

ESTBB_25-26_GM-INBIOT_B4_683 - INS7ETDES01- ECUE Méthodes d'analyses de
données de séquençage haut débit

Next generation sequencing technologies for a successful diagnosis in a cold case of Leigh syndrome



Table des matières

I. Introduction générale du cas et du contexte scientifique	3
1. Présentation de la maladie et justification de l'étude	3
2. Particularité du cas étudié	3
3. Finalité de l'article	3
II. Objectifs précis de l'étude	4
III. Stratégie et pipeline analytique	4
1. Protocole de séquençage et filtrage bioinformatique	4
2. Analyse des variations structurelles et validation orthogonales	5
IV. Résultats moléculaires et caractérisation génotypique	5
1. Résolution du mécanisme génétique : l'hétérozygotie composite	5
2. Validation biologique et cohérence phénotypique.....	6
V. Discussion : Apport de l'approche génomique intégrée.....	6
1. Limites des analyses standards face à la complexité génomique	6
2. Limites et perspectives	6
VI. Conclusion	7
VII. Bibliographie	8

I. Introduction générale du cas et du contexte scientifique

1. Présentation de la maladie et justification de l'étude

Le syndrome de Leigh est une encéphalopathie mitochondriale infantile sévère caractérisée par une défaillance de la phosphorylation oxydative et des lésions bilatérales des noyaux gris centraux. Sa prévalence est estimée entre 1/40 000 et 1/50 000 naissances, et plus de 75 gènes — mitochondriaux ou nucléaires — ont été impliqués dans sa physiopathologie, rendant le diagnostic génétique particulièrement complexe.

Dans le cas étudié, malgré un tableau clinique et radiologique évocateur dès la première enfance, les analyses ciblées du génome mitochondrial (notamment MILS et NARP) n'ont révélé aucune mutation. Cette absence de résultat souligne les limites des approches classiques et justifie le recours à des méthodes de séquençage élargies, capables de détecter aussi bien des mutations ponctuelles que des délétions génomiques impliquées dans des formes nucléaires du syndrome de Leigh.

2. Particularité du cas étudié

Le patient présenté un tableau typique de syndrome de Leigh dès la première année de vie, avec régression psychomotrice, anomalies métaboliques (lactate et alanine élevés) et lésions bilatérales caractéristiques des noyaux gris centraux. Malgré cette présentation fortement évocatrice, aucune mutation mitochondriale connue n'a été identifiée lors des analyses initiales, notamment celles ciblant les syndromes MILS et NARP, ce qui a empêché pendant plus d'une décennie l'obtention d'un diagnostic moléculaire précis.

Cette absence de confirmation génétique, alors que le phénotype était clairement compatible avec une encéphalopathie mitochondriale, fait de ce cas un véritable cold case. La situation illustre les limites des approches ciblées classiques et souligne la nécessité d'outils génomiques plus larges, capables d'explorer simultanément les gènes nucléaires et de détecter des anomalies structurales complexes. C'est ce contexte d'impasse diagnostique qui justifie l'utilisation du séquençage d'exome et des analyses complémentaires dans cette étude.

3. Finalité de l'article

L'étude a pour objectif principal de résoudre un cas de syndrome de Leigh resté sans diagnostic moléculaire malgré de nombreuses investigations cliniques, métaboliques et génétiques réalisées sur plus de dix ans. Les analyses ciblées du génome mitochondrial n'ayant révélé aucune mutation, il devenait essentiel d'explorer

l'ensemble des gènes nucléaires pouvant être responsables de formes récessives du syndrome.

La démarche vise ainsi à démontrer l'intérêt d'une approche intégrée combinant le séquençage d'exome (WES), la détection de variations du nombre de copies (CNV) et des techniques de confirmation comme l'array-CGH et la qPCR. L'étude montre comment cette stratégie permet non seulement d'identifier une mutation pathogène dans ECHS1, mais également de mettre en évidence une délétion maternelle passée inaperçue avec les méthodes classiques, aboutissant enfin à un diagnostic complet pour un cas longtemps resté non résolu.

II. Objectifs précis de l'étude

L'objectif principal de l'étude est d'identifier la cause moléculaire du tableau clinique évocateur de syndrome de Leigh chez un patient pour lequel les analyses ciblées du génome mitochondrial, notamment celles relatives aux formes MILS et NARP, n'avaient révélé aucune mutation pathogène. Cette impasse diagnostique orientait vers une origine nucléaire et nécessitait une exploration élargie des gènes associés à la fonction mitochondriale.

L'étude vise également à montrer l'apport d'une approche génomique intégrée combinant le séquençage d'exome, la détection de variations du nombre de copies et des méthodes de confirmation cytogénétiques. Cette stratégie doit permettre de mettre en évidence des anomalies génétiques complexes comme des mutations ponctuelles, des délétions et de les interpréter de manière cohérente avec les données cliniques, radiologiques et métaboliques, afin d'aboutir à un diagnostic complet.

III. Stratégie et pipeline analytique

1. Protocole de séquençage et filtrage bioinformatique

Face à l'échec des analyses ciblées mitochondrielles (panels MILS et NARP) et compte tenu de la grande hétérogénéité génétique du syndrome de Leigh, qui peut impliquer plus de 75 gènes distincts, l'étude a adopté une stratégie de séquençage d'exome complet (Whole Exome Sequencing - WES). Cette approche non biaisée permet d'interroger simultanément l'ensemble des régions codantes du génome nucléaire, là où les panels de gènes traditionnels se révèlent limités pour des phénotypes complexes. Le séquençage a été réalisé selon une configuration en « trio », incluant le patient (malade) et ses deux parents sains. Cette méthodologie est indispensable en génétique médicale moderne car elle permet de déterminer leur origine parentale, et de filtrer efficacement les variants non pathogènes ou les erreurs de séquençage.

Les données brutes issues de la plateforme de séquençage ont été traitées par un pipeline bioinformatique. L'objectif était de réduire le bruit de fond, constitué de milliers de variants, pour isoler les candidats pertinents. Un premier filtre a été appliqué pour

exclure les polymorphismes fréquents ($MAF > 1\%$) en se basant sur des banques de données de populations contrôles telles que l'ExAC. Ensuite, une priorisation a été effectuée sur la base de l'impact fonctionnel prédit *in silico* par des algorithmes comme Polyphen2, en focalisant l'analyse sur les gènes dont la fonction est compatible avec le métabolisme mitochondrial.

2. Analyse des variations structurelles et validation orthogonales

L'analyse standard des variants nucléotidiques (SNV) a révélé une incohérence : le patient apparaissait homozygote pour une mutation pathogène du gène ECHS1, alors que seule la paternité de la mutation était confirmée, la mère étant sauvage (non porteuse). Cette discordance mendélienne a conduit à suspecter une fausse homozygotie causée par une délétion de l'allèle maternel, une anomalie invisible aux pipelines standards d'appel de variants SNV.

Pour vérifier cette hypothèse, l'équipe a intégré un outil bioinformatique spécifique nommé CeQer. Contrairement aux outils classiques qui analysent la séquence des bases, CeQer examine la profondeur de couverture ("read depth") des données d'exome pour détecter les variations du nombre de copies (CNV). Cette étape démontre l'importance de la réanalyse bioinformatique : les données brutes contenaient l'information de la délétion sous forme d'une chute de couverture, mais seul un algorithme dédié pouvait l'extraire. Enfin, pour valider biologiquement ces résultats *in silico*, des méthodes orthogonales ont été appliquées : l'hybridation génomique comparative sur microréseau (array-CGH) a été utilisée pour caractériser l'étendue chromosomique précise de la délétion, et la PCR quantitative en temps réel (RT-qPCR) a permis de confirmer la perte d'une copie du gène en comparant le dosage génique du patient à celui de son père.

IV. Résultats moléculaires et caractérisation génotypique

1. Résolution du mécanisme génétique : l'hétérozygotie composite

L'analyse des SNV a permis d'identifier une mutation faux-sens paternelle c.713C>T (p.Ala238Val) dans le gène ECHS1. Cette mutation, affectant le domaine catalytique de l'enoyl-CoA hydratase, est prédictive comme délétère pour la stabilité de la protéine. Parallèlement, l'algorithme CeQer, confirmé ultérieurement par l'array-CGH, a mis en évidence une microdélétion de 35 kb en région 10q26.3 chez le patient et sa mère. Cette délétion englobe l'intégralité du gène ECHS1 ainsi que quatre autres gènes adjacents (ZNF511, CALY, PRAP1, FUOM).

Le génotype du patient a donc été résolu comme étant une hétérozygotie composite : il associe un allèle paternel porteur de la mutation ponctuelle et un allèle maternel délété (hémizygote). La validation par RT-qPCR a corroboré ce mécanisme en

montrant une réduction de 50% du dosage génique pour les cinq gènes de la région chez la mère et l'enfant par rapport au père. Le patient ne possède donc aucune copie fonctionnelle du gène ECHS1, l'allèle paternel étant muté et l'allèle maternel étant absent.

2. Validation biologique et cohérence phénotypique

La caractérisation de ce génotype permet d'expliquer l'ensemble du tableau clinique et métabolique observé. Le gène ECHS1 code pour une enzyme mitochondriale essentielle à la beta-oxydation des acides gras et au catabolisme de la valine¹². Le déficit enzymatique résultant de l'hétérozygotie composite concorde parfaitement avec les biomarqueurs identifiés chez le patient, notamment l'élévation des taux de lactate, de valine et l'excréition urinaire d'acide 3-méthylglutaconique. De plus, les lésions des noyaux gris centraux observées à l'IRM sont caractéristiques des déficits énergétiques mitochondriaux sévères induits par le dysfonctionnement de cette voie métabolique.

V. Discussion : Apport de l'approche génomique intégrée

1. Limites des analyses standards face à la complexité génomique

Ce cas clinique illustre les limites des pipelines d'analyse génétique standard lorsqu'ils sont confrontés à des mécanismes d'hérédité complexes. Une analyse WES classique, focalisée uniquement sur les variants ponctuels (SNV) et les petites insertions/délétions (Indels), aurait conduit à une impasse diagnostique ou à une erreur d'interprétation. En effet, l'homozygote apparente du patient aurait pu être rejetée, ou bien, aurait masqué le risque de récurrence lié à la transmission maternelle de la délétion.

C'est l'intégration d'une analyse des variations du nombre de copies (CNV) directement à partir des données de séquençage qui a permis de résoudre le problème. L'étude démontre que le WES ne doit pas être considéré uniquement comme un outil de lecture de séquences, mais comme une source de données quantitatives permettant, via des algorithmes comme CeQer, de détecter des anomalies structurelles.

2. Limites et perspectives

Cette étude souligne également l'importance cruciale de la réévaluation des données anciennes pour les patients en errance diagnostique (les "Cold Cases"). Le patient présenté avait reçu un diagnostic clinique de syndrome de Leigh dès l'âge de 3 ans, mais est resté sans confirmation moléculaire pendant 15 ans. Les technologies disponibles à l'époque (Sanger ciblé, panels mitochondriaux) étaient insuffisantes pour détecter une anomalie combinant un gène nucléaire rare et une variation structurelle. La résolution de ce cas prouve que l'évolution des outils bioinformatiques permet de valoriser des données existantes pour aboutir à un diagnostic précis, indispensable

pour le conseil génétique familial et la prise en charge thérapeutique.

VI. Conclusion

La résolution de ce cas clinique, resté en impasse diagnostique pendant quinze ans, montre la puissance de l'approche génomique intégrée dans l'élucidation des maladies rares. Le diagnostic final de syndrome de Leigh par déficit en ECHS1 repose sur l'identification d'une hétérozygotie composite complexe, associant une mutation faux-sens paternelle (c.713C>T) et une microdélétion maternelle de 35 kb en 10q26.3. Ce résultat démontre que le séquençage d'exome (WES) ne doit plus être considéré uniquement comme une méthode de lecture de variants nucléotidiques, mais comme une plateforme de données multidimensionnelle nécessitant une exploitation bioinformatique approfondie pour révéler les variations structurelles (CNV).

VII. Bibliographie

Article analysé

1. **Aretini P, Mazzanti CM, La Ferla M, et al.** Next generation sequencing technologies for a successful diagnosis in a cold case of Leigh syndrome. *BMC Neurology*. 2018;18:99. doi:10.1186/s12883-018-1103-7.

Références complémentaires

2. **UniProt Consortium.** UniProtKB - P30084 (ECHS1_HUMAN).
<https://www.uniprot.org/uniprotkb/P30084/entry>
3. **CeQer (Outil de détection des CNV)** : Piazza R, Magistroni V, Redaelli S, et al. CeQer: a graphical tool for copy number aneuploidy detection from whole-exome sequencing data.
4. **ExAC (Base de données de fréquences alléliques)** : Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536(7616):285-291.
5. **Polyphen-2 (Outil de prédition de pathogénicité)** : Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7(4):248-249.