigault De Longrais R, Leombruni S, Pinessi L, Cerrato P, Brusco A, Brussino A. Two families with novel missense mutations in COL4A1: When diagnosis can be missed. *J Neurol Sci* 2015; **352**: 99–104.
15. Vermeulen RJ, Peeters-Scholte C, Van Vugt JJM, Van Vught JJMG, Barkhof F, Rizzu P, van der Schoor SRD, van der Knaap MS. Fetal origin of brain damage in 2 infants with a COL4A1 mutation: fetal and neonatal MRI. *Neuropediatrics* 2011; **42**: 1–3.
16. Meuwissen MEC, de Vries LS, Verbeek HA, Lequin MH, Govaert PP, Schot R, Cowan FM, Hennekam R, Rizzu P, Verheijen FW, Wessels MW, Mancini GMS. Sporadic COL4A1 mutations with extensive prenatal porencephaly resembling hydranencephaly. *Neurology* 2011; **76**: 844–846.
17. Harada T, Uegaki T, Arata K, Tsunetou T, Taniguchi F, Harada T. Schizencephaly and Porencephaly Due to Fetal Intracranial Hemorrhage: A Report of Two Cases. *Yonago Acta Med* 2017; **60**: 241–245.
18. Sato Y, Shibasaki J, Aida N, Hiiragi K, Kimura Y, Akahira-Azuma M, Enomoto Y, Tsurusaki Y, Kurosawa K. Novel COL4A1 mutation in a fetus with early prenatal onset of schizencephaly. *Hum Genome Var* 2018; **5**: 4.
19. Vahedi K, Alamowitch S. Clinical spectrum of type IV collagen (COL4A1) mutations: a novel genetic multisystem disease. *Curr Opin Neurol* 2011; **24**: 63–68.
20. Manganaro L, Bernardo S, La Barbera L, Noia G, Masini L, Tomei A, Fierro F, Vinci V, Sollazzo P, Silvestri E, Giancotti A, Marini M. Role of foetal MRI in the evaluation of ischaemic-haemorrhagic lesions of the foetal brain. *J Perinat Med* 2012; **40**: 419–426.
21. Naidich TP, Griffiths PD, Rosenbloom L. Central nervous system injury in utero: selected entities. *Pediatr Radiol* 2015; **45** (Suppl 3): S454–S462.
22. Howe DT, Rankin J, Draper ES. Schizencephaly prevalence, prenatal diagnosis and clues to etiology: a register-based study. *Ultrasound Obstet Gynecol* 2012; **39**: 75–82.
23. Eller KM, Kuller JA. Fetal porencephaly: a review of etiology, diagnosis, and prognosis. *Obstet Gynecol Surv* 1995; **50**: 684–687.
24. Khalid R, Krishnan P, Andres K, Blaser S, Miller S, Moharir M, Dlamini N. COL4A1 and fetal vascular origins of schizencephaly. *Neurology* 2018; **90**: 232–234.
25. Yoneda Y, Haginoya K, Kato M, Osaka H, Yokochi K, Arai H, Kakita A, Yamamoto T, Otsuki Y, Shimizu S, Wada T, Koyama N, Mino Y, Kondo N, Takahashi S, Hirabayashi S, Takanashi J, Okumura A, Kumagai T, Hirai S, Nabetani M, Saitoh S, Hattori A, Yamasaki M, Kumakura A, Sugo Y, Nishiyama K, Miyatake S, Tsurusaki Y, Doi H, Miyake N, Matsumoto N, Saitsu H. Phenotypic spectrum of COL4A1 mutations: porencephaly to schizencephaly. *Ann Neurol* 2013; **73**: 48–57.

