Université de Bourgogne

UFR des Sciences de Santé

Circonscription Médecine

ANNEE 2017

N°

Le séquençage d'exome dans le diagnostic des maladies rares avec anomalies du développement: intérêt de la réanalyse des données annuelle prospective

THESE

présentée

à l'UFR des Sciences de Santé de Dijon
Circonscription Médecine

et soutenue publiquement le

06 Janvier 2017

pour obtenir le grade de Docteur en Médecine

par NAMBOT Sophie

Née le 21 Décembre 1987

A Besançon

Université de Bourgogne

UFR des Sciences de Santé

Circonscription Médecine

ANNEE 2017

N°

Le séquençage d'exome dans le diagnostic des maladies rares avec anomalies du développement: intérêt de la réanalyse des données annuelle prospective

THESE

présentée

à l'UFR des Sciences de Santé de Dijon
Circonscription Médecine

et soutenue publiquement le

06 Janvier 2017

pour obtenir le grade de Docteur en Médecine

par NAMBOT Sophie

Née le 21 Décembre 1987

A Besançon

Université de Bourgogne

UFR des Sciences de Santé

Circonscription Médecine

Année Universitaire 2016-2017
au **1er Janvier 2017**

Doyen :	**M. Frédéric HUET**
1er Assesseur **:**	M. Yves ARTUR
Assesseurs :	Mme Laurence DUVILLARD
	M. Pablo ORTEGA-DEBALLON
	M. Marc MAYNADIE

PROFESSEURS DES UNIVERSITES – PRATICIENS HOSPITALIERS

			Discipline
M.	Marc	**BARDOU**	Pharmacologie clinique
M.	Jean-Noël	**BASTIE**	Hématologie - transfusion
M.	Emmanuel	**BAULOT**	Chirurgie orthopédique et traumatologie
M.	Laurent	**BEDENNE**	Gastroentérologie et hépatologie
M.	Yannick	**BEJOT**	Neurologie
M.	Alain	**BERNARD**	Chirurgie thoracique et cardiovasculaire
M.	Jean-François	**BESANCENOT**	Médecine interne
Mme	Christine	**BINQUET**	Epidémiologie, économie de la santé et prévention
M.	Bernard	**BONIN**	Psychiatrie d'adultes
M.	Philippe	**BONNIAUD**	Pneumologie
M.	Alain	**BONNIN**	Parasitologie et mycologie
M.	Bernard	**BONNOTTE**	Immunologie
M.	Olivier	**BOUCHOT**	Chirurgie cardiovasculaire et thoracique
M.	Belaid	**BOUHEMAD**	Anesthésiologie - réanimation chirurgicale
M.	Alexis	**BOZORG-GRAYELI**	ORL
M.	Alain	**BRON**	Ophtalmologie
M.	Laurent	**BRONDEL**	Physiologie
M.	François	**BRUNOTTE**	Biophysique et Médecine Nucléaire
M.	Patrick	**CALLIER**	Génétique
M.	Jean-Marie	**CASILLAS-GIL**	Médecine physique et réadaptation
Mme	Catherine	**CHAMARD-NEUWIRTH**	Bactériologie - virologie; hygiène hospitalière
M.	Pierre-Emmanuel	**CHARLES**	Réanimation
M.	Pascal	**CHAVANET**	Maladies infectieuses
M.	Nicolas	**CHEYNEL**	Anatomie
M.	Alexandre	**COCHET**	Biophysique et médecine nucléaire
M.	Luc	**CORMIER**	Urologie
M.	Yves	**COTTIN**	Cardiologie
M.	Charles	**COUTANT**	Gynécologie-obstétrique
M.	Gilles	**CREHANGE**	Oncologie-radiothérapie
Mme	Catherine	**CREUZOT-GARCHER**	Ophtalmologie
M.	Frédéric	**DALLE**	Parasitologie et mycologie
M.	Serge	**DOUVIER**	Gynécologie-obstétrique
Mme	Laurence	**DUVILLARD**	Biochimie et biologie moléculaire
Mme	Laurence	**FAIVRE-OLIVIER**	Génétique médicale
Mme	Patricia	**FAUQUE**	Biologie et Médecine du Développement
Mme	Irène	**FRANCOIS-PURSSELL**	Médecine légale et droit de la santé
M.	Pierre	**FUMOLEAU**	Cancérologie
M.	François	**GHIRINGHELLI**	Cancérologie
M.	Claude	**GIRARD**	Anesthésiologie – réanimation chirurgicale
M.	Vincent	**GREMEAUX**	Médecine physique et réadaptation
M.	Frédéric	**HUET**	Pédiatrie
M.	Pierre	**JOUANNY**	Gériatrie

M.	Denis	**KRAUSÉ**	Radiologie et imagerie médicale
M.	Sylvain	**LADOIRE**	Histologie
M.	Gabriel	**LAURENT**	Cardiologie
M.	Côme	**LEPAGE**	Hépato-gastroentérologie
M.	Romaric	**LOFFROY**	Radiologie et imagerie médicale
M.	Luc	**LORGIS**	Cardiologie
M.	Jean-Francis	**MAILLEFERT**	Rhumatologie
M.	Cyriaque Patrick	**MANCKOUNDIA**	Gériatrie
M.	Sylvain	**MANFREDI**	Hépato-gastroentérologie
M.	Laurent	**MARTIN**	Anatomie et cytologie pathologiques
M.	David	**MASSON**	Biochimie et biologie moléculaire
M.	Marc	**MAYNADIE**	Hématologie - transfusion
M.	Thibault	**MOREAU**	Neurologie
M.	Klaus Luc	**MOURIER**	Neurochirurgie
Mme	Christiane	**MOUSSON**	Néphrologie
M.	Paul	**ORNETTI**	Rhumatologie
M.	Pablo	**ORTEGA-DEBALLON**	Chirurgie Générale
M.	Jean-Michel	**PETIT**	Endocrinologie, diabète et maladies métaboliques
M.	Lionel	**PIROTH**	Maladies infectieuses
Mme	Catherine	**QUANTIN**	Biostatistiques, informatique médicale
M.	Jean-Pierre	**QUENOT**	Réanimation
M.	Patrick	**RAT**	Chirurgie générale
M.	Jean-Michel	**REBIBOU**	Néphrologie
M.	Frédéric	**RICOLFI**	Radiologie et imagerie médicale
M.	Paul	**SAGOT**	Gynécologie-obstétrique
M.	Emmanuel	**SAPIN**	Chirurgie Infantile
M.	Henri-Jacques	**SMOLIK**	Médecine et santé au travail
M.	Éric	**STEINMETZ**	Chirurgie vasculaire
Mme	Christel	**THAUVIN**	Génétique
M.	Pierre	**VABRES**	Dermato-vénéréologie
M.	Bruno	**VERGÈS**	Endocrinologie, diabète et maladies métaboliques
M.	Narcisse	**ZWETYENGA**	Chirurgie maxillo-faciale et stomatologie

PROFESSEURS EN SURNOMBRE

M.	Roger	**BRENOT**	(surnombre jusqu'au 31/08/2018)
M.	Philippe	**CAMUS**	(surnombre jusqu'au 31/08/2019)
Mme	Monique	**DUMAS-MARION**	(surnombre jusqu'au 31/08/2018)
M.	Maurice	**GIROUD**	(surnombre jusqu'au 21/08/2018)
M.	Pierre	**TROUILLOUD**	(surnombre du 05/02/2014 au 31/08/2017)

Université de Bourgogne

UFR des Sciences de Santé

Circonscription Médecine

MAITRES DE CONFERENCES DES UNIVERSITES
PRATICIENS HOSPITALIERS DES DISCIPLINES MEDICALES

			Discipline Universitaire
M.	Sylvain	**AUDIA**	Médecine interne
Mme	Shaliha	**BECHOUA**	Biologie et médecine du développement
Mme	Marie-Claude	**BRINDISI**	Nutrition
M.	Jean-Christophe	**CHAUVET-GELINIER**	Psychiatrie, psychologie médicale
		(Mobilité Novembre 2016 à 2017)	
M.	Alexis	**DE ROUGEMONT**	Bactériologie-virologie ; hygiène hospitalière
M.	Hervé	**DEVILLIERS**	Médecine interne
M.	Olivier	**FACY**	Chirurgie générale
Mme	Ségolène	**GAMBERT-NICOT**	Biochimie et biologie moléculaire
Mme	Françoise	**GOIRAND**	Pharmacologie fondamentale
Mme	Agnès	**JACQUIN**	Physiologie
M.	Alain	**LALANDE**	Biophysique et médecine nucléaire
M.	Louis	**LEGRAND**	Biostatistiques, informatique médicale
Mme	Stéphanie	**LEMAIRE-EWING**	Biochimie et biologie moléculaire
M	Maxime	**SAMSON**	Médecine interne
		(Mobilité Novembre 2016 à 2017)	
M.	Benoit	**TROJAK**	Psychiatrie d'adultes ; addictologie
M.	Paul-Mickaël	**WALKER**	Biophysique et médecine nucléaire

PROFESSEURS EMERITES

M.	Jean	**CUISENIER**	(01/09/2014 au 31/08/2017)
M.	Jean	**FAIVRE**	(01/09/2012 au 31/08/2018)
M.	Marc	**FREYSZ**	(01/09/2016 au 28/02/2017)
M	Philippe	**GAMBERT**	(01/09/2014 au 31/08/2017)
M.	Patrick	**HILLON**	(01/09/2016 au 31/08/2019)
M.	François	**MARTIN**	(01/09/2015 au 31/08/2018)
M.	Pierre	**POTHIER**	(01/09/2015 au 31/08/2018)

PROFESSEURS DES UNIVERSITES DE MEDECINE GENERALE

M.	Jean-Noël	**BEIS**	Médecine Générale

PROFESSEURS ASSOCIES DE MEDECINE GENERALE

M.	Didier	**CANNET**	Médecine Générale
M.	Gilles	**MOREL**	Médecine Générale
M.	François	**MORLON**	Médecine Générale

MAITRES DE CONFERENCES ASSOCIES DE MEDECINE GENERALE

M.	Clément	**CHARRA**	Médecine Générale
M.	Rémi	**DURAND**	Médecine Générale
M.	Arnaud	**GOUGET**	Médecine Générale
Mme	Anne	**WALDNER-COMBERNOUX**	Médecine Générale

Université de Bourgogne

UFR des Sciences de Santé

Circonscription Médecine

MAITRES DE CONFERENCES DES UNIVERSITES

M.	Didier	**CARNET**	Anglais
M.	Jean-Pierre	**CHARPY**	Anglais
Mme	Catherine	**LEJEUNE**	Pôle Epidémiologie
M.	Gaëtan	**JEGO**	Biologie Cellulaire

PROFESSEURS DES UNIVERSITES

Mme	Marianne	**ZELLER**	Physiologie

PROFESSEURS AGREGES de L'ENSEIGNEMENT SECONDAIRE

Mme	Marceline	**EVRARD**	Anglais
Mme	Lucie	**MAILLARD**	Anglais

PROFESSEURS CERTIFIES

Mme	Anaïs	**CARNET**	Anglais
M.	Philippe	**DE LA GRANGE**	Anglais
Mme	Virginie	**ROUXEL**	Anglais (Pharmacie)

PROFESSEURS DES UNIVERSITES – PRATICIENS HOSPITALIERS DES DISCIPLINES PHARMACEUTIQUES

Mme	Evelyne	**KOHLI**	Immunologie
M.	François	**GIRODON**	Sciences biologiques, fondamentales et cliniques

MAITRES DE CONFERENCES DES UNIVERSITES PRATICIENS HOSPITALIERS DES DISCIPLINES PHARMACEUTIQUES

M.	Mathieu	**BOULIN**	Pharmacie clinique
M.	Philippe	**FAGNONI**	Pharmacie clinique
M.	Frédéric	**LIRUSSI**	Toxicologie
M.	Marc	**SAUTOUR**	Botanique et cryptogamie
M.	Antonin	**SCHMITT**	Pharmacologie

Université de Bourgogne

UFR des Sciences de Santé

Circonscription Médecine

L'UFR des Sciences de Santé de Dijon, Circonscription Médecine, déclare que les opinions émises dans les thèses qui lui sont présentées doivent être considérées comme propres à leurs auteurs, et qu'elle n'entend ne leur donner ni approbation, ni improbation.

COMPOSITION DU JURY

Président : Madame le Professeur Laurence OLIVIER-FAIVRE, service de Génétique médicale, CHU Dijon

Membres : Madame le Professeur Christel THAUVIN-ROBINET, service de Génétique médicale, CHU Dijon, Directeur de thèse

Monsieur le Professeur Patrick CALLIER, Laboratoire de Cytogénétique et Biologie moléculaire, CHU Dijon

Monsieur le Professeur Pierre VABRES, service de Dermatologie, CHU Dijon

Monsieur le Docteur Julien THEVENON, CHU Grenoble

REMERCIEMENTS

Je souhaite remercier Laurence, Christel et Julien de m'avoir proposé un si beau sujet de thèse et de m'avoir accordé leur confiance et leur soutien dans la réalisation de ce travail.

Je remercie Laurence de s'impliquer autant dans ma formation et de me pousser sans cesse à me dépasser, en me proposant de participer à de nombreux projets. J'admire ton envie d'en faire toujours plus et ta façon de travailler, tout en restant accessible, à l'écoute et humaine

Je remercie Christel pour son encadrement, sa disponibilité et sa bienveillance. Merci pour ton aide et tes conseils dans ce travail. J'admire ta rigueur, ta capacité de travail et ta prévenance

Je remercie Julien pour sa disponibilité, sa gentillesse et son accompagnement au cours de ses trois années d'internat. Ce travail est aussi le tien et j'espère que tu en es fier. J'ai toujours pu compter sur toi et j'ai beaucoup de chance d'avoir pu profiter de ton enseignement avant ton départ dans les montagnes.

Je remercie Patrick de faire partie de mon jury, mais aussi pour sa bienveillance, sa bonne humeur et son enseignement

Je remercie le Pr Vabres d'avoir accepté de faire partie du jury de ma thèse et de son encadrement au cours de mes six premiers mois d'internat

Je remercie toute l'équipe du service de génétique clinique et du laboratoire, pour leur accueil chaleureux et leur aide au cours de ces trois années. C'est un bonheur de travailler avec des personnes aussi dynamiques et agréables, toujours prêtes à rendre service.

Je remercie particulièrement Justine, Tiphaine et Sonia qui m'ont grandement aidé dans ce travail en cherchant et en réussissant à dénicher le moindre dossier

Je remercie Claudine, Sylvie, Sonia, Annick pour leur aide et efficacité, et qui ne me disent jamais non lorsque j'ai besoin d'elles

Je remercie Daphné pour sa gentillesse et ses talents de pédagogue, ainsi qu'Amandine et Caroline de m'avoir encadré et accompagné lors de mes toutes premières consultations

Je remercie Yannis et Emilie d'avoir toujours pris le temps de me sortir de situations bioinformatiques « épineuses », ainsi qu'Ange-Line pour sa disponibilité et son aide

A mes co-internes Mathilde, Julian, et maintenant Aurélien, sur qui je peux compter et avec qui j'ai plaisir à partager ses années d'internat

A ma mère, cette force de la nature, qui m'a appris l'importance du travail et la persévérance, soutenue par son amour et son attention dans les moments difficiles, gâtée à outrance dans les moments de joie. C'est grâce à toi si j'en suis là aujourd'hui et je ne pourrais jamais assez te remercier. Mon admiration et mon amour pour toi sont incommensurables et inconditionnels

A mon père, la force tranquille, qui a toujours cru en moi sans jamais me mettre la pression, avec cette même petite phrase avant chaque examen « fais de ton mieux », et qui a toujours été aux petits soins pour moi pendant toutes ces années de révision sans fin. Je serai toujours là pour toi, comme tu l'as été pour moi

A mon frère, que j'aime et que j'admire pour ce qu'il est et ce qu'il fait. Pourvu que notre complicité dure malgré la distance et nos modes de vie bien différents, et que l'on pourra toujours compter l'un sur l'autre

A Sylvain, qui m'a toujours soutenu et s'est toujours impliqué dans ma formation, en me téléchargeant à ses risques et périls tous les livres de médecine qu'il pouvait, et m'a donné de précieux conseils pour mes études et ma vie professionnelle future. Même si je le montre peu, j'ai beaucoup de respect pour toi et te remercie pour ce que tu fais pour moi

A mes grands-parents, qui m'ont permis à plusieurs reprises de garder la motivation quand je voyais la fierté dans leurs yeux

A toute ma famille, Nambot et Verdot, pour votre soutien et votre présence aujourd'hui. C'est toujours un plaisir de vous retrouver et de partager des moments avec vous

A tous les amis de mes parents qui me font l'honneur d'être présents aujourd'hui, que j'adore et qui m'ont toujours encouragé dans mes études. A Josiane et Claude, mes « deuxièmes parents » que j'aime

énormément et qui ont toujours veillé sur moi. A Bernard, sans qui je n'en serai peut-être pas là aujourd'hui, sans son aide pour passer le cap de la première année de médecine

A Clara, ma partenaire pendant l'externat, dont le grain de folie m'a fait tellement de bien pendant ces trois longues années de révision intense, à tous ces fous rires et ces soirées que l'on a partagées

A JD

A Julia, Justine, Mathilde, Camille, JF et mes amis de Besançon, que je vois de moins en moins souvent mais que je retrouve comme si c'était hier à chaque fois, et avec qui j'ai passé, et j'espère je passerai encore, des moments mémorables

A mes copains de l'internat et de Dijon, avec qui j'ai découvert les joies et travers de la vie d'interne. Particulièrement, Myriam et Rym, qui sont maintenant des amies chères et font partie intégrante de ma vie, avec qui je rigole et me sent tellement bien. J'ai de la chance que vous soyez venues à Dijon, même si vous pensez le contraire. A Sandra et Jean Charles, avec qui j'ai découvert les joies de la dermatologie, qui sont aussi devenus des amis proches et je l'espère pour longtemps

A Florence, ma petite bouillabaisse, qui m'a soutenu et aidé à tenir dans les dernières lignes droites de ce travail, et qui me fait le plaisir d'être là aujourd'hui, accompagnée de Vincent que je remercie aussi

A mes nouveaux copains de la SIGF, les internes en génétique, ces individus étranges que j'apprécie plus que tout et que je suis ravie de retrouver plusieurs fois par an, pour de studieuses semaines de cours, évidemment

SERMENT D'HIPPOCRATE

"Au moment d'être admis(e) à exercer la médecine, je promets et je jure d'être fidèle aux

lois de l'honneur et de la probité.

Mon premier souci sera de rétablir, de préserver ou de promouvoir la santé dans tous ses éléments, physiques et mentaux, individuels et sociaux.

Je respecterai toutes les personnes, leur autonomie et leur volonté, sans aucune discrimination selon leur état ou leurs convictions.

J'interviendrai pour les protéger si elles sont affaiblies, vulnérables ou menacées dans leur intégrité ou leur dignité.

Même sous la contrainte, je ne ferai pas usage de mes connaissances contre les lois de l'humanité.

J'informerai les patients des décisions envisagées, de leurs raisons et de leurs conséquences.

Je ne tromperai jamais leur confiance et n'exploiterai pas le pouvoir hérité des circonstances pour forcer les consciences.

Je donnerai mes soins à l'indigent et à quiconque me les demandera.

Je ne me laisserai pas influencer par la soif du gain ou la recherche de la gloire.

Admis(e) dans l'intimité des personnes, je tairai les secrets qui me seront confiés. Reçu(e) à l'intérieur des maisons, je respecterai les secrets des foyers et ma conduite ne servira pas à corrompre les mœurs.

Je ferai tout pour soulager les souffrances. Je ne prolongerai pas abusivement les agonies. Je ne provoquerai jamais la mort délibérément.

Je préserverai l'indépendance nécessaire à l'accomplissement de ma mission. Je n'entreprendrai rien qui dépasse mes compétences. Je les entretiendrai et les perfectionnerai pour assurer au mieux les services qui me seront demandés.

J'apporterai mon aide à mes confrères ainsi qu'à leurs familles dans l'adversité.

Que les hommes et mes confrères m'accordent leur estime si je suis fidèle à mes promesses ; que je sois déshonoré(e) et méprisé(e) si j'y manque."

LISTE DES SIGLES ET ABREVIATIONS

ACMG: American College of Medical Genetics and Genomics

ACPA: Analyse Chromosomique sur Puce à ADN

AD: Autosomal Dominant

AR: Autosomal Recessive

Bp: Base Pair

BWA: Burrows-Wheeler Aligner

CcuB: Centre de Calcul de l'Université de Bourgogne

CNG: Centre National de Génotypage

CNV: Copy Number Variation

DA : Developmental anomalies

DNA: Deoxyribonucleic Acid

EE: Epileptic encephalopathy

GATK: Genome Analysis Toolkit

GC: Guanine Cytosine

HPO: Human Phenotype Ontology

ID : Intellectual Disability

WES: Whole-Exome Sequencing

WGS: Whole-genome sequencing

MAF: Minor Allele Frequency

MCA: Multiple Congenital Anomalies

NCBI: National Center for Biotechnology Information

NGS: Next Generation Sequencing

PCR: Polymerase Chain Reaction

PMD: Psychomotor Delay

QC: Quality Control

SNP: Single Nucleotide Polymorphism

SNV: Single Nucleotide Variation

VUS: Variant of Unknown Significance

Le séquençage d'exome dans le diagnostic des maladies rares avec anomalies du développement: intérêt de la réanalyse des données annuelle prospective

INTRODUCTION

Les anomalies du développement et la déficience intellectuelle (AD/DI) représentent un vaste groupe hétérogène de pathologies, impliquant plus de 3000 entités cliniques différentes. Alors que chacune est individuellement rare, elles concernent collectivement 1 à 3% des naissances dans le monde [1]. Les anomalies du développement sont des malformations liées à une anomalie lors de l'embryogenèse, déterminées majoritairement par des facteurs génétiques, les distinguant des deux autres types d'anomalies congénitales, à savoir les déformations et les disruptions. La déficience intellectuelle se caractérise par une altération précoce et durable des fonctions cognitives, motrices et du langage, impactant le niveau global d'intelligence et le fonctionnement social de l'individu. Les causes peuvent être génétiques, acquises et/ou environnementales. Ces troubles chroniques de début précoce sont responsables de handicaps lourds, entraînant une forte morbi-mortalité et d'importantes dépenses de santé publique [2].

La plupart des AD/DI d'origine génétique suivent un mode d'hérédité mendélien, mais présentent une très grande hétérogénéité génétique, c'est à dire que plusieurs types de variations de l'ADN, survenant dans différents gènes, peuvent être la cause de ces pathologies. Actuellement, une grande partie de leurs bases moléculaires reste encore inconnue, représentant un véritable défi diagnostique dans le domaine de la génétique médicale. En effet, établir le diagnostic étiologique de ces maladies rares est essentiel pour la mise en place d'un suivi adapté du patient, la prévention de certaines complications,

l'instauration d'un traitement ou l'inclusion dans des essais cliniques [3]. Il est aussi indispensable pour fournir un conseil génétique fiable, en évaluant le risque de récidive à chaque future grossesse selon le mode de transmission, et éventuellement proposer un diagnostic prénatal.

La démarche diagnostique actuelle des AD/DI repose sur l'expertise clinique, des analyses cytogénétiques telles que l'analyse chromosomique sur puce à ADN (ACPA) et la recherche du syndrome X-fragile, et éventuellement l'analyse de gènes ciblés selon l'orientation phénotypique et l'histoire familiale. Mais cette démarche longue et éprouvante pour le patient et sa famille, s'apparentant à une succession de consultations auprès de plusieurs spécialistes dans différents centres, et d'examens paracliniques souvent invasifs, laisse près de la moitié des patients sans diagnostic [4].

Depuis 2009, le séquençage nouvelle génération (NGS) a révolutionné le monde de la génétique médicale en étendant ses perspectives dans la résolution moléculaire des maladies génétiques rares. En effet, les séquenceurs haut débit sont aujourd'hui capables de lire plusieurs millions de séquences ADN en parallèle en un temps réduit, et permettent de s'affranchir d'un certain nombre de biais et contraintes spécifiques de la méthode Sanger (précédent gold-standard dans la lecture du code génétique), tels que le clonage de l'ADN à séquencer et l'individualisation des clones. Toutes les régions codantes du génome, appelées exons, peuvent désormais être séquencées en quelques heures au cours d'un seul examen. Il est donc possible de détecter quasiment tous les variants nucléotidiques, mais aussi depuis peu, les variants structuraux tels que les délétions et les duplications de l'ADN, couverts par l'analyse (90 à 95% des régions codantes). Ces variations peuvent être localisées dans des gènes connus en pathologie humaine, mais aussi dans des gènes non associés à une pathologie humaine, ou même dont la fonction est inconnue. L'immense volume de données produites

par ces nouvelles techniques a dû être appréhendé par les chercheurs et les cliniciens, et les règles de leur interprétation a nécessité quelques années avant d'être clairement établies.

C'est pourquoi cette technique a initialement été utilisée dans le domaine de la recherche pour des pathologies ultrarares, résistantes aux techniques diagnostiques traditionnelles, concernant des groupes homogènes de patients. Cette approche de séquençage non biaisée a démontré un taux de succès sans précédent dans l'identification de gènes pathogènes [5-7]. L'exome a ensuite été appliqué dans de grandes séries de patients présentant des maladies plus fréquentes, avec une grande hétérogénéité moléculaire et des spectres phénotypiques larges, telle que la DI syndromique [8,9], le retard de développement [10], l'autisme [11,12], l'épilepsie [13,14] ou les anomalies cardiaques congénitales [15,16].

Dans un second temps, une interprétation plus précise des données et une réduction des coûts de séquençage a permis son implantation dans le domaine diagnostique. Cela a permis la délimitation de plus de 500 nouvelles associations génotype-phénotype entre 2010 et 2015, plaçant l'exome comme examen de choix pour le diagnostic de maladies rares avec suspicion d'hérédité mendélienne [17]. En effet, le taux diagnostique actuel de l'exome est compris entre 25 et 32%, selon la pathologie et l'hétérogénéité de la cohorte étudiées [18-22]. Ce taux diagnostique correspond à l'identification d'un variant pathogène dans un gène précédemment impliqué en pathologie humaine, avec des corrélations génotype-phénotype publiées. Mais dans certains cas, des preuves moléculaires ou phénotypiques peuvent manquer pour conclure à la pathogénicité d'un variant, aboutissant à un résultat non concluant ou négatif. Par exemple, pour un variant faux-sens détecté chez un patient avec un phénotype concordant avec la pathologie associée au gène d'intérêt, mais dans lequel seuls des variants troncants ont été rapportés, ou dans le cas où une unique famille atteinte a été décrite dans la littérature. Le partage international de données, à travers des projets tels que MatchMaker

Exchange [23], permettent d'augmenter rapidement la récurrence par l'identification de patients supplémentaires avec des variants candidats dans un même gène.

Tous ces progrès ont mené inévitablement à une explosion des connaissances, avec plus de deux cents nouveaux gènes par an, que le clinicien ne peut suivre et assimiler [24,25]. Les méta-analyses de grands groupes de patients atteints de AD/DI n'ayant pu obtenir un diagnostic grâce à l'exome ont souligné l'importance de réévaluer les données de séquençage étant donné la croissance rapide des connaissances. Dans la littérature, très peu d'articles évaluent la pertinence d'une analyse et interprétation périodiques des données d'exome de larges séries diagnostiques. L'Université de Standford a été la première a rapporté son expérience chez 40 patients ayant reçu un résultat négatif après une première analyse diagnostique en solo. Les auteurs ont obtenu 10% de taux diagnostique additionnel, mettant en évidence le potentiel de cette stratégie [26]. Tous les diagnostics établis après réanalyse étaient liés à des publications récentes, survenues quelques mois après la première analyse.

Au-delà de l'intérêt diagnostique, l'exome offre un atout économique majeur. En 2016, Monroe et al. [27] ont rapporté que l'approche diagnostique traditionnelle dans le domaine de la déficience intellectuelle était quatre fois plus coûteuse que la stratégie du séquençage d'exome en trio (analyse du cas index et de ses parents). Ceci a été confirmé par plusieurs études médico-économiques mettant toutes en évidence la meilleure capacité de l'exome à établir un diagnostic moléculaire accompagnée d'une réduction du délai diagnostique et du nombre d'examens paracliniques, parfois non pertinents et invasifs [28-30].

Cette étude a pour but d'évaluer la faisabilité et l'intérêt de l'exome diagnostique pour les maladies avec AD/DI dans notre pratique clinique quotidienne, ainsi que de la réanalyse annuelle systématique des données. Nous rapportons les résultats moléculaires et cliniques de 418 patients, inclus de Juin 2013 à Juin 2016, ayant tous bénéficié d'une analyse d'exome en

solo. Les données de 156 patients, ayant reçu un résultat non positif après l'analyse initiale, ont été réanalysées de manière prospective, après mise à jour du pipeline informatique et de la littérature. La détection des variants structuraux, possible à partir de Novembre 2015, a été réalisée pour tous les patients. Le taux diagnostique selon la présentation phénotypique, le délai diagnostique et le nombre d'examens complémentaires par année ont également été collectés et rapportés.

ORIGINAL ARTICLE

TITLE

Clinical whole-exome sequencing for the diagnosis of rare disorders with developmental anomalies: substantial interest of prospective annual reanalysis

AUTHORS

Sophie Nambot (1,2,3)*, Julien Thevenon (1,3,4)*, Paul Kuentz (2,3,4), Yannis Duffourd (3,4), Emilie Tisserant (3,4), Ange-Line Bruel (3,4), Anne-Laure Mosca-Boidron (2), Alice Masurel-Paulet (1,3), Daphné Lehalle (1), Nolwenn Jean-Marçais (1,3), Mathilde Lefebvre (1,2), Pierre Vabres (3,4), Salima El Chehadeh-Djebbar (5), Orphanomix physicians' group (6), Judith St-Onge (2,4), Nadège Gigot (2,3,4), Thibaud Jouan (2,3,4), Martin Chevarin (2,3,4), Charlotte Poé (2,3,4), Virginie Carmignac (4), Antonio Vitobello (2,3,4), Patrick Callier (2,3,4), Jean-Baptiste Rivière (2,3,4), Laurence Faivre (1,3,4)[#], Christel Thauvin-Robinet (1,2,3,4)[#].

* equally contributed to the work

[#] equally codirected the work

1. Centre de Génétique et Centre de référence «Anomalies du Développement et Syndromes Malformatifs», Hôpital d'Enfants, Centre Hospitalier Universitaire de Dijon, Dijon, France.
2. Laboratoire de Génétique Moléculaire, Plateau Technique de Biologie, Centre Hospitalier Universitaire de Dijon, Dijon, France.
3. Fédération Hospitalo-Universitaire Médecine Translationnelle et Anomalies du Développement (FHU TRANSLAD), Centre Hospitalier Universitaire de Dijon et Université de Bourgogne-Franche Comté, Dijon, France.

4. Equipe d'Accueil 4271, Génétique des Anomalies du Développement, Université de Bourgogne, Dijon, France.
5. Service de génétique médicale, Centre Hospitalier Universitaire de Strasbourg, Strasbourg, France
6. Jean-Luc Alessandri, Yline Capri, Marie-Pierre Cordier-Alex, Christine Coubes, Salima El Chehadeh, François Feillet, Mélanie Fradin, David Geneviève, Marion Gérard-Blanluet, Dominique Germain, Brigitte Gilbert-Dussardier, Cyril Goizet, Alice Goldenberg, Delphine Heron, Marie-Line Jacquemont, Aurelia Jacquette, Elodie Lacaze, Didier Lacombe, Fanny Laffargue, Laëtitia Lambert, Béatrice Laudier, Marine Lebrun, Isabelle Maystadt, Fanny Morice-Picard, Sébastien Moutton, Laurence Perrin, Florence Petit, Christophe Philippe, Fabienne Prieur, Pierre Sarda, Elise Schaefer, Renaud Touraine, Marjolaine Willems

Key words: whole-exome sequencing, developmental anomalies, intellectual disability, datasharing, reanalysis

Address for correspondence:

Dr Julien Thevenon, Pr Christel Thauvin-Robinet and Pr Laurence Faivre

Centre de Génétique, Hôpital d'Enfants

14 rue Gaffarel

21079 Dijon Cedex

Tel: +33 380 295 313 Fax: +33 380 293 266

Emails: julien.thevenon@chu-dijon.fr; christel.thauvin@chu-dijon.fr; laurence.faivre@chu-dijon.fr

ABSTRACT

Purpose

Developmental anomalies and intellectual disability (DA/ID) represent a great diagnostic challenge in genetic since up to half of patients remain molecularly undiagnosed after traditional genetic evaluations and a long and stressful diagnostic "odyssey". To enhance the diagnostic yield of our generalist consulting genetic center, we developed clinical whole-exome sequencing (WES) in solo strategy.

Methods

To evaluate the effectiveness of this strategy, we report on clinical and molecular findings of 418 consecutive patients with DA/ID who benefited from WES in solo strategy, with a systematic annual reanalysis of the updated data with negative or non-conclusive result.

Results

We obtained a total diagnostic yield of 31%. The mean diagnostic yield per prospective annual reanalysis was 7.5%, represented mainly by pro-active international datasharing allowing the identification of 5 new disease-causing genes. The mean number of prior negative tests clearly decreased after the implementation of WES in our clinical routine practice, leading to subsequent cost reductions in etiological investigations.

Conclusion

Our results show the great clinical effectiveness of the WES implementation in daily clinical practice in the field of DA/ID, even in small clinical centers.

INTRODUCTION

Developmental anomalies and intellectual disability (DA/ID) represent a vast and heterogeneous group of disorders, involving more than 3.000 different clinical and genetic entities. While each is individually rare, they may collectively affect 1-3% of living birth worldwide [1]. Most of the DA/ID are of genetic origin and are supporting by Mendelian inheritance. Because the prevalence of each disorder is individually low and a large portion of DA/ID molecular bases is still unresolved, their diagnosis remains challenging. These chronic and early onset disorders significantly contribute to morbidity, mortality and healthcare expenses [2]. An etiologic diagnosis for rare disorders is essential for genetic counselling, prenatal testing, accurate follow-up, complications prevention, adapted treatment or clinical trial inclusion [3]. The current standard of care for diagnosis of DA/ID includes multiple clinical evaluations by highly specialized physicians, and countless paraclinical investigations such as imaging and biological tests, which can be invasive for the patients. The genetic investigations include cytogenetic tests and successive single-gene testing, and more recently gene-panels that need regular updates. This long and tedious traditional approach leaves approximately half of the families with no diagnosis [4].

Next generation sequencing (NGS) revolutionized medical genetics by expanding its perspectives for the molecular resolution of rare genetic diseases. Initially applied in research, different strategies were considered to achieve this goal. From tailored sequencing projects aiming at discovering the molecular bases of a recognizable syndrome in a homogeneous group of patients to the systematic application of pan-genomic sequencing in large heterogeneous cohorts: whole-exome sequencing (WES) has demonstrated an unprecedented success rate in disease-causing genes identification. [5-7]. Among other heterogeneous disorders, DA/ID in various categories highlighted the usefulness of an unbiased sequencing

approach for syndromic ID [8,9], developmental delay [10], autism [11,12], epilepsy [13,14], or congenital heart defects [15,16].

Later, a more accurate interpretation of the data and a reduction of sequencing costs enabled its large implementation in clinical practice. Between 2010 and 2015, the delineation of more than 500 new genotype-phenotype descriptions placed WES as the current standard of care for the diagnosis of highly heterogeneous rare disorders with suspected Mendelian inheritance [17]. WES is a powerful diagnostic tool that blurs the line between diagnosis and research. The widespread application of this test in non-specific cohorts of patients with DA/ID allows a diagnostic yield of 25 to 32% [18-22]. This diagnostic yield corresponds to the identification of a disease-causing variant in a gene previously implicated in a human disorder and with published genotype-phenotype correlations. Although the diagnostic yield should not vary, the likelihood of identifying a candidate variant for a new disorder may depend on the phenotype and the sequencing strategy (trio-based or proband-based WES) [20,31]. But molecular or phenotypic evidences could lack to conclude: (1) in patients with missense variants and a phenotype concordant with the pathology related to the mutated gene, but in which only causal truncating variants have been reported; (2) when patients carry the same type of variant reported as pathogenic in a gene associated to a pathology, but with an atypical phenotype; (3) when only one family with gene mutations is described in the literature; or (4) for affected girls with X-linked disorder, in which only boys have been reported. International datasharing systems, such as the MatchMaker exchange initiative [23], catalyze the identification of additional patients with candidate variants in a same gene. This allows fast and accurate reverse phenotyping to assess the clinical relevance of the candidate variants and genes.

Meta-analyses of large groups of patients with DA/ID and a negative WES interpretation have recently highlighted the critical importance of re-assessing the

interpretation of WES data in the light of scientific advances [24,25]. In the literature, very few articles assess the relevance of a periodic analysis and interpretation of WES data from large series of diagnostic exome. Stanford University team has been the first to report its experience in 40 patients with an additional diagnostic yield of 10% that gives an interesting preview of the potential of this strategy [26].

This study aims at assessing the feasibility and usefulness of a yearly reanalysis and interpretation of data from patients with DA/ID and a negative clinical interpretation of a proband-based WES.

PATIENTS AND METHODS

Considering the large distress and expectance of numerous undiagnosed patients, the growing power of WES in identifying genetic causes of rare diseases, and the massive costs for health institutions of successive etiological investigations in patients with DA/ID, diagnostic WES was implemented in Dijon University Hospital in clinical practice in 2013 for patients with rare diseases in Burgundy, and for undiagnosed patients with strong demand of genetic counselling in France.

Patients

From June 2013 to June 2016, we proposed WES to 325 patients referred for etiological diagnosis to the Reference Center for Developmental Anomalies and Malformative Syndromes of Dijon and to 93 patients addressed by French external centers to the Orphanomix service (see URL) (Supplemental Figure 1a).

The inclusion criteria were: (i) a patient with at least one congenital anomaly with or without ID, of suspected genetic origin, (ii) with a negative prior diagnostic work-up, and (iii) an informed consent for the inclusion of the family and proband when feasible. Foetuses with polymalformative associations were included in another research project especially focused on WES in the prenatal field. Before inclusion, every patient had a diagnostic work-up including visceral malformation screened by imaging studies, and a variable number of biological and metabolic investigations depending on the clinical orientation. Array-comparative genomic hybridization (array-CGH) had systematically previously been performed for patients with developmental delay, isolated or syndromic ID (associated to dysmorphism or one congenital malformation), autism spectrum disorders, pre or post-natal malformations (≥2), and the characterization of an anomaly detected by another cytogenetic method. In case of suspected diagnostic etiology, a targeted genetic test (single-gene or gene

panel) was first ordered. However, in some cases, WES could be directly ordered. This decision, based on the sensibility, the time of response and the cost, was discussed between the different physicians of the center on weekly clinic-biological meetings.

WES was proposed to the patient and parents/guardians in a dedicated pre-test consultation by the referring physician, who described the aims, benefits and limitations of the test, the different possible results with explanation of result of uncertain significance, and the possibility to opt in or out of receiving the results of secondary findings in the 56 genes following the ACMG recommendations [32]. Patients and/or parents and/or legal guardians signed a specific informed consent for WES mentioning anonymized medical datasharing for diagnostic purpose, with a dedicated section for secondary findings [33]. The local ethical committee approved this process.

Standardized deep-phenotyping

The patients were separated in three major phenotypic groups according to the clinical indication: neurodevelopmental disorders, neuromuscular disorders, and DA without ID/psychomotor delay (PMD). The neurodevelopmental disorders group was divided in four subgroups: non syndromic ID, syndromic ID defined as ID with dysmorphism and/or malformation(s), epileptic encephalopathy, and syndromic PMD when the age of the patient did not permit to differentiate ID and learning difficulties. Detailed phenotypic data were anonymously collected in the PhenomeCentral database (see URL) using the HPO terms.

Collection of prior negative tests

For each patient, all previous diagnostic tests ordered for etiological investigations by the physicians before WES were retrospectively collected. It included cytogenetic tests, and molecular genetic tests (array-CGH, single-gene or gene panels tests) but also imagery,

sensorial, electrophysiological, biological, metabolic and histological analyses. Tests performed for the medical follow-up were excluded. The dates of ordering, realization and result of each test were also collected, with the aim to compare the time to diagnosis before and after the complete integration of clinical WES in solo in our routine practice in June 2014.

Whole exome sequencing

Sequencing and bioinformatics analysis

In all index cases, librairies of genomic DNA samples were prepared using the Agilent Sureselect Human All Exon v5 kit, and were sequenced on a HiSeq2000 (Illumina) according to the manufacturer's recommendation for paired-end 76 bp reads. Bioinformatic pipeline, alignement processes and quality procedures are detailed in Supplemental methods.

Interpretation strategy

In first step, the rare variants affecting the reading-frame and responding to the quality criteria (see Supplemental methods) were prioritized if reported as pathogenic or probably pathogenic in public variant databases of clinical interest (Clinvar, see URLS) or as affecting a gene associated with a human disorder referenced in OMIM (See URLs) or manually curated via Pubmed **(**See URLs).

In second step, for patients with variants in non human-disease gene but presenting potential evidences of pathogenicity, a translational research approach was used in order to identify strong candidate genes. The interpretation approach is detailed in the Supplemental methods and represented in the Supplemental Figure 1b.

Annual reanalysis

Reanalysis of the negative and uncertain results was realized from the original raw sequencing data stored in fastq folders. Softwares used for variant detection were updated on a yearly basis, and databases used for the annotation on a quarterly basis. The in-house

pipeline has been optimized to enable the CNV detection in November 2015. This analysis was retrospectively applied to all patients, including positive results. SNV and CNV analysis and interpretation are described in the Supplemental methods.

RESULTS

Patients

423 patients were included between June 1st, 2013 and June 30th, 2016. Five patients were excluded because of failed quality control. 325 patients (78%) were received to the Reference Center for Developmental Anomalies and Malformative Syndromes of Dijon by 7 physicians, and blood samples of 93 patients (22%) were addressed by 33 physicians from 18 French institutions. The 418 patients included 172 females (41%) and 246 males (59%) (sex ratio=1.4), with 84.2% children (352 patients) and 15.8% adults (66 patients). The mean age at inclusion was 6.7 years for children and 30.7 for adults (global average: 10.5 years). Thirty-seven patients were born from consanguineous parents. The proportion of patients in each phenotypic category was as described in Figure 1.

Molecular results

Primary molecular results

Among the 418 patients, 82/418 (19.5%) were analyzed during the year 1, 119/418 (28.5%) during the year 2 and 217/418 (52%) during the year 3. Initial diagnostic yield varied from 22% in the year 1 to 28% in the year 3. Variants in strong candidate genes were identified in 13/82 patients (15%) in the year 1, 9/119 patients (7.5%) in the year 2, and 19/217 patients (9%) in the year 3 (Figure 2).

Annual reanalysis

Annual reanalysis in patients with non-positive diagnosis was performed in 156/418 patients, including 64/156 and 92/156 patients referred in year 1 or 2, respectively. After first reanalysis, additional diagnostic yield ranged from 6 to 8.4%. The 59/64 patients from year 1 with persistent non-positive diagnosis benefited from a second reanalysis 24 months after the initial analysis, harboring an additional diagnostic yield at 8.5% (Figure 2).

Finally, among the 156 reanalysed patients, 22 results became positive (15 after reanalysis at 12 months, and 7 at 24 months). On these 22 positive results, 14 were initially negative, and 8 were initially non conclusive because of lack of recurrence in the literature or because of a missing second event in a recessive disorder (*CLN3*). Ten cases have been resolved through international datasharing and publications of our team or collaboration with another team, allowing the identification of 5 novel genes (*KCNA2 [36], FIBP [37], AP3B2 [38], SLC13A5 [39],* and *MAB21L1 [40]*). Seven cases were linked to recent publications in the literature, 4 variants were missed at the first analysis, and 1 case was resolved by the detection of a CNV (Supplemental Table 1).

Overall the 418 patients, 128 received a positive molecular diagnosis, with a total diagnostic yield of 31%. The Mendelian disease patterns included autosomal dominant (61%), autosomal recessive (27%), and X-linked (12%) (Figure 3). Of the 78 autosomal dominant conditions diagnosed, 64 (82%) arose as a result of a *de novo* mutation, 10 (13%) were inherited, and 4 (5%) were undetermined due to lack of parental samples. Among the 34 autosomal recessive disorders, there were 16 (47%) cases of homozygosity, and 18 (53%) cases of compound heterozygosity. Of the 15 X-linked disorders, 8 (53%) were linked to a *de novo* mutation, and 7 (47%) were inherited. The 147 pathogenic or probably pathogenic variants included missense (42%), frameshift (24%), nonsense (18%), splicing variant (10%), small in frame insertion/deletion (2%), silencing variant (1%), and CNVs (3%). The 5

pathogenic or probably pathogenic CNVs included only deletions ranging from 0.63 to 12 Mb (Supplemental Table 2). The total of 147 pathogenic or likely pathogenic SNVs or CNVs occurred in 102 different genes.

In addition, 41/418 patients (9.8%) remained with variants in strong candidate genes related to their phenotype. The modes of inheritance included 24 autosomal dominant, 13 autosomal recessive, and 4 X-linked. The 50 variants including 32 missense, 8 nonsense, 5 splicing variants, 2 frameshift, 2 in-frame in/del, and 1 silencing variant (Figure 3).

Six of the 418 patients had medically actionable secondary finding in one of the 56 ACMG genes (Supplemental Table 2), including 3 *BRCA2,* 1 *BRCA1,* 1 *KCNQ1* and 1 *SCN5A* variants. 2/6 patients harbored a secondary finding without an etiological diagnosis.

1 patient was a healthy carrier of a pathogenic variant implicated in a recessive disorder (Supplemental Table 2).

Phenotypic distribution and diagnostic yield per phenotypic group

Over the three years, we observed a global similar repartition of the phenotypic groups with a majority of patients presenting neurodevelopmental disorders, concordant with our cohort's recruitment (Figure 1). The distribution inside this group showed a clear increase of the syndromic ID group (22.4%), associated to a strong decrease of the epileptic encephalopathy (EE) group (27.4%). This initial over-representation of EE group can be explained by a work realized in our center during the first year of inclusion focused on diagnosis of EE patients by WES [19].

The highest global diagnostic yield after two reanalyses was obtained for the neuromuscular disorders group (33%), followed by the neurodevelopmental disorders group (31%) (Figure 4). For the neurodevelopmental disorders group, we observed an increase of

the positive results at each reanalysis, with a clear improvement for the syndromic ID and epileptic encephalopathy groups (Supplemental Figure 2).

Time to diagnosis and etiological investigations

The time to diagnosis was calculated only for the patients with a positive result seen in our local center. It corresponded to the time of global diagnostic process, delimited by the first consultation in our center and the date of WES report. The median time to diagnosis was about 54 months for the first year, 48 months for the second and 45 months for the last year, traducing a global reduction of 9 months over the three years (Figure 5).

The mean number of tests for all local patients also significantly decreased over the three years. However, the strongest decrease was observed in the metabolic and standard biological categories, followed by the imaging and molecular genetics categories (Figure 5).

DISCUSSION

The recent high acceleration in gene discovery makes difficult for the physicians to follow genetic medical knowledge. The WES routine use first demonstrated the several limitations of usual phenotype-driven strategies based on clinical diagnosis, despite the famous clinical expertise of physicians in a reference center for rare disease: (1) atypical presentations of known diseases making difficult the diagnosis at first sight; (2) ultra-rare diseases described in only few cases and therefore unknown to most specialists, (3) patients exhibiting a rare disease with a specific, but only recently discovered phenotype. These limitations pushed us to adapt our clinical practice and diagnostic process, with the set-up of systematic reanalysis and international datasharing.

In this study, we report the molecular results of 418 patients, who benefited from clinical WES in solo strategy, with the results of systematic annual and prospective reanalysis

of data from patients with non-positive diagnosis. Among the 156 initially undiagnosed patients, 22 (14%) received a molecular diagnosis by one or two successive WES reanalyses. The majority of the results turned positive thanks to a secondary research reanalysis and active international datasharing through databases as Matchmaker exchange or PhenomeCentral, leading to a publication by our team or collaboration with another team [36-39, 41-43] (Supplemental Table 1). Indeed, to face the lack of recurrence in case of ultra-rare disorders, it is essential that the scientific community provide efforts to share their molecular and clinical data with standardized terms and methods, so that the access and the submission from the clinical laboratory become easier. But this active research of additional patients through international collaborations induces a strong work, taking time and needing the support of a research team inside the center. This translational integrated organization explains why we obtained a high global diagnostic yield for the first year through successive reanalyzes, because we had time to collect sufficient molecular and phenotypic evidences to conclude on the pathogenicity of the candidate variants. Recent publications of previously unknown genes, revealed by current reviews of the literature and pipeline updates, were the second cause of positive result after reanalysis. The latter principally concerned the patients of the second year, traducing the recent and exponential evolution of genetic discoveries [17].

So, despite absence of additional costs for sequencing, the organization of systematic reanalysis led to a sizeable workload at different steps of the process: (i) the bioinformaticians update the in-house pipeline and reanalyzed the raw data, (ii) the technicians confirm causal variants by additional validating technics, and (iii) the physicians provide accurate and updated phenotypic data by regular consultations or re-convening, participate to international datasharing and collaborations, and to the final discussion by reverse phenotyping. This multidisciplinary interpretation was set up in order to offer the best possible analysis with a minimal of missed pathogenic variants. This annual reanalysis strategy is a great advantage

for patients, but had to be integrated, and valued, to the current activity of the clinical and laboratory teams.

Aware of the strong capacities of clinical WES and easier access to this technology, we progressively position it more precociously in the diagnosis strategy, and proposed it in a more prospective way to new addressed patients, without diagnostic delay and prior extensive work up. The implementation of diagnostic WES in current practice has been effective at the end of the first year of this study, concordant with a significant reduction of the time to diagnosis. Indeed, between the first and the second year, the median duration from the first consultation to diagnosis has been reduced of 6 months, against 3 months between the second and the third year, showing the high impact of this practice's change. This also has been traduced by a significant reduction of the mean number of etiological tests, in particular metabolic, standard biological, imaging and molecular genetics tests, limiting consecutive substantial costs. This could obviously be explained by earlier use of WES, but also because our diagnostic reasoning initially based on a phenotype-driven approach has been disrupted. In case of ultrarare or atypic phenotype, reverse phenotyping is an important tool to help in the interpretation of WES results, by an accurate reexamination of the clinical data, in order to confront them to the phenotype described in the literature. The use of this genotype-driven approach, in complement with the traditional phenotype-driven approach, allows a more efficient ordering of complementary exams. For example, one of the patients benefited from a negative extensive etiological work-up of 57 exams before WES detected a homozygous truncating variant in the *SPR* gene, responsible for the Dopa-responsive dystonia due to sepiapterin reductase deficiency (MIM #612716). This result was concordant with the patient's partial phenotype, which was still uncharacteristic at this age (6 years) [44]. The result was delivered to the family, immediately followed by L-DOPA treatment, which greatly improved the neuromuscular signs of the patients. Thus, even if the WES price could

still seem elevated regarding other technics and targeted genetic tests, its early implementation in the diagnostic strategy, the genotype-driven approach, its constant decreasing cost and the availability of the data over the patient's lifetime allowing reanalyses, made that WES more advantageous. This has been also illustrated by several medico-economic studies, concerning subjects suspected of genetic condition, all highlighting the better capacity of WES to establish a molecular diagnosis at a lower cost than the classical technics [28-30].

For the 31% of cases who received a positive molecular diagnosis, this step ends a diagnostic "odyssey". Next to this end, molecular diagnostic provided precious medical care information and accurate care to the patient and his family, as already described. Novel therapeutic options or clinical care recommendations were introduced following WES positive diagnosis in 9 patients. For example, a cetogenic diet was started in a girl carrying a variation in *SLC2A1* (GLUT1 deficiency syndrome 1 (MIM #606777)), stopping seizures and clearly improving the psychomotor development. In an another boy presenting epileptic encephalopathy, microcephaly and elevated hepatic transaminases, 2 compound heterozygous variants in the *NGLY1* gene were identified as responsible for a congenital disorder of deglycosylation (MIM #615273). The patient could be included in a survey and phenotypic description protocol, set up in aim to develop a targeted therapeutic, and the family could benefited of a pre-implantation diagnosis.

Secondary findings in the 56 ACMG genes were identified in 1.4 % patients in concordance with the estimation that 1 to 3% of adult's patients undergoing WES have actionable pathogenic or likely pathogenic variants. Results were delivered to the patients and their family, when consent was given, allowing specific preventive medical care and follow-up, as well as familial screening.

In the remaining negative cases, WES prospective reanalysis should be continued because it is likely that the causal variant(s) may be localized in human disease genes unknown at this time [17], but it is most likely that some key elements are still present in the raw data generated by sequencing. The progressive technological evolution from WES to whole-genome sequencing (WGS) will made discuss WGS implementation and other genomics technologies since a smaller part of molecular etiologies include epigenetic alterations or causal variant(s) that may be localized in non-coding regions, such as regulatory or deep intronic regions, or in coding regions not well covered by WES (about 7% of the coding regions), or that repetitive sequences and pseudogenes crowd and scramble the analysis and the interpretation of the raw data. WGS will so resolve some limitations and will probably also lead to novel discoveries in understanding the mechanisms of Mendelian diseases.

In conclusion, our study shows that WES can be used as a powerful diagnostic test for patients with DA and/or ID, limiting time to diagnosis and extensive etiological investigations. The exponential explosion of genetic discoveries rapidly increases its diagnostic yield, comforting the high interest of prospective reanalysis of WES data in undiagnosed patients. These convergent results make discuss the place of WES in first intention in the diagnostic process of DA/ID, especially as solo strategy is at least efficient than trio, and as the detection of CNV is now possible with the computational performances. But the borders between diagnostic and research yields become increasingly fine because of the substantial interest of prospective pipeline update and reanalysis, the intensive international datasharing, the phenotypic spectrum enlargement of known human diseases, and the rapid identification of novel causal genes, requiring a translational integrated organization from diagnosis to research. Finally, if WES offers a better depth of coverage for

coding regions, which are actually our first interest in diagnostic practice, WGS is so the next genetic challenge in a very close future with new necessary inherent changes. But, if the diagnostic contribution is huge as one of the WES, it is definitively worth it for the patients and their families.

ACKNOWLEDGMENTS

We wish to thank the patients and their families involved in the study, and the physicians who addressed their patients to our center. We also thank the University of Burgundy Centre de Calcul (CCuB) for technical support and management of the informatics platform and the Regional Council of Burgundy and the Centre Hospitalo-Universitaire de Dijon for financial support. Finally, we thank the Centre National de Génotypage (CNG) for the sequencing support and the Biomedicine Agency for structural variation software XHMM.

AUTHOR'S CONTRIBUTION

JT, AMP, DL, NJM, CTR, LOF and Orphanomix physicians' group contributed to patient recruitment and phenotyping. SN, ML, JD and AS contributed to the record of phenotypic data in Phenome Central. PC and ALMB contributed to the structural variations characterization. JstO, NG, MC, CP, and TJ performed the molecular laboratory work. YD, ET and JBR performed the bioinformatics analysis. JT, JBR, PK, and ALB interpreted exome data. CTR and JT designed the study. All the authors contributed to the writing and review of the paper.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

URLs

ClinGen: https://www.clinicalgenome.org

ClinVar: www.ncbi.nlm.nih.gov/clinvar

DGV: http://dgv.tcag.ca/dgv/app/home

Ensembl: http://www.ensembl.org

EVS: http://evs.gs.washington.edu

ExAC Browser: https://exac.broadinstitute.org

GeneReviews®: www.ncbi.nlm.nih.gov/books

IGV: www.broadinstitute.org/igv/

Integrative Genome Viewer: www.broadinstitute.org/igv/

LOVD: http://databases.lovd.nl/shared

Matplotlib python librairy: www.matplotlib.org

NGLY1.org: http://www.ngly1.org

NHLBI GO Exome Sequencing Project

OMIM: www.omim.org

Orphanomix: http://www.orphanomix.com/index.html

PhenomeCentral: https://phenomecentral.org/

Pubmed: https://www.ncbi.nlm.nih.gov/pubmed/

RefSeq: http://www.ncbi.nlm.nih.gov/refseq/

UCSC browser: www.genome.ucsc.edu

TITLES AND LEGENDS TO FIGURES

Figure 1: Overview of phenotypic distribution in the 418 patients

Figure 2: Diagnostic yield evolution regarding prospective reanalysis

Figure 3: Distribution of Mendelian mode of inheritance and variant type for cases with positive molecular diagnosis, strong candidate genes and secondary findings

Figure 4: Overall molecular diagnostic yield of phenotypic groups

Figure 5: Overview of the time to diagnosis in months and etiological investigations in the local patients

Supplemental figures and tables

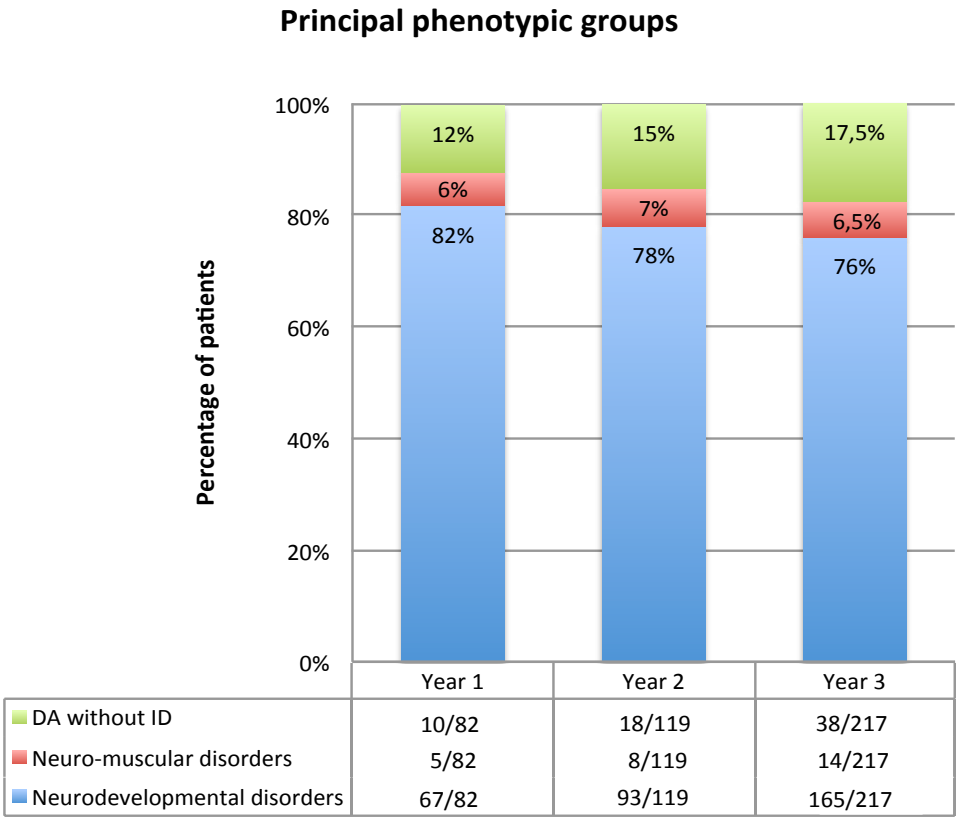
Supplemental Figure 1a: Representations of Patient's workflow

Supplemental Figure 1b: Representations of interpretation approach

Supplemental Figure 2: Detailed molecular diagnostic yield of phenotypic neurodevelopmental disorders group

Supplemental Table 1: Diagnosis obtained by WES reanalysis

Supplemental Table 2: List of positive, VUS and secondary findings variants



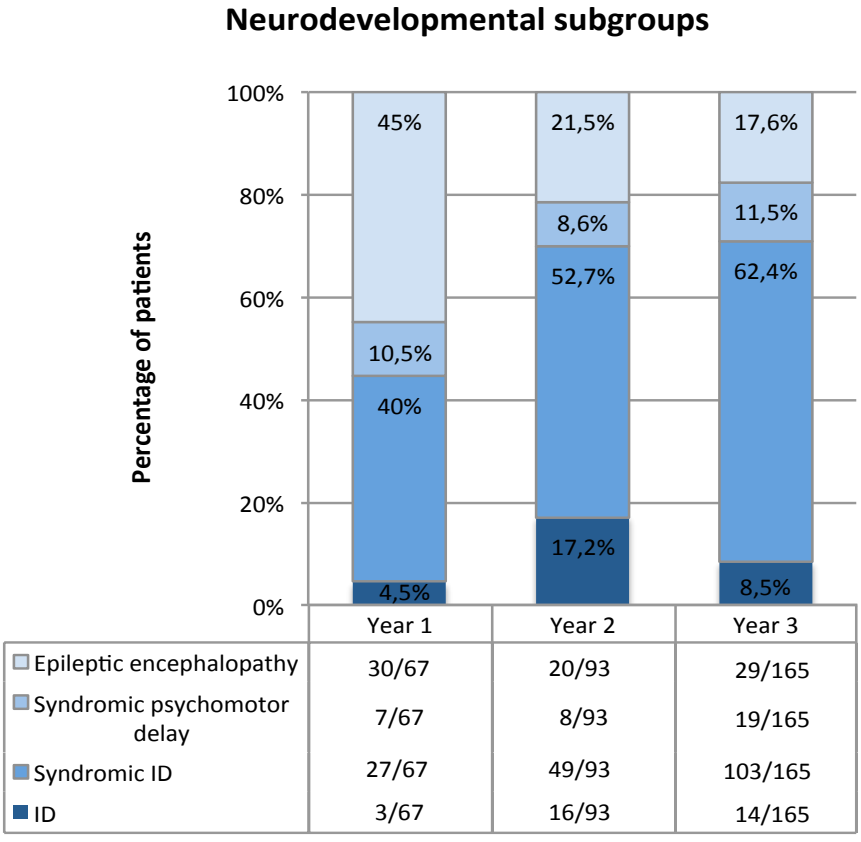


Figure 1 : Overview of phenotypic distribution in the 418 patients

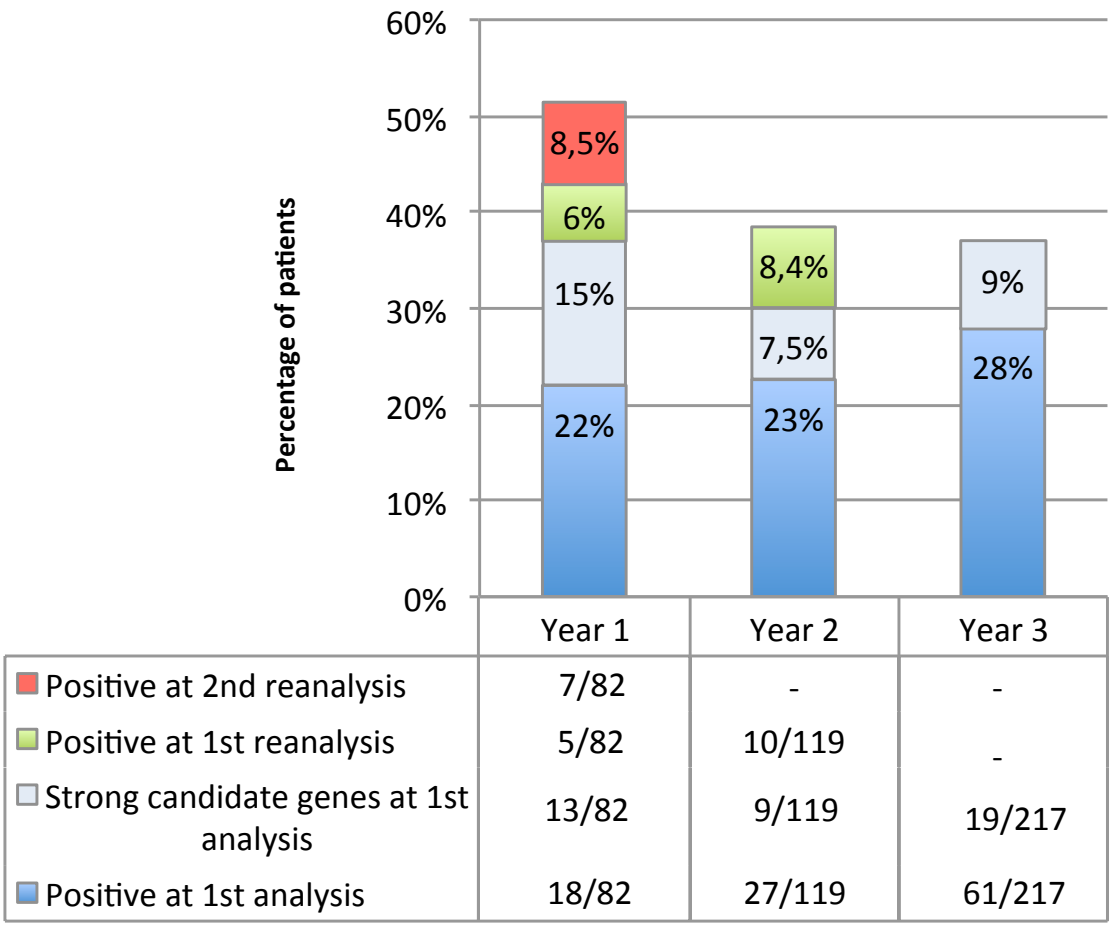


Figure 2 : Diagnostic yield evolution regarding prospective reanalysis

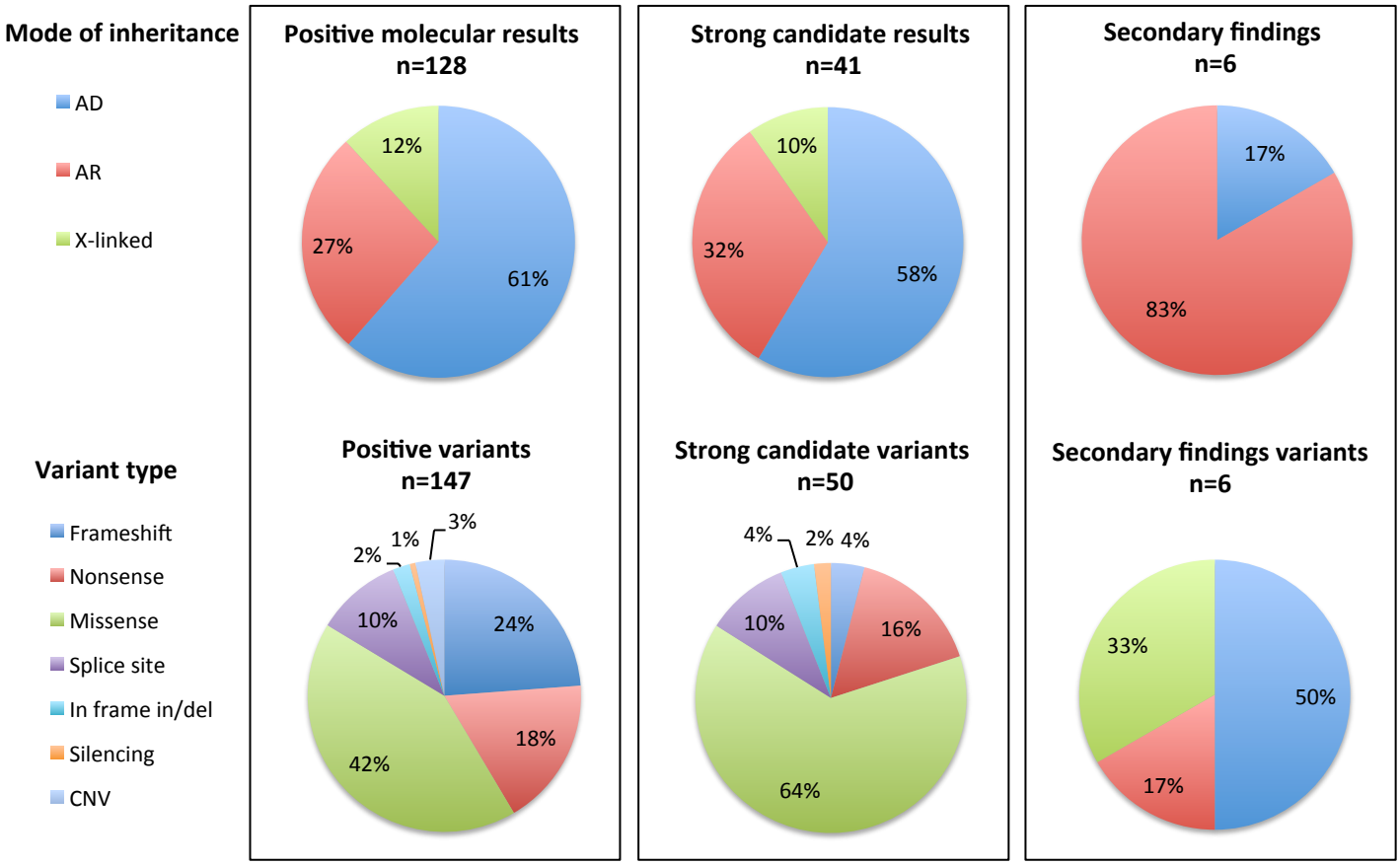


Figure 3 : Distribution of Mendelian mode of inheritance and variant type for cases with positive molecular diagnosis, strong candidate genes and secondary findings

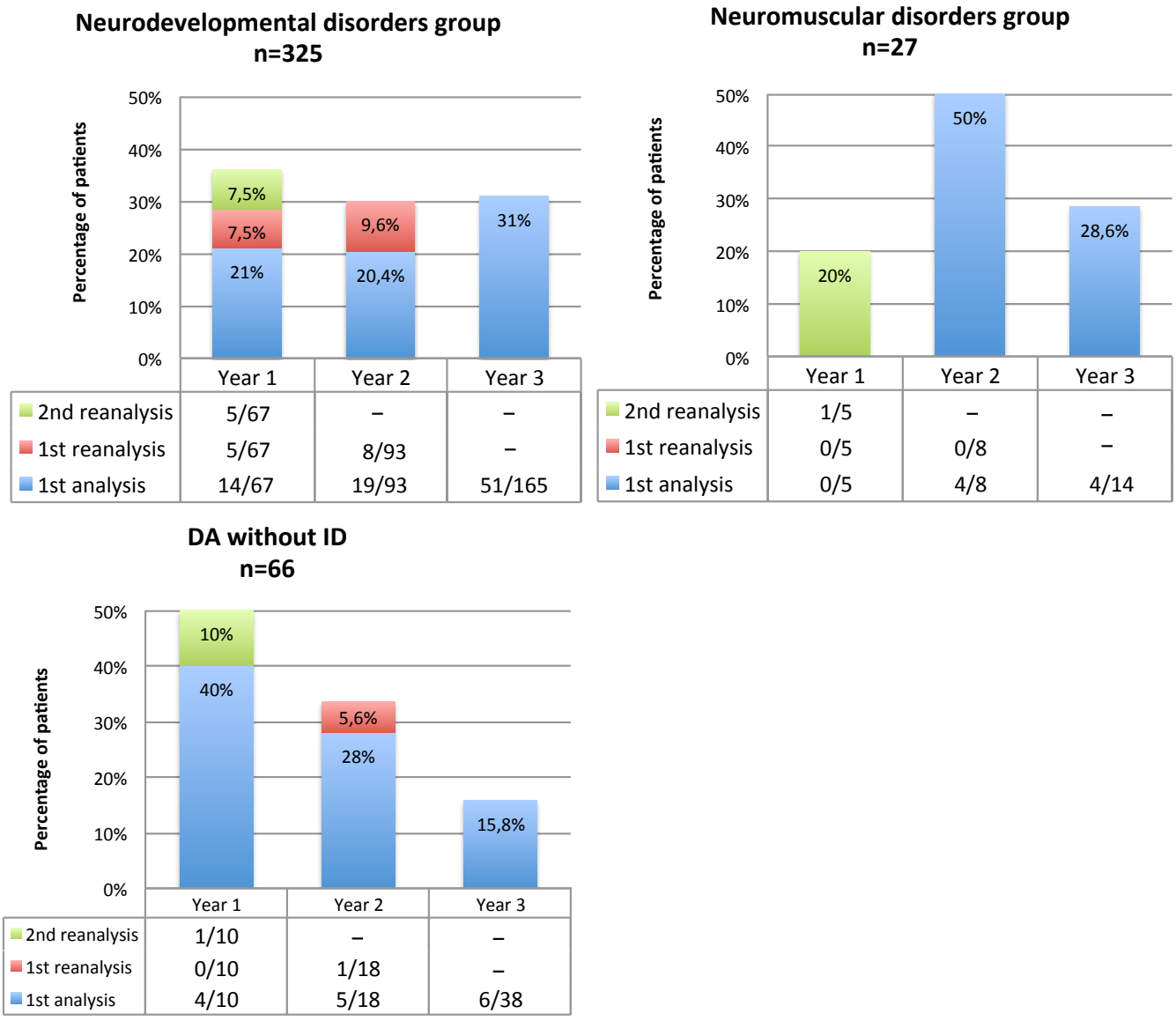
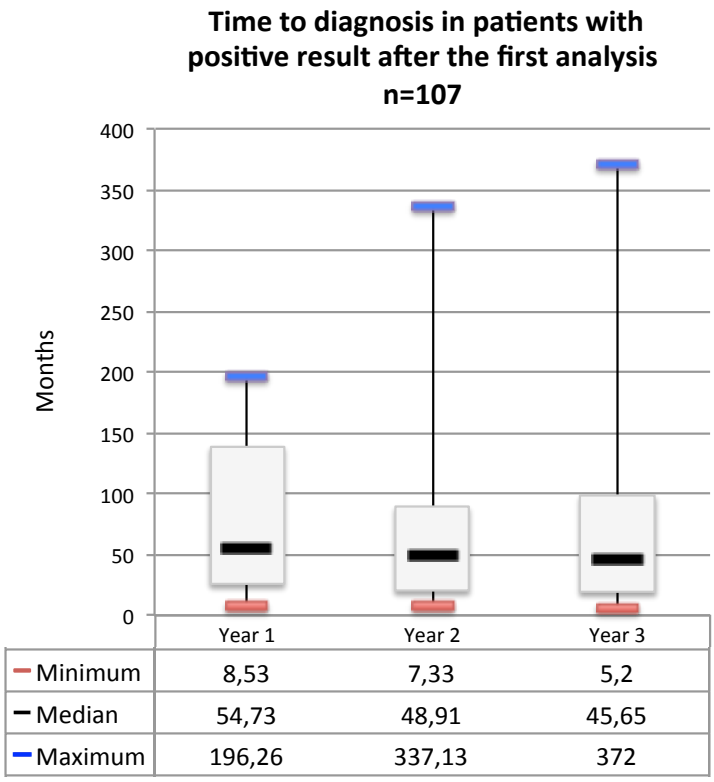
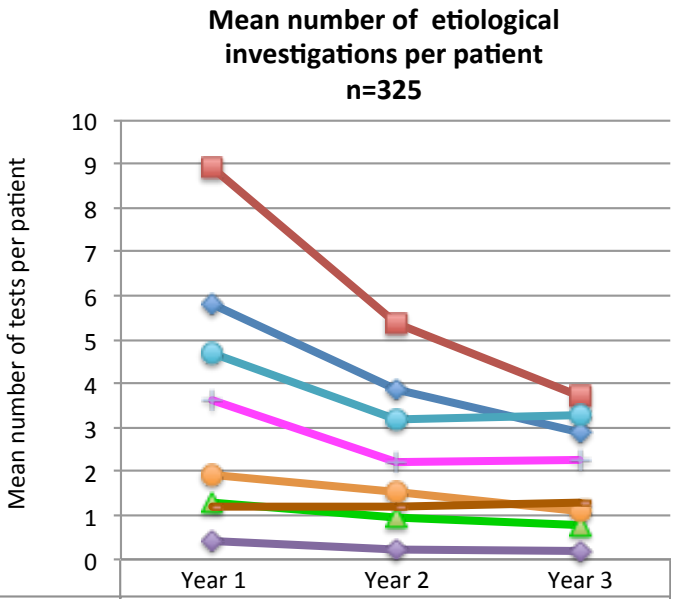


Figure 4 : Overall molecular diagnostic yield of phenotypic groups



	Year 1	Year 2	Year 3
Minimum	8,53	7,33	5,2
Median	54,73	48,91	45,65
Maximum	196,26	337,13	372



	Year 1	Year 2	Year 3
Standard biology	5,8	3,8	2,9
Metabolic testing	8,9	5,4	3,7
Neurophysiology	1,3	0,9	0,8
Histology	0,4	0,2	0,2
Imaging	4,7	3,2	3,3
Cytogenetics	1,9	1,5	1,1
Molecular genetics	3,6	2,2	2,3
Sensorial	1,2	1,2	1,3
Total	27,8	18,5	15,5

Figure 5 : Overview of the time to diagnosis in months and etiological investigations in the local patients

UNIVERSITE DE BOURGOGNE

THESE SOUTENUE PAR Me NAMBOT Sophie

CONCLUSIONS

En conclusion, cette étude montre que le séquençage d'exome est un examen diagnostique puissant dans le domaine des anomalies du développement et/ou de la déficience intellectuelle. Il permet l'identification de nouveaux gènes, l'extension du spectre phénotypique de certaines maladies humaines connues, la limitation des investigations étiologiques préalables et la réduction du délai diagnostique. Le diagnostic moléculaire permet d'établir un conseil génétique précis, d'affiner le pronostic, de prévenir certaines complications et de proposer des pistes thérapeutiques, éléments clés d'une prise en charge personnalisée.

Compte tenu de l'augmentation exponentielle des connaissances génétiques, son rendement diagnostique continuera de croître dans les années à venir, soutenu par la ré-analyse régulière et le partage international de données. Mais, même si la couverture des régions codantes ne cesse de s'améliorer, l'exploration des régions non-codantes du génome deviendra inévitable pour poursuivre cet essor. Le séquençage du génome est donc le prochain défi génétique à relever, induisant à nouveau des changements inhérents à la technique auxquels il faudra s'adapter. Mais, si sa contribution diagnostique est aussi importante que celle du séquençage d'exome, cela en vaut indéniablement la peine pour les patients et leurs familles.

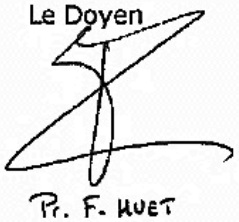
Le Président du jury,

Pr Laurence OLIVIER-FAIVRE

Vu et permis d'imprimer

Dijon, le mardi 13 Décembre 2016

Le Doyen



Pr. F. HUET

REFERENCES

1. Maulik PK, Mascarenhas MN, Mathers CD, Dua T, Saxena S. Prevalence of intellectual disability: a meta-analysis of population-based studies. Res Dev Disabil. 2011;32(2):419–436.

2. McCandless SE, Brunger JW, Cassidy SB. The burden of genetic disease on inpatient care in a children's hospital. Am J Hum Genet 2004: 74: 121–127.

3. Schieppati A, Henter J-I, Daina E, Aperia A. Why rare diseases are an important medical and social issue. Lancet 2008: 371: 2039–2041

4. Shashi V, McConkie-Rosell A, Rosell B, Schoch K, Vellore K, McDonald M, Jiang YH, Xie P, Need A, Goldstein DB. The utility of the traditional medical genetics diagnostic evaluation in the context of next-generation sequencing for undiagnosed genetic disorders. Genet Med. 2014 Feb;16(2):176-82. doi: 10.1038/gim.2013.99. Epub 2013 Aug 29.

5. Ng SB, Buckingham KJ, Lee C, Bigham AW, Tabor HK, Dent KM, Huff CD, Shannon PT, Jabs EW, Nickerson DA, Shendure J, Bamshad MJ. Exome sequencing identifies the cause of a mendelian disorder . Nat Genet 2010;42:30-5.

6. Thauvin-Robinet C, Auclair M, Duplomb L, Caron-Debarle M, Avila M, St-Onge J, Le Merrer M, Le Luyer B, Héron D, Mathieu-Dramard M, Bitoun P, Petit JM, Odent S, Amiel J, Picot D, Carmignac V, Thevenon J, Callier P, Laville M, Reznik Y, Fagour C, Nunes ML, Capeau J, Lascols O, Huet F, Faivre L, Vigouroux C, Rivière JB. PIK3R1 mutations cause syndromic insulin resistance with lipoatrophy. Am J Hum Genet 2013;93:141-9.

7. Thauvin-Robinet C, Lee JS, Lopez E, Herranz-Pérez V, Shida T, Franco B, Jego L, Ye F, Pasquier L, Loget P, Gigot N, Aral B, Lopes CA, St-Onge J, Bruel AL, Thevenon J, González-Granero S, Alby C, Munnich A, Vekemans M, Huet F, Fry AM, Saunier S, Rivière JB, Attié-Bitach T, Garcia-Verdugo JM, Faivre L, Mégarbané A, Nachury MV. The oral-

facial-digital syndrome gene C2CD3 encodes a positive regulator of centriole elongation. Nat Genet 2014;46:905-11.

8. Vissers LE, Gilissen C, Veltman JA. Genetic studies in intellectual disability and related disorders. Nat Rev Genet. 2016 Jan;17(1):9-18. doi: 10.1038/nrg3999. Review.

9. Gilissen C, Hehir-Kwa JY, Thung DT, van de Vorst M, van Bon BW, Willemsen MH, Kwint M, Janssen IM, Hoischen A, Schenck A, Leach R, Klein R, Tearle R, Bo T, Pfundt R, Yntema HG, de Vries BB, Kleefstra T, Brunner HG, Vissers LE, Veltman JA. Genome sequencing identifies major causes of severe intellectual disability. Nature. 2014 Jul 17;511(7509):344-7. doi: 10.1038/nature13394.

10. Deciphering Developmental Disorders Study. Large-scale discovery of novel genetic causes of developmental disorders. Nature. 2015 Mar 12;519(7542):223-8. doi: 10.1038/nature14135.

11. Iossifov I, O'Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, Stessman HA, Witherspoon KT, Vives L, Patterson KE, Smith JD, Paeper B, Nickerson DA, Dea J, Dong S, Gonzalez LE, Mandell JD, Mane SM, Murtha MT, Sullivan CA, Walker MF, Waqar Z, Wei L, Willsey AJ, Yamrom B, Lee YH, Grabowska E, Dalkic E, Wang Z, Marks S, Andrews P, Leotta A, Kendall J, Hakker I, Rosenbaum J, Ma B, Rodgers L, Troge J, Narzisi G, Yoon S, Schatz MC, Ye K, McCombie WR, Shendure J, Eichler EE, State MW, Wigler M. The contribution of de novo coding mutations to autism spectrum disorder. Nature. 2014 Nov 13;515(7526):216-21. doi: 10.1038/nature13908.

12. Stessman HA, Willemsen MH, Fenckova M, Penn O, Hoischen A, Xiong B, Wang T, Hoekzema K, Vives L, Vogel I, Brunner HG, van der Burgt I, Ockeloen CW, Schuurs-Hoeijmakers JH, Klein Wassink-Ruiter JS, Stumpel C, Stevens SJ, Vles HS, Marcelis CM, van Bokhoven H, Cantagrel V, Colleaux L, Nicouleau M, Lyonnet S, Bernier RA, Gerdts J, Coe BP, Romano C, Alberti A, Grillo L, Scuderi C, Nordenskjöld M, Kvarnung M, Guo H,

Xia K, Piton A, Gerard B, Genevieve D, Delobel B, Lehalle D, Perrin L, Prieur F, Thevenon J, Gecz J, Shaw M, Pfundt R, Keren B, Jacquette A, Schenck A, Eichler EE, Kleefstra T. Disruption of POGZ Is Associated with Intellectual Disability and Autism Spectrum Disorders. Am J Hum Genet. 2016 Mar 3;98(3):541-52. doi: 10.1016/j.ajhg.2016.02.004.

13. Helbig KL, Farwell Hagman KD, Shinde DN, Mroske C, Powis Z, Li S, Tang S, Helbig I. Diagnostic exome sequencing provides a molecular diagnosis for a significant proportion of patients with epilepsy. Genet Med. 2016 Sep;18(9):898-905. doi: 10.1038/gim.2015.186.

14. Kodera H, Ohba C, Kato M, Maeda T, Araki K, Tajima D, Matsuo M, Hino-Fukuyo N, Kohashi K, Ishiyama A, Takeshita S, Motoi H, Kitamura T, Kikuchi A, Tsurusaki Y, Nakashima M, Miyake N, Sasaki M, Kure S, Haginoya K, Saitsu H, Matsumoto N. De novo GABRA1 mutations in Ohtahara and West syndromes. Epilepsia. 2016 Apr;57(4):566-73. doi: 10.1111/epi.13344.

15. Long PA, Larsen BT, Evans JM, Olson TM. Exome Sequencing Identifies Pathogenic and Modifier Mutations in a Child With Sporadic Dilated Cardiomyopathy. J Am Heart Assoc. 2015 Dec 9;4(12). pii: e002443. doi: 10.1161/JAHA.115.002443.

16. Blue GM, Kirk EP, Giannoulatou E, Dunwoodie SL, Ho JW, Hilton DC, White SM, Sholler GF, Harvey RP, Winlaw DS. Targeted next-generation sequencing identifies pathogenic variants in familial congenital heart disease. J Am Coll Cardiol. 2014 Dec 16;64(23):2498-506. doi: 10.1016/j.jacc.2014.09.048.

17. Chong JX, Buckingham KJ, Jhangiani SN, Boehm C, Sobreira N, Smith JD, Harrell TM, McMillin MJ, Wiszniewski W, Gambin T, Coban Akdemir ZH, Doheny K, Scott AF, Avramopoulos D, Chakravarti A, Hoover-Fong J, Mathews D, Witmer PD, Ling H, Hetrick K, Watkins L, Patterson KE, Reinier F, Blue E, Muzny D, Kircher M, Bilguvar K, López-Giráldez F, Sutton VR, Tabor HK, Leal SM, Gunel M, Mane S, Gibbs RA, Boerwinkle E, Hamosh A, Shendure J, Lupski JR, Lifton RP, Valle D, Nickerson DA; Centers for Mendelian

Genomics., Bamshad MJ. The Genetic Basis of Mendelian Phenotypes: Discoveries, Challenges, and Opportunities. Am J Hum Genet. 2015 Aug 6;97(2):199-215. doi: 10.1016/j.ajhg.2015.06.009. Review.

18. Lee H, Deignan JL, Dorrani N, Strom SP, Kantarci S, Quintero-Rivera F, Das K, Toy T, Harry B, Yourshaw M, Fox M, Fogel BL, Martinez-Agosto JA, Wong DA, Chang VY, Shieh PB, Palmer CGS, Dipple KM, Grody WW, Vilain E, Nelson SF. Clinical Exome Sequencing for Genetic Identification of Rare Mendelian Disorders. JAMA. 2014 Nov 12;312(18):1880-7. doi: 10.1001/jama.2014.14604.

19. Thevenon J, Duffourd Y, Masurel-Paulet A, Lefebvre M, Feillet F, El Chehadeh-Djebbar S, St-Onge J, Steinmetz A, Huet F, Chouchane M, Darmency-Stamboul V, Callier P, Thauvin-Robinet C, Faivre L, Rivière JB. Diagnostic odyssey in severe neurodevelopmental disorders: toward clinical whole-exome sequencing as a first-line diagnostic test. Clin Genet. 2016 Jun;89(6):700-7. doi: 10.1111/cge.12732. Epub 2016 Apr 26.

20. Yang Y, Muzny DM, Xia F, Niu Z, Person R, Ding Y, Ward P, Braxton A, Wang M, Buhay C, Veeraraghavan N, Hawes A, Chiang T, Leduc M, Beuten J, Zhang J, He W, Scull J, Willis A, Landsverk M, Craigen WJ, Bekheirnia MR, Stray-Pedersen A, Liu P, Wen S, Alcaraz W, Cui H, Walkiewicz M, Reid J, Bainbridge M, Patel A, Boerwinkle E, Beaudet AL, Lupski JR, Plon SE, Gibbs RA, Eng CM. Molecular Findings Among Patients Referred for Clinical Whole-Exome Sequencing. JAMA. 2014 Nov 12;312(18):1870-9. doi: 10.1001/jama.2014.14601.

21. Willemsen MH, Kleefstra T. Making headway with genetic diagnostics of intellectual disabilities. Clin Genet. 2014 Feb;85(2):101-10. doi: 10.1111/cge.12244. Epub 2013 Aug 26. Review.

22. Iglesias A, Anyane-Yeboa K, Wynn J, Wilson A, Truitt Cho M, Guzman E, Sisson R, Egan C, Chung WK. The usefulness of whole-exome sequencing in routine clinical practice. Genet Med. 2014 Dec;16(12):922-31. doi: 10.1038/gim.2014.58. Epub 2014 Jun 5.

23. Philippakis AA, Azzariti DR, Beltran S, Brookes AJ, Brownstein CA, Brudno M, Brunner HG, Buske OJ, Carey K, Doll C, Dumitriu S, Dyke SO, den Dunnen JT, Firth HV, Gibbs RA, Girdea M, Gonzalez M, Haendel MA, Hamosh A, Holm IA, Huang L, Hurles ME, Hutton B, Krier JB, Misyura A, Mungall CJ, Paschall J, Paten B, Robinson PN, Schiettecatte F, Sobreira NL, Swaminathan GJ, Taschner PE, Terry SF, Washington NL, Züchner S, Boycott KM, Rehm HL. The Matchmaker Exchange: a platform for rare disease gene discovery. Hum Mutat. 2015 Oct;36(10):915-21. doi: 10.1002/humu.22858.

24. Lelieveld SH, Reijnders MR, Pfundt R, Yntema HG, Kamsteeg EJ, de Vries P, de Vries BB, Willemsen MH, Kleefstra T, Löhner K, Vreeburg M, Stevens SJ, van der Burgt I, Bongers EM, Stegmann AP, Rump P, Rinne T, Nelen MR, Veltman JA, Vissers LE, Brunner HG, Gilissen C. Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. Nat Neurosci. 2016 Sep;19(9):1194-6. doi: 10.1038/nn.4352.

25. Akawi N, McRae J, Ansari M, Balasubramanian M, Blyth M, Brady AF, Clayton S, Cole T, Deshpande C, Fitzgerald TW, Foulds N, Francis R, Gabriel G, Gerety SS, Goodship J, Hobson E, Jones WD, Joss S, King D, Klena N, Kumar A, Lees M, Lelliott C, Lord J, McMullan D, O'Regan M, Osio D, Piombo V, Prigmore E, Rajan D, Rosser E, Sifrim A, Smith A, Swaminathan GJ, Turnpenny P, Whitworth J, Wright CF, Firth HV, Barrett JC, Lo CW, FitzPatrick DR, Hurles ME; DDD study. Discovery of four recessive developmental disorders using probabilistic genotype and phenotype matching among 4,125 families. Nat Genet. 2015 Nov;47(11):1363-9. doi: 10.1038/ng.3410.

26. Wenger AM, Guturu H, Bernstein JA, Bejerano G. Systematic reanalysis of clinical exome data yields additional diagnoses: implications for providers. Genet Med. 2016 Jul 21. doi: 10.1038/gim.2016.88. [Epub ahead of print]

27. Monroe GR, Frederix GW, Savelberg SM, de Vries TI, Duran KJ, van der Smagt JJ, Terhal PA, van Hasselt PM, Kroes HY, Verhoeven-Duif NM, Nijman IJ, Carbo EC, van Gassen KL, Knoers NV, Hövels AM, van Haelst MM, Visser G, van Haaften G. Effectiveness of whole-exome sequencing and costs of the traditional diagnostic trajetory in children with intellectual disability. Genet Med. 2016 Sep;18(9):949-56. doi: 10.1038/gim.2015.200.

28. Heger M. Dutch Study Aims to Demonstrate Cost-Effectiveness of Reimbursing for Exome Sequencing Dx. 2011. URL: https://www.genomeweb.com/sequencing/dutch-study-aims-demonstrate-cost-effectiveness-reimbursing-exome-sequencing-dx

29. Towne C, Beggs A, Agrawal P. Efficiency of Whole Exome/Genome Sequencing for Achieving a Diagnosis in Rare Presentations. 2013. URL: http://www.ashg.org/2013meeting/abstracts/fulltext/ f130122905.htm

30. Valencia CA, Husami A, Holle J, Johnson JA, Qian Y, Mathur A, Wei C, Indugula SR, Zou F, Meng H, Wang L, Li X, Fisher R, Tan T, Hogart Begtrup A, Collins K, Wusik KA, Neilson D, Burrow T, Schorry E, Hopkin R, Keddache M, Harley JB, Kaufman KM, Zhang K. Clinical Impact and Cost-Effectiveness of Whole Exome Sequencing as a Diagnostic Tool: A Pediatric Center's Experience. Front Pediatr 2015;3:67.

31. Wright CF, Fitzgerald TW, Jones WD, Clayton S, McRae JF, van Kogelenberg M, King DA, Ambridge K, Barrett DM, Bayzetinova T, Bevan AP, Bragin E, Chatzimichali EA, Gribble S, Jones P, Krishnappa N, Mason LE, Miller R, Morley KI, Parthiban V, Prigmore E, Rajan D, Sifrim A, Swaminathan GJ, Tivey AR, Middleton A, Parker M, Carter NP, Barrett JC, Hurles ME, FitzPatrick DR, Firth HV; DDD study. Genetic diagnosis of developmental

disorders in the DDD study: a scalable analysis of genome-wide research data. Lancet. 2015 Apr 4;385(9975):1305-14. doi: 10.1016/S0140-6736(14)61705-0. Epub 2014 Dec 17.

32. Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL, McGuire AL, Nussbaum RL, O'Daniel JM, Ormond KE, Rehm HL, Watson MS, Williams MS, Biesecker LG; American College of Medical Genetics and Genomics. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. Genet Med. 2013 Jul;15(7):565-74. doi: 10.1038/gim.2013.73. Epub 2013 Jun 20.

33. Rehm HL, Bale SJ, Bayrak-Toydemir P, Berg JS, Brown KK, Deignan JL, Friez MJ, Funke BH, Hegde MR, Lyon E, Working Group of the American College of Medical Genetics and Genomics Laboratory Quality Assurance Commitee: ACMG clinical laboratory standards for next-generation sequencing. Genet Med. 2013 Sep;15(9):733-47. doi: 10.1038/gim.2013.92.

34. Richards CS, Bale S, Bellissimo DB, Das S, Grody WW, Hegde MR, Lyon E, Ward BE; Molecular Subcommittee of the ACMG Laboratory Quality Assurance Committee. ACMG recommendations for standards for interpretation and reporting of sequence variations: Revisions 2007. Genet Med. 2008 Apr;10(4):294-300. doi: 10.1097/GIM.0b013e31816b5cae.

35. 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 May;17(5):405-24. doi: 10.1038/gim.2015.30.

36. Allou L, Julia S, Amsallem D, El Chehadeh S, Lambert L, Thevenon J, Duffourd Y, Saunier A, Bouquet P, Pere S, Moustaïne A, Ruaud L, Roth V, Jonveaux P, Philippe C. Rett-like phenotypes: expanding the genetic heterogeneity to the KCNA2 gene and first familial

case of CDKL5-related disease. *Clin Genet.* 2016 Apr 7. doi:10.1111/cge.12784. [Epub ahead of print]

37. Thauvin-Robinet C, Duplomb-Jego L, Limoge F, Picot D, Masurel A, Terriat B, Champilou C, Minot D, St-Onge J, Kuentz P, Duffourd Y, Thevenon J, Rivière JB, Faivre L. Homozygous FIBP nonsense variant responsible of syndromic overgrowth, with overgrowth, macrocephaly, retinal coloboma and learning disabilities. Clin Genet. 2016 May;89(5):e1-4. doi: 10.1111/cge.12704.

38. Assoum M, Philippe C, Isidor B, Perrin L, Makrythanasis P, Sondheimer N, Paris C, Douglas J, Lesca G, Antonarakis S, Hamamy H, Jouan T, Duffourd Y, Auvin S, Saunier A, Begtrup A, Nowak C, Chatron N, Ville D, Mireskandari K, Milani P, Jonveaux P, Lemeur G, Milh M, Amamoto M, Kato M, Nakashima M, Miyake N, Matsumoto N, Masri A, Thauvin-Robinet C, Rivière JB, Faivre L, Thevenon J. Autosomal-Recessive Mutations in AP3B2, Adaptor-Related Protein Complex 3 Beta 2 Subunit, Cause an Early-Onset Epileptic Encephalopathy with Optic Atrophy. Am J Hum Genet. 2016 Dec 1;99(6):1368-1376. doi: 10.1016/j.ajhg.2016.10.009.

39. Thevenon J, Milh M, Feillet F, St-Onge J, Duffourd Y, Jugé C, Roubertie A, Héron D, Mignot C, Raffo E, Isidor B, Wahlen S, Sanlaville D, Villeneuve N, Darmency-Stamboul V, Toutain A, Lefebvre M, Chouchane M, Huet F, Lafon A, de Saint Martin A, Lesca G, El Chehadeh S, Thauvin-Robinet C, Masurel-Paulet A, Odent S, Villard L, Philippe C, Faivre L, Rivière JB. Mutations in SLC13A5 cause autosomal-recessive epileptic encephalopathy with seizure onset in the first days of life. Am J Hum Genet. 2014 Jul 3;95(1):113-20. doi: 10.1016/j.ajhg.2014.06.006.

40. Bruel AL, Masurel-Paulet A, Rivière JB, Duffourd Y, Lehalle D, Bensignor C, Huet F, Borgnon J, Roucher F, Kuentz P, Deleuze JF, Thauvin-Robinet C, Faivre L, Thevenon J.

Autosomal recessive truncating MAB21L1 mutation associated with a syndromic scrotal agenesis. Clin Genet. 2016 Apr 22. doi: 10.1111/cge.12794. [Epub ahead of print]

41. El Chehadeh S, Kerstjens-Frederikse WS, Thevenon J, Kuentz P, Bruel AL, Thauvin-Robinet C, Bensignor C, Dollfus H, Laugel V, Rivière JB, Duffourd Y, Bonnet C, Robert MP, Isaiko R, Straub M, Creuzot-Garcher C, Calvas P, Chassaing N, Loeys B, Reyniers E, Vandeweyer G, Kooy F, Hančárová M, Havlovicová M, Prchalová D, Sedláček Z, Gilissen C, Pfundt R, Wassink-Ruiter JS, Faivre L. Dominant variants in the splicing factor PUF60 cause a recognizable syndrome with intellectual disability, heart defects and short stature. Eur J Hum Genet. 2016 Jan;25(1):43-51. doi: 10.1038/ejhg.2016.133.

42. S. Nambot, D. Gavrilov, J. Thevenon, AL. Bruel, M. Bainbridge, M. Rio, C. Goizet, A. Rötig, J. Jaeken, N. Niu, F. Xia, A. Vital, N.Houcinat, F. Mochel, P. Kuentz, D. Lehalle, Y. Duffourd, JB. Rivière, C. Thauvin-Robinet, AL. Beaudet, L. Faivre. Further delineation of a rare recessive encephalomyopathy linked to mutations in GFER thanks to Data Sharing of whole exome sequencing data. Clin Genet. 2016 in submission

43. Amos JS, Huang L, Thevenon J, Kariminedjad A, Beaulieu CL, Masurel-Paulet A, Najmabadi H, Fattahi Z, Beheshtian M, Tonekaboni SH, Tang S, Helbig KL, Alcaraz W, Rivière JB, Faivre L, Innes AM, Lebel RR, Boycott KM. Autosomal recessive mutations in THOC6 cause intellectual disability: Syndrome delineation requiring forward and reverse phenotyping. Clin Genet. 2016 Apr 22. doi: 10.1111/cge.12793. [Epub ahead of print].

44. Friedman J, Roze E, Abdenur JE, Chang R, Gasperini S, Saletti V, Wali GM, Eiroa H, Neville B, Felice A, Parascandalo R, Zafeiriou DI, Arrabal-Fernandez L, Dill P, Eichler FS, Echenne B, Gutierrez-Solana LG, Hoffmann GF, Hyland K, Kusmierska K, Tijssen MA, Lutz T, Mazzuca M, Penzien J, Poll-The BT, Sykut-Cegielska J, Szymanska K, Thöny B, Blau N. Sepiapterin reductase deficiency: a treatable mimic of cerebral palsy.Ann Neurol. 2012 Apr;71(4):520-30. doi: 10.1002/ana.22685.

SUPPLEMENTAL METHODS

Detailed methods of whole-exome sequencing

Sequencing

Peripheral-blood samples of index cases were provided in most cases but purified genomic DNA from other tissue was accepted (skin, thymus, saliva...). DNA was extracted using the QIAamp DNA Blood Mini Kit (Qiagen) following standard procedures. Libraries were prepared from 1 µg of genomic DNA per individual using the Agilent Sureselect Human All Exon v5 kit, and were sequenced on a HiSeq2000 (Illumina) according to the manufacturer's recommendation for paired-end 76 bp reads. An average of 4.8 gigabases of mappable sequences per individual were generated, resulting in a depth of coverage of at least ten reads for more than 95% of RefSeq coding exons. Reads were aligned to the human genome reference sequence GRCh37/hg19 build of UCSC Genome Browser with the Burrows-Wheeler Aligner (BWA, v.0.6.2), and potential duplicate paired-end reads were marked with Picard v.2.01.The Genome Analysis Toolkit (GATK) v.3.4-46 was used for base quality score recalibration, indel realignment, and variant discovery (both single-nucleotide variants (SNVs) and indels).

SNV analysis and interpretation

Detected variants were annotated with SeattleSeq SNP Annotation. The priorization of the variants first selected variants responding to the folowing criteria: quality scores of the supporting bases superior to 30 (Phred score), sequencing depth superior to 5 reads, and at least 10% of the reads supporting the alternative allele. A mean of 300.000 variants were identified per individual. Then, priorization focused on variants affecting the reading frame (non-synonymous substitutions, splice variants (10bp), small indels) present at a frequency

less than 1% in dbSNP 138, in the NHLBI GO Exome Sequencing Project and in the ExAC database and in a local database of WES.

In the first step of diagnostic analysis, the remained variants (mean 500 per individual) were then priorized if reported as pathogenic or probably pathogenic in public variant databases of clinical interest (Clinvar, see URLS) or as affecting a gene associated with a human disorder referenced in OMIM (See URLs) or manually curated via Pubmed **(**See URLs). Each of the remaining variant (mean 80 per individual) were interpreted according to the ACMG guidelines [34,35] and classified in 5 categories: pathogenic, probably pathogenic, variant of unknown significance (VUS), probably benign or benign. The classification was based on i) the type of variant, its frequency and status in the public databases used for priorization; ii) the global patient's clinical presentation and its concordance with the phenotype described in the literature; iii) the familial history, the predict mode of transmission, and the familial segregation; iv) the prevalence of the pathology compared to the frequency of the variant in general population; v) the predicted effect of the variant by *in-silico* tools including scoreCADD, polyPhen, consScoreGERP, and GranthamScore; vi) the functional data available in the literature. All putative causative variants were subjected to non-exhaustive literature and databases searches (See URLs). Results were multidisciplinary discussed by laboratory directors and physicians with clinical expertise in regular meetings to reverse phenotyping. Variants considered as pathogenic and probably pathogenic, further classified as related to the patient's phenotype, were concluded as positive diagnosis. The 56 genes of the list of medically actionable incidental findings defined by the ACMG were also studied and interpreted according to the ACMG recommandations. The results were returned to the patient when consent was given [32].

In second step, with translational research approach for patients with non positive diagnosis approach, interpretation focused on all other variants in order to identify strong

candidate genes. These non human-disease genes were considered as strong candidates because of variant type, consistent family segregation (mainly sporadic or biallelic variants), encoded protein function, published functional studies and available animals models. But they could not be initially considered as causative because of lack of phenotypic or molecular evidences. For these cases, intensive international datasharing through databases as Matchmaker exchange or PhenomeCentral was used to identify additional patients carrying variants in the same gene, in order to increase the recurrence and definitively conclude on their implication in patient's phenotype.

CNV detection and interpretation

In-house pipeline was developed for CNV detection in November 2015. The procedure for CNV detection is reported elsewhere [19]. The whole in-house database of WES data with the same capture kit (n = 846) was used as a normalization dataset. Among the 193 342 captured exons, 2 with extreme GC (over 90% or below 10%) and 2691 with low complexity were excluded from analysis. Moreover, 8602 exons considered outliers were also excluded from analysis (targets with outlier read depth values or targets with a outlier post-normalized standard deviation). Overall, 12 228 CNVs were identified, and were then filtered (n=3410) on three attributes : XHMM quality score (SQ) > 60, exon spanned > 2, and minor allele frequency (MAF) < 1%. All the detected CNV were plotted and visualized before consideration for interpretation using the matplotlib python library (See URLs). The ressources used for the annotation included Database of Genomic Variants (structural variants, see URLs), Clinical Genome Resource dosage sensitive genes annotation (download at NCBI ftp://ftp.ncbi.nlm.nih.gov/pub/dbVar/clingen/), Clinical Genome Resource CNVs found in patients referred for genetic testing for indications such as intellectual disability, developmental delay, autism and congenital anomalies (download at UCSC: ClinGen CNVs Track), and Copy Number Variation Morbidity Map of Developmental Delay CNVs from

cases of developmental delay along with healthy control sets (download at UCSC: Development Delay Track).

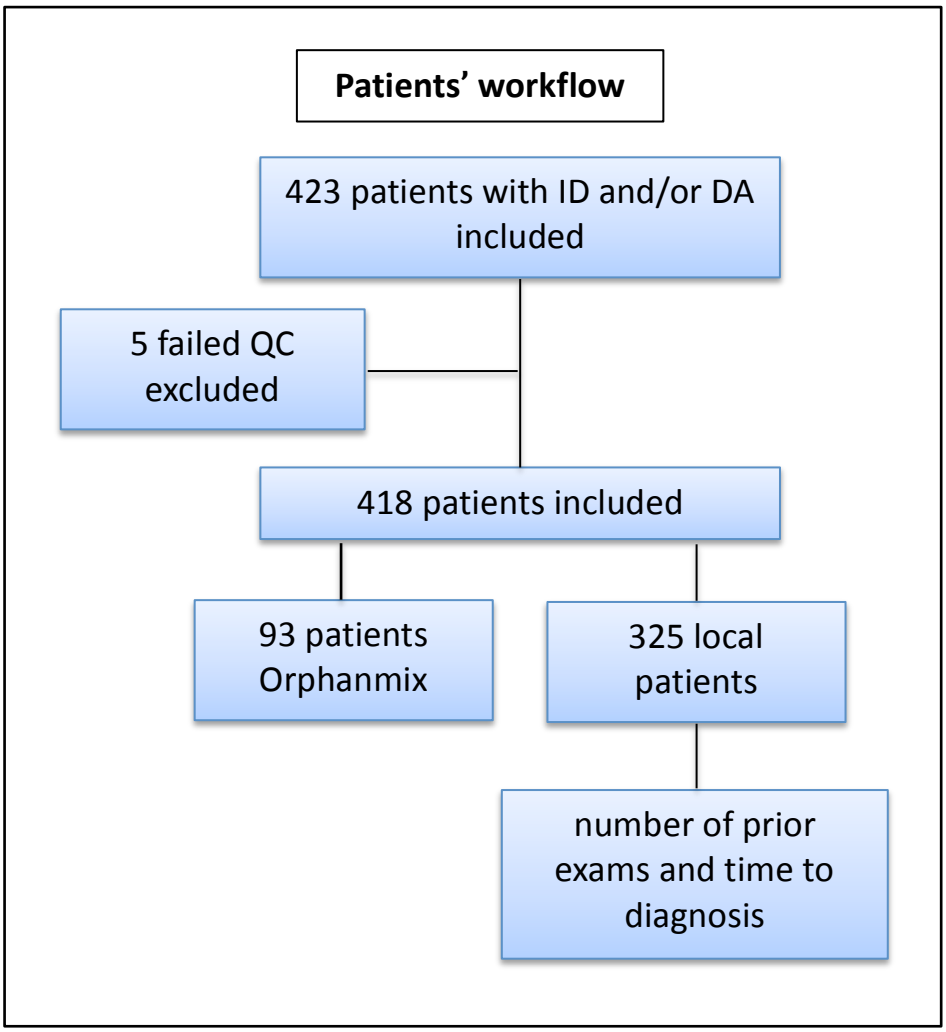
Candidate CNVs were classified considering i) specific CNV reported in public databases to be associated with a human disorder; ii) CNV encompassing/interrumpting gene(s) associated with a human disorder referenced in OMIM (See URLs) or manually curated via Pubmed (See URLs); iii) CNV type consistent with both the chromosomic rearrangment reported in the considered disorder and the suspected mode of inheritance (i.e. X-linked, autosomal recessive or autosomal dominant); iv) the frequency of the CNV in the databases used for the annotation cited above.

SNV and CNV confirmation

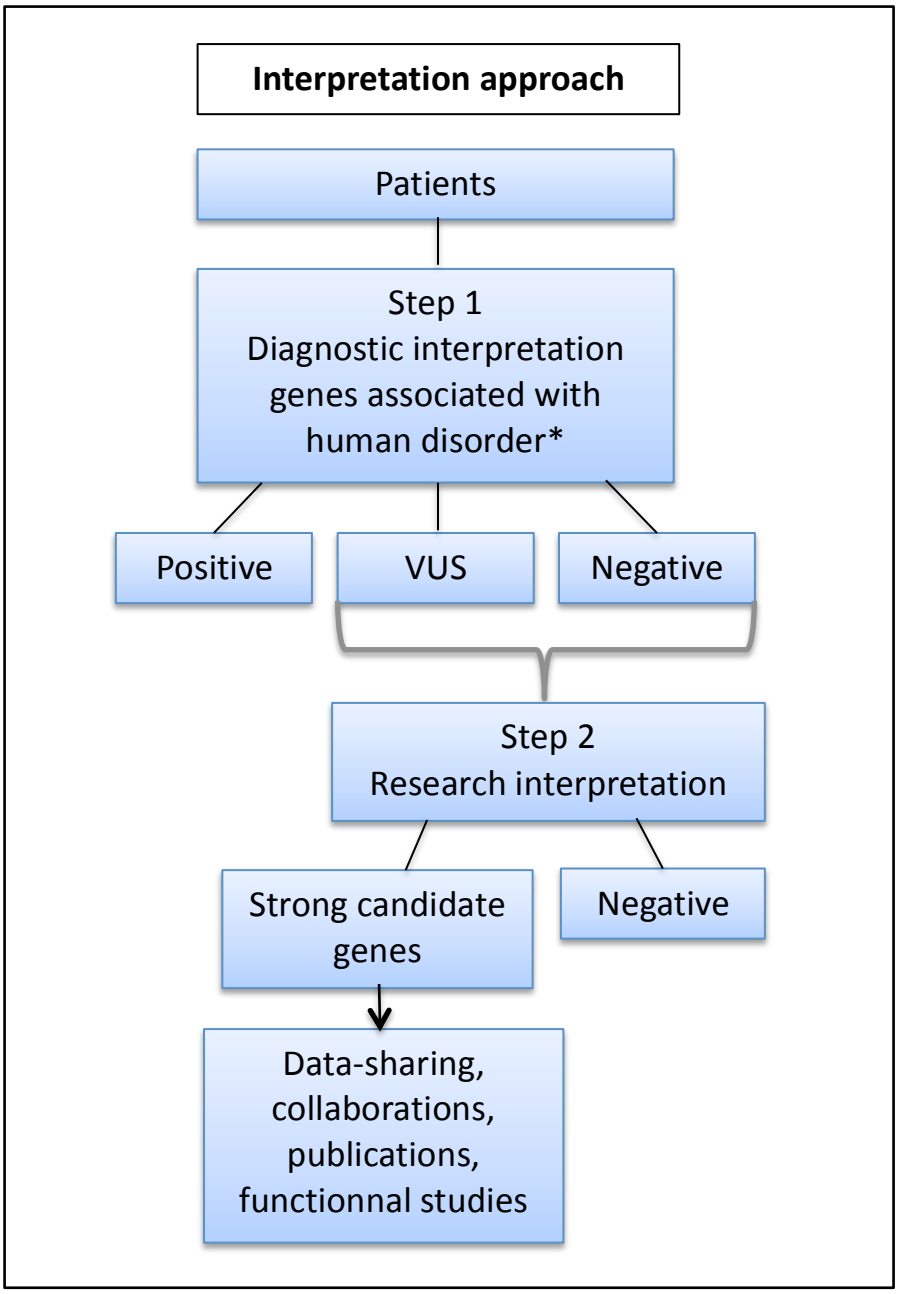
All sequence rare and deemed clinically relevant variants were confirmed by Sanger sequence analysis using a separate DNA preparation, along with segregation analysis of the variants in the family members if confirmed. DNA was extracted from blood samples of the family members, using the QIAamp DNA Blood Mini Kit (Qiagen) following standard procedures. Genomic DNA was amplified by polymerase chain reaction (PCR) using the HotStarTaq PCR kit (Qiagen) according to the manufacturer's protocol. PCR products were purified using the Agencourt CleanSEQ system (Beckman Coulter) and were sequenced with the BigDye Terminator Cycle Sequencing kit, v3.1 (Applied Biosystems) in ABI 3730 sequencer (Applied Biosystems). Sequence data were analysed using Mutation Surveyor v4.0.9.

CNVs were confirmed either using existing array-CGH data, or by quantitative PCR (qPCR) or direct Sanger sequencing. qPCR was performed on an CFX96 (Biorad) instrument with each sample as duplicate. Results were analyzed with normalization of all assays on two control genes and further compared to DNA from a control subject without any known CNV

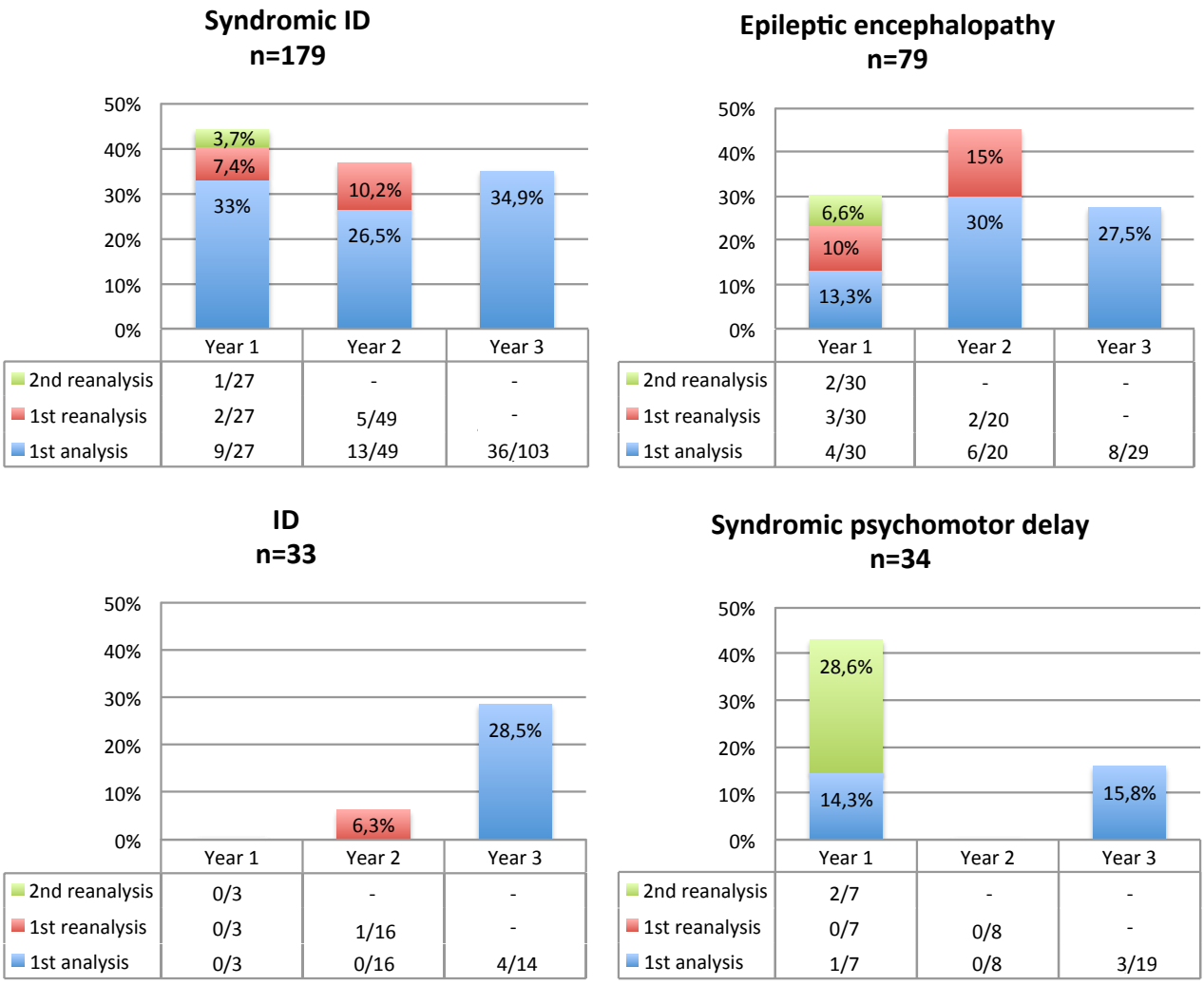
at this position. Biological concordance between samples was attested by the VeriFiler Direct PCR Amplification Kit (Life Technologies) for each suspected *de novo* variant.



Supplemental Figure 1a: Representation of Patient's workflow



Supplemental Figure 1b: Representation of Interpretation approach
***reported as: pathogenic/probably pathogenic in public variant databases of clinical interest (Clinvar) or affecting a gene associated with a human disorder referenced in OMIM or manually curated via Pubmed



Supplemental Figure 2 : Detailed molecular diagnostic yield of phenotypic neurodevelopmental disorders group

Supplemental Table 1: Diagnosis obtained by WES reanalysis

Cause of change	Diagnosis from WES data re-analysis	Date of first WES analysis	Publication date of article of interest
Collaboration with another team			
THOC6	Beaulieu-Boycott-Innes syndrome (MIM #613680)	November 2014	April 2016
KLHL7		December 2014	In submission
KCNA2	Early infantile epileptic encephalopathy-32 (MIM #616366)	March 2015	April 2016
Team's publication			
SLC13A5(X2)	Early-onset epileptic encephalopathy-25 (MIM #615905)	November 2013	July 2014
GFER	Myopathy with cataract and combined respiratory chain deficiency (MIM #613076)	June 2014	In submission
MAB21L1	Syndromic scrotal agenesis	November 2014	April 2016
AP3B2	Early-Onset Epileptic Encephalopathy with Optic Atrophy.	March 2014	December 2016
FIBP	Thauvin-Robinet-Faivre syndrome (MIM #617107)	June 2014	May 2016
PUF60	Verheij syndrome (MIM #615583)	May 2014	January 2016
Recent publication			
EEF1A2	Early infantile epileptic encephalopathy-33 (MIM #616409)	January 2015	July 2013 and April 2015 (OMIM report 's creation in April 2015
KDM5C	Claes-Jensen type syndromic X-linked mental retardation (MIM #300534)	January 2015	January 2012 and February 2015
NAA10	Syndromic X-linked ID	February 2014	May 2015
TBL1XR1	Autosomal dominant mental retardation-41 (MIM #616944)	January 2015	January 2015
PPP2R5D	Autosomal dominant mental retardation-35 (MIM #616355)	March 2015	July 2015
DDX3X (X2)	X-linked mental retardation-10 (MIM #300958)	August and November 2015	August 2015
Missed variant			
TCF4	Pitt-Hopkins syndrome (MIM #610954)		
GABRA1	Early infantile epileptic encephalopathy-19 (MIM		
KMT2A	Wiedemann-Steiner syndrome (MIM #605130)		
TUBB3	Congenital fibrosis of extraocular muscles-3A (MIM #600638)		
CNV			
CLN3	Neuronal ceroid lipofuscinosis-3 (MIM #204200)		

Supplemental Table 2: List of positive results

SNVs									
Autosomal dominant de novo									
Year	Patient's ID	Gene	Pathology	Genomic annotation	cDNA	Protein	Mutation type	Exon	Classification
Year 2	PED0051.1	*SHANK3*	Phelan-McDermid syndrome (MIM #606232)	chr22:g.51169584_51169586delCCA	NM_001080420.1:c.5090_5092del	p.(His1697del)	In/del	22	Probably pathogenic
Year 2	PED1082.1	*ARID1B*	Coffin-Siris syndrome-1 (MIM #135900)	chr6:g.157222569delA	NM_020732.3:c.1836del	p.(Tyr613Ilefs*55)	Frameshift	4	Probably pathogenic
Year 1	PED1306.1	*ARID1B*	Coffin-Siris syndrome-1 (MIM #135900)	chr6:g.157405927_157405946del	NM_017519.2:c.2130_2149del	p.(His710Glnfs*33)	Frameshift	6	Probably pathogenic
Year 1	PED1315.1	*CTNNB1*	Autosomal dominant mental retardation-19 (MIM #615075)	chr3:g.41280759CA>C	NM_001098209.1:c.2273del	p.(His758Leufs*33)	Frameshift	15	Probably pathogenic
Year 1	PED1318.1	*TBR1*	PMID: 25232744	chr2:g.162280277_162280283dup	NM_006593.2:c.1588_1594dup	p.(Thr532Argfs*144)	Frameshift	6	Probably pathogenic
Year 1	PED1320.1	*SHANK3*	Phelan-McDermid syndrome (MIM #606232)	chr22:g.51160594C>T	NM_001080420.1:c.4381C>T	p.(Gln1461*)	Nonsense	21	Probably pathogenic
Year 1	PED1343.1	*DYRK1A*	Autosomal dominant mental retardation-7 (MIM #614104)	chr21:g.38884305C>A	NM_001396.3:c.1763C>A	p.(Thr588Asn)	Missense	11	Probably pathogenic
Year 1	PED1433.3	*KMT2A*	Wiedemann-Steiner syndrome (MIM #605130)	chr11:g.118362545C>T	NM_005933.2:c.4897C>T	p.(Arg1633*)	Nonsense	15	Probably pathogenic
Year 1	PED1597.2	*SLC2A1*	GLUT1 deficiency syndrome 1 (MIM #606777)	chr1:g.43408909A>C	NM_006516.2:c.102T>G	p.(Asn34Lys)	Missense	2	Probably pathogenic
Year 1	PED1691.3	*NR2F1*	Bosch-Boonstra-Schaaf optic atrophy syndrome (MIM #615722)	chr5:g.92921018C>T	NM_005654.4:c.289C>T	p.(His97Tyr)	Missense	1	Probably pathogenic
Year 1	PED1909.1	*TUBB3*	Complex cortical dysplasia with other brain malformations-1 (MIM #614039)	chr16:g.90001721G>A	NM_001197181.1:c.646G>A	p.(Glu216Lys)	Missense	4	Probably pathogenic
Year 2	PED2064.1	*SLC6A1*	Myoclonic-atonic epilepsy (MIM #616421)	chr3:g.11076216G>C	NM_003042.3:c.1528-1G>C	intron 14/15	Splice site	15	Probably pathogenic
Year 2	PED2066.1	*SHOC2*	Noonan-like syndrome with loose anagen hair (MIM #607721)	chr10:g.112724120A>G	NM_001269039.1:c.4A>G	p.(Ser2Gly)	Missense	1	Pathogenic
Year 2	PED2102.1	*CHD8*	Susceptibility to autism 18 (MIM #615032)	chr14:g.21862492_21862501delTGGAAGTACT	NM_001170629.1:c.5534_5543del	p.(Lys1845Metfs*43)	Frameshift	30	Probably pathogenic
Year 3	PED2119.1	*SMARCA4*	Coffin-Siris syndrome-4 (MIM #614609)	chr19:g.11144847C>T	NM_003072.3:c.3922C>T	p.(Arg1308Trp)	Missense	28	Probably pathogenic
Year 3	PED2139.1	*GATAD2B*	Autosomal dominant mental retardation-18 (MIM # 615074)	chr1:g.153791266C>T	NM_020699.2:c.597+1G>A	Intron 5	Splice site	4	Probably pathogenic

Year 2	PED2186.1	*SETD5*	Autosomal dominant mental retardation-23 (MIM #615761)	chr3:g.9483840A>G	NM_001080517.1:c.988A>G	p.(Lys330Glu)	Missense	10	Probably pathogenic
Year 2	PED2231.1	*ANKRD11*	KBG syndrome (MIM #148050)	chr16:g.89349866G>T	NM_001256183.1:c.3084C>A	p.(Tyr1028*)	Nonsense	9	Probably pathogenic
Year 2	PED2241.1	*ASXL1*	Bohring-Opitz syndrome (MIM #605039)	chr20:g.31024284delG	NM_015338.4:c.3769del	p.(Ala1257Leufs*23)	Frameshift	12	Probably pathogenic
Year 2	PED2257.1	*ADNP*	Helsmoortel-Van der Aa syndrome (MIM #615873)	chr20:g.49509094_49509095insT	NM_181442.1:c.2156_2157insA	p.(Tyr719*)	Frameshift	4	Pathogenic
Year 2	PED2267.1	*DYNC1H1*	Autosomal dominant axonal Charcot-Marie-Tooth disease type 2O (MIM #614228)	chr14:g.102446124T>G	NM_001376.4:c.587T>G	p.(Leu196Trp)	Missense	4	Probably pathogenic
Year 2	PED2375.1	*GNAO1*	Early infantile epileptic encephalopathy-17 (MIM #615473)	chr16:g.56226265G>A	NM_020988.2:c.118G>A	p.(Gly40Arg)	Missense	1	Probably pathogenic
Year 2	PED2382.1	*TCF4*	Pitt-Hopkins syndrome (MIM #610954)	chr18:g.52896231G>A	NM_001083962.1:c.1726C>T	p.(Arg576*)	Nonsense	18	Probably pathogenic
Year 3	PED2438.1	*KMT2D*	Kabuki syndrome 1 (MIM #147920)	chr12:g.49420108C>T	NM_003482.3:c.15641G>A	p.(Arg5214His)	Missense	48	Pathogenic
Year 2	PED2504.1	*DNM1L*	Encephalopathy due to defective mitochondrial and peroxisomal fission-1 (MIM #614388)	chr12:g.32883953G>A	NM_005690.4:c.1085G>A	p.(Gly362Asp)	Missense	10	Probably pathogenic
Year 2	PED2574.1	*NR5A1*	46,XY sex reversal 3 (MIM #612965)	chr9:g.127255309C>T	NM_004959.4:c.990G>A	p.(Glu330Glu)	Missense	5	Probably pathogenic
Year 2	PED2595.1	*NSD1*	Sotos syndrome 1 (MIM #117550)	chr5:g.176675222_176675223insTCCT	NM_022455.4:c.4538_4539insTCCT	p.(Glu1513Aspfs*5)	Frameshift	11	Probably pathogenic
Year 3	PED2641.1	*TUBB3*	Complex cortical dysplasia with other brain malformations-1 (MIM #614039)	chr16:g.90001151G>A	NM_006086.3:c.292G>A	p.(Gly98Ser)	Missense	4	Probably pathogenic
Year 3	PED2642.1	*UBE3A*	Angelman syndrome (MIM #105830)	chr15:g.25615845_25615849delATAAT	NM_000462.3:c.1481_1485del	p.(Tyr494*)	Frameshift	7	Pathogenic
Year 3	PED2653.1	*BCL11A*	Dias-Logan syndrome (MIM #617101)	chr2:g.60773337G>A	NM_018014.3:c.154C>T	p.(Gln52*)	Nonsense		Probably pathogenic
Year 3	PED2708.1	*PURA*	Autosomal dominant mental retardation-31 (MIM #616158)	chr5:g.139494196A>G	NM_005859.4:c.430A>G	p.(Lys144Glu)	Missense	1	Probably pathogenic
Year 3	PED2709.1	*PPP2R1A*	Autosomal dominant mental retardation-36 (MIM #616362)	chr19:g.52716212C>T	NM_014225.5:c.656C>T	p.(Ser219Leu)	Missense	6	Probably pathogenic
Year 3	PED2740.1	*DYRK1A*	Autosomal dominant mental retardation-7 (MIM #614104)	chr21:g.38862764_38862767delGTAA	NM_001396.3:c.951+1_951+4del	Intron 6/10	Splice site	6	Probably pathogenic
Year 3	PED2757.1	*POGZ*	White-Sutton syndrome (MIM #616364)	chr1:g.151379386delC	NM_001194937.1:c.2518+1delG	Intron 17/18	Splice site	17	Probably pathogenic

Year 3	PED2763.1	*PPP2R5C*	PMID: 25972378, 26168268	chr14:g.102348567G>A	NM_002719.3:c.364G>A	p.(Glu122Lys)	Missense	3	Probably pathogenic
Year 3	PED2764.1	*SETD5*	Autosomal dominant mental retardation-23 (MIM #615761)	chr3:g.9482242C>T	NM_001080517.1:c.670C>T	p.(Gln224*)	Nonsense	8	Probably pathogenic
Year 3	PED2847.1	*POGZ*	White-Sutton syndrome (MIM #616364)	chr1:g.151378675delC	NM_001194937.1:c.2809delG	p.(Asp937Metfs*12)	Frameshift	19	Probably pathogenic
Year 3	PED2870.1	*SETD5*	Autosomal dominant mental retardation-23 (MIM #615761)	chr3:g.9488916_9488917delCT	NM_001080517.1:c.1709_1710del	p.(Ser570Cysfs*42)	Frameshift	14	Probably pathogenic
Year 3	PED2880.1	*KANSL1*	Koolen-De Vries syndrome (MIM #610443)	chr17:g.44248634delA	NM_015443.3:c.876delT	p.(Asp293Thrfs*37)	Frameshift	2	Probably pathogenic
Year 3	PED3002.1	*COL2A1*	Stanescu type of spondyloepiphyseal dysplasia (MIM #616583)	chr12:g.48370332C>G	NM_001844.4:c.3454G>C	p.(Gly1152Arg)	Missense	49	Pathogenic
Year 3	PED3005.1	*CHD7*	CHARGE syndrome (MIM #214800)	chr8:g.61729019C>T	NM_017780.3:c.2572C>T	p.(Arg858*)	Nonsense	8	Pathogenic
Year 3	PED3295.1	*CREBBP*	Rubinstein-Taybi syndrome-1 (MIM #180849)	chr16:g.3779564_3779566delGTA	NM_004380.2:c.5482_5484del	p.(Tyr1828del)	In/del	31	Probably pathogenic
Year 3	PED3301.1	*EP300*	Rubinstein-Taybi syndrome-2 (MIM #613684)	chr22:g.41551022C>T	NM_001429.3:c.3166C>T	p.(Gln1056*)	Nonsense	17	Pathogenic
Year 3	PED3368.1	*DYRK1A*	Autosomal dominant mental retardation-7 (MIM #614104)	chr21:g.38862511_38862512delTC	NM_001396.3:c.702_703del	p.(Cys235Phefs*4)	Frameshift	6	Pathogenic
Year 3	PED3453.1	*BRAF*	Cardiofaciocutaneous syndrome-1 (MIM #115150)	chr7:g.140449165A>T	NM_004333.4:c.1914T>A	p.(Asp638Glu)	Missense	16	Pathogenic
Year 3	PED3489.1	*AHDC1*	Xia-Gibbs syndrome (MIM #615829)	chr1:g.27876565G>A	NM_001029882.2:c.2062C>T	p.(Arg688*)	Nonsense	6	Pathogenic
Year 3	PED3490.1	*SYNGAP1*	Autosomal dominant mental retardation-5 (MIM #612621)	chr6:g.33403367C>T	NM_006772.2:c.739C>T	p.(Gln247*)	Nonsense	7	Pathogenic
Year 3	PED3501.1	*ACTG2*	Visceral myopathy (MIM #155310)	chr2:g.74135881C>G	NM_001615.3:c.337C>G	p.(Pro113Ala)	Missense	4	Probably pathogenic
Year 3	PED3627.1	*PURA*	Autosomal dominant mental retardation-31 (MIM #616158)	chr5:g.139494277C>G	NM_005859.3:c.511C>G	p.(Arg171Gly)	Missense	1	Probably pathogenic
Year 3	PED3646.1	*HCN1*	Early infantile epileptic encephalopathy-24 (MIM #615871)	chr5:g.45695782delG	NM_021072.3:c.414delC	p.(Tyr138*)	Frameshift	1	Pathogenic
Year 3	PED3647.1	*NFIX*	Sotos syndrome-2 (MIM #614753)	chr19:g.13136274G>C	NM_002501.2:c.467G>C	p.(Cys156Ser)	Missense	2	Probably pathogenic
Year 3	PED3677.1	*PUF60*	Verheij syndrome (MIM #615583)	chr8:g.144898991T>C	NM_001136033.2:c.1252-2A>G	Intron 11/11	Splice site	intron 11	Probably pathogenic

Year 3	PED3716.1	*OTX2*	Microphthalmia syndromic-5 (MIM #610125)	chr14:g.57268563_57268566delTGGT	NM_021728.3:c.781_784del	p.(Thr261Leufs*40)	Frameshift	5	Pathogenic
	Reanalysis								
Year 2	PED0047.1	*EEF1A2*	Early infantile epileptic encephalopathy-33 (MIM #616409)	chr20:g.62126409C>T	NM_001958.3:c.370G>A	p.(Glu124Lys)	Missense	4	Probably pathogenic
Year 1	PED1478.2	*TCF4*	Pitt-Hopkins syndrome (MIM #610954)	chr18:g.52896247C>G	NM_001083962.1:c.1710G>C	p.(Arg570Ser)	Missense	18	Probably pathogenic
Year 1	PED1525.1	*PUF60*	Verheij syndrome (MIM #615583)	chr8:g.144911449C>G	NM_014281.3:c.24+1G>C	Intron 1/10	Splice site	1	Unknown significance
Year 2	PED2120.1	*TUBB3*	Congenital fibrosis of extraocular muscles-3A (MIM #600638)	chr16:g.90001644G>A	NM_006086.3:c.785G>A	p.(Arg262His)	Missense	4	Probably pathogenic
Year 2	PED2185.1	*TBL1XR1*	Autosomal dominant mental retardation-41 (MIM #616944)	chr3:g.176750839A>G	NM_024665.4:c.1336T>C	p.(Tyr446His)	Missense	14	Probably pathogenic
Year 2	PED2237.1	*GABRA1*	Early infantile epileptic encephalopathy-19 (MIM #615744)	chr5:g.161324264G>C	NM_000806.5:c.1207G>C	p.(Glu403Gln)	Missense	11	Probably pathogenic
Year 2	PED2373.1	*KMT2A*	Wiedemann-Steiner syndrome (MIM #605130)	chr11:g.118348811G>A	NM_005933.3:c.3464G>A	p.(Cys1155Tyr)	Missense	5	Pathogenic
Year 2	PED2604.1	*PPP2R5D*	Autosomal dominant mental retardation-35 (MIM #616355)	chr6:g.42975003G>A	NM_006245.3:c.592G>A	p.(Glu198Lys)	Missense	5	Pathogenic
Year 1	PED2646.1	*KCNA2*	Early infantile epileptic encephalopathy-32 (MIM #616366)	chr1:g.111146182A>G	NM_004974.3:c.1223T>C	p.(Val408Ala)	Missense	3	Pathogenic
Autosomal dominant inherited									
Year 1	PED1028.1	*PAX9*	Selective tooth agenesis-3 (MIM #604625)	chr14:g.37132148C>G	NM_006194.3:c.51C>G	p.(Asn17Lys)	Missense	2	Probably pathogenic
Year 1	PED1181.1	*FGFR2*	Lacrimoauriculodentodigital syndrome (MIM #149730)	chr10 :g.123247549C>T	NM_000141.4:c.1942G>A	p.(Ala648Thr)	Missense	14	Pathogenic
Year 1	PED1187.1	*ABCC9*	Cantu syndrome (MIM #239850)	chr12:g.22047065G>A	NM_005691.2:c.1703C>T	p.(Pro568Leu)	Missense	12	Probably pathogenic
Year 1	PED1431.1	*LMX1B*	Nail-patella syndrome (MIM #161200)	chr9 :g.129455806C>G	NM_001174146.1:c.745C>G	p.(Arg249Gly)	Missense	5	Probably pathogenic
Year 2	PED2087.1	*SHH*	Isolated microphthalmia with coloboma-5 (MIM #611638)	chr7:g.155595881C>A	NM_000193.2:c.1102G>T	p.(Glu368*)	Nonsense	1	Probably pathogenic
Year 3	PED2332.1	*DEPDC5*	Familial focal epilepsy with variable foci (MIM #604364)	chr22:g.32174093_32174094insCTGG	NM_001242896.1:c.422_423insCTGG	p.(Gly142Trpfs*3)	Frameshift	8	Probably pathogenic

Year 3	PED2612.1	*NR2F1*	Bosch-Boonstra-Schaaf optic atrophy syndrome (MIM #615722)	chr5:g.92920819_92920828delCCGCGGCGGC	NM_005654.4:c.90_99delCCGCGGCGGC	p.(Arg31Alafs*85)	Frameshift	1	Probably pathogenic
Year 3	PED2654.1	*TNNI3K*	Cardiac conduction disease with or without dilated cardiomyopathy (MIM #616117)	chr1:g.74957901G>A	NM_015978.2:c.2302G>A	p.(Glu768Lys)	Missense	23	Probably pathogenic
Year 3	PED2683.1	*KMT2A*	Wiedemann-Steiner syndrome (MIM #605130)	chr11:g.118374788_118374791delAAGG	NM_005933.3:c.8174_8177delAAGG	p.(Glu2725Valfs*28)	Frameshift	27	Probably pathogenic
Year 3	PED3384.1	*SLC6A1*	Myoclonic-atonic epilepsy (MIM #616421)	chr3:g.11067221delC	NM_003042.3:c.801delC	p.(Ile268Serfs*36)	Frameshift	8	Probably pathogenic
Autosomal dominant undetermined (one parent unavailable or deceased)									
Year 3	PED2491.1	*NSD1*	Sotos syndrome 1 (MIM #117550)	chr5:g.176715924A>T	NM_022455.4:c.6256A>T	p.(Lys2086*)	Nonsense	21	Probably pathogenic
Year 3	PED3024.1	*ANKRD11*	KBG syndrome (MIM #148050)	chr16:g.89350623A>C	NM_001256183.1:c.2327T>G	p.(Leu776*)	Nonsense	9	Probably pathogenic
Year 3	PED3340.1	*MED13L*	Mental retardation and distinctive facial features with or without cardiac defects (MIM #616789)	chr12:g.116445389G>A	NM_015335.4:c.2065C>T	p.(Gln689*)	Nonsense	11	Probably pathogenic
Autosomal recessive Homozygous									
Year 1	PED1336.1	*TAF2*	Autosomal recessive mental retardation-40 (MIM #615599)	chr8:g.120774682G>A	NM_003184.3:c.2531C>T	p.(Pro844Leu)	Missense	19	Probably pathogenic
Year 2	PED1481.2	*SPR*	Dopa-responsive dystonia due to sepiapterin reductase deficiency (MIM #612716)	chr2:g.73114566_73114567insGGGCGGGCTG	NM_003124.4:c.18_19insGGGCGGGCTG	p.(Arg7Glyfs*37)	Frameshift	1	Probably pathogenic
Year 2	PED2193.1	*TTC21B*	Nephronophthisis-12 (MIM #613820)	chr2:g.166797621G>A	NM_024753.3:c.626C>T	p.(Pro209Leu)	Missense	6	Pathogenic
Year 2	PED2239.1	*POC1A*	Short stature, onychodysplasia, facial dysmorphism, and hypotrichosis (MIM #614813)	chr3:g.52183854C>G	NM_001161580.1:c.253G>C	p.(Val85Leu)	Missense	3	Probably pathogenic
Year 2	PED2255.1	*HPS3*	Hermansky-Pudlak syndrome-3 (MIM #614072)	chr3:g.148876559C>T	NM_032383.3:c.1798C>T	p.(Gln600*)	Nonsense	10	Probably pathogenic
Year 3	PED2274.1	*TRAPPC9*	Autosomal recessive mental retardation-13 (MIM #613192)	chr8:g.141370230G>A	NM_001160372.1:c.1414C>T	p.(Arg472*)	Nonsense	9	Probably pathogenic
Year 2	PED2336.4	*RNASEH2B*	Aicardi-Goutieres syndrome-2 (MIM #610181)	chr13:g.51519581G>A	NM_001142279.1:c.529G>A	p.(Ala177Thr)	Missense	7	Pathogenic
Year 2	PED2599.1	*ST3GAL5*	Amish infantile epilepsy syndrome (MIM #609056)	chr2:g.86073609C>T	NM_003896.3:c.740G>A	p.(Gly247Asp)	Missense	5	Probably pathogenic

Year 3	PED2700.1	*TRIM37*	Mulibrey Syndrome (MIM #253250)	chr17:g.57134281_57134282insGCGAGTAAGTC	NM_015294.3:c.1153_1154insGACTTACTCGC	p.(Ala385Glyfs*10)	Frameshift	13	Probably pathogenic
Year 3	PED2712.1	*MRE11A*	Ataxia-telangiectasia-like disorder (MIM #604391)	chr11:g.94212892T>C	NM_005591.3:c.350A>G	p.(Asn117Ser)	Missense	5	Pathogenic
Year 3	PED3257.1	*UBE3B*	Kaufman oculocerebrofacial syndrome (MIM #244450)	chr12:g.109967742_109967743dupA	NM_130466.3:c.2675_2676dupA	p.(Lys893*)	Frameshift	25	Probably pathogenic
Year 3	PED3298.1	*VPS13B*	Cohen syndrome (MIM #216550)	chr8:g.100160140C>T	NM_152564.4:c.1915C>T	p.(Arg639*)	Nonsense	14	Probably pathogenic
	Reanalysis								
Year 1	PED1065.1	*KLHL7*	PMID: 27392078	chr7:g.23180564G>A	NM_018846.4:c.474+1G>A	Intron 6/11	Splice site		Probably pathogenic
Year 1	PED1188.1	*MAB21L1*	PMID: 27103078	chr13:g.36049541dupC	NM_005584.3:c.735dupG	p.(Cys246Leufs*18)	Frameshift	2	Probably pathogenic
Year 1	PED1236.4	*SLC13A5*	Early-onset epileptic encephalopathy-25 (MIM #615905)	chr17:g.6590960A>G	NM_177550.3:c.1463T>C	p.(Leu488Pro)	Missense	11	Probably pathogenic
Year 1	PED1995.1	*FIBP*	Thauvin-Robinet-Faivre syndrome (MIM #617107)	chr11:g.65652652G>A	NM_004214.4:c.652C>T	p.(Gln218*)	Nonsense	6	Probably pathogenic
Autosomal recessive	Compound heterozygous								
Year 1	PED1356.1	*PKHD1*	Autosomal recessive polycystic kidney disease (MIM #263200)	chr6:g.51524480G>A	NM_138694.3:c.10444C>T	p.(Arg3482Cys)	Missense	61	Probably pathogenic
	PED1356.1		Autosomal recessive polycystic kidney disease (MIM #263200)	chr6:g.51882287C>T	NM_138694.3:c.5521G>A	p.(Glu1841Lys)	Missense	34	Probably pathogenic
Year 2	PED1540.1	*SH3TC2*	Charcot-Marie-Tooth disease type 4C (MIM #601596)	chr5:g.148406864G>A	NM_024577.3:c.2431C>T	p.(Gln811*)	Nonsense	11	Probably pathogenic
	PED1540.1		Charcot-Marie-Tooth disease type 4C (MIM #601596)	chr5:g.148384473C>T	NM_024577.3:c.3676-8G>A	Intron 16/16)	Splice site		Unknown significance
Year 2	PED1674.1	*CLN3*	Neuronal ceroid lipofuscinosis-3 (MIM #204200)	chr16:g.28493821C>T	NM_000086.2:c.883G>A	p.(Glu295Lys)	Missense	11	Pathogenic
	PED1674.1		second event is CNV reported below						
Year 1	PED2037.1	*ADCK3*	Primary coenzyme Q10 deficiency-4 (MIM #612016)	chr1:g.227169808C>T	NM_020247.4:c.811C>T	p.(Arg271Cys)	Missense	6	Unknown significance
	PED2037.1		Primary coenzyme Q10 deficiency-4 (MIM #612016)	chr1:g.227173005CAT>C	NM_020247.4:c.1625_1626del	p.(Ile542Argfs*31)	Frameshift	14	Probably pathogenic

Year 2	PED2238.1	*MAN2B1*	Alpha-mannosidosis (MIM#248500)	chr19:g.12758091_12758106delTGGTTGGCCACCAGCG	NM_000528.3:c.2864_2879delCGCTGGTGGCCAACCA	p.(Thr955Serfs*73)	Frameshift	23	Pathogenic
	PED2238.1		Alpha-mannosidosis (MIM#248500)	chr19:g.12776184G>A	NM_000528.3:c.418C>T	p.(Arg140*)	Nonsense	3	Pathogenic
Year 2	PED2300.1	*TTN*	Early-onset myopathy with fatal cardiomyopathy (MIM #611705)	chr2:g.179463790C>T	NM_001256850.1:c.51725-1G>A	Intron 240	Splice site		Probably pathogenic
	PED2300.1		Early-onset myopathy with fatal cardiomyopathy (MIM #611705)	chr2:g.179464280C>T	NM_001256850.1:c.51424+1G>A	Intron 239	Splice site		Probably pathogenic
Year 2	PED2335.1	*NGLY1*	Congenital disorder of deglycosylation (MIM #615273)	chr3:g.25770747_25770754delACAAAGGT	NM_001145293.1:c.1427_1434delACCTTTGT	p.(His476Leufs*14)	Frameshift	10	Probably pathogenic
	PED2335.1		Congenital disorder of deglycosylation (MIM #615273)	chr3:g.25778897C>T	NM_001145293.1:c.931G>A	p.(Glu311Lys)	Missense	6	Probably pathogenic
Year 3	PED3072.1	*VPS13B*	Cohen syndrome (MIM #216550)	chr8:g.100789002G>A	NM_017890.3:c.7323-1G>A	Intron 40/61	Splice site	41?	Probably pathogenic
	PED3072.1		Cohen syndrome (MIM #216550)	chr8:g.100866389_100866391delTCT	NM_017890.3:c.10847_10849delTCT	p.(Phe3616del)	In/del	56	Pathogenic
Year 3	PED3167.1	*ANO10*	Autosomal recessive spinocerebellar ataxia-10 (MIM #613728)	chr3:g.43596901A>G	NM_018075.3:c.1537T>C	p.(Cys513Arg)	Missense	10	Probably pathogenic
	PED3167.1		Autosomal recessive spinocerebellar ataxia-10 (MIM #613728)	chr3:g.43647212_43647213insT	NM_018075.3:c.132dupA	p.(Asp45Argfs*9)	Frameshift	2	Probably pathogenic
Year 3	PED3327.1	*PPT1*	Neuronal ceroid lipofuscinosis-1 (MIM #256730)	chr1:g.40546155C>T	NM_000310.3:c.541G>A	p.(Val181Met)	Missense	6	Pathogenic
	PED3327.1		Neuronal ceroid lipofuscinosis-1 (MIM #256730)	chr1:g.40555147delA	NM_000310.3:c.471delT	p.(His158Thrfs*10)	Frameshift	5	Probably pathogenic
Year 3	PED3374.1	*ADCK3*	Primary coenzyme Q10 deficiency-4 (MIM #612016)	chr1:g.227153421G>A	NM_020247.4:c.638G>A	p.(Arg213Gln)	Missense	4	Probably pathogenic
	PED3374.1		Primary coenzyme Q10 deficiency-4 (MIM #612016)	chr1:g.227172321T>A	NM_020247.4:c.1471T>A	p.(Trp491Arg)	Missense	12	Unknown significance
Year 3	PED3414.1	*ATIC*	AICAR transformylase/IMP cyclohydrolase deficiency (MIM#608688)	chr2:g.216190736G>A	NM_004044.6:c.406G>A	p.(Ala136Thr)	Missense	6	Unknown significance
	PED3414.1		AICAR transformylase/IMP cyclohydrolase deficiency (MIM#608688)	chr2:g.216213967A>T	NM_004044.6:c.1654A>T	p.(Lys552*)	Nonsense	15	Pathogenic
Year 3	PED3508.1	*LARP7*	Alazami syndrome (MIM #615071)	chr4:g.113567491_113567531delTTTTCATTTTCTTTAGATGTTGATATATCACTACTTGTGTC	NM_015454.1:c.203-16_227delTTTTCATTTTCTTTAGATGTTGATATATCACTACTTGTGTC	Intron 4/14	Splice site		Probably pathogenic

	PED3508.1		Alazami syndrome (MIM #615071)	chr4:g.113568938_113568941delAAAG	NM_015454.1:c.1091_1094delAAGA	p.(Lys364Argfs*12)	Frameshift	10	Probably pathogenic
Year 3	PED3570.1	*PGAP3*	Hyperphosphatasia with mental retardation syndrome-4 (MIM #615716)	chr17:g.37829358T>C	NM_033419.3:c.845A>G	p.(Asp282Gly)	Missense	7	Pathogenic
	PED3570.1		Hyperphosphatasia with mental retardation syndrome-4 (MIM #615716)	chr17:g.37841002delA	NM_033419.3:c.280delT	p.(Trp94Glyfs*27)	Frameshift	3	Pathogenic
	Reanalysis								
Year 1	PED1312.1	*SLC13A5*	Early-onset epileptic encephalopathy-25 (MIM #615905)	chr17:g.6606325G>A	NM_001143838.1:c.680C>T	p.(Thr227Met)	Missense	5	Pathogenic
	PED1312.1		Early-onset epileptic encephalopathy-25 (MIM #615905)	chr17:g.6606350C>T	NM_001143838.1:c.655G>A	p.(Gly219Arg)	Missense	5	Pathogenic
Year 1	PED1906.1	*GFER*	Myopathy with cataract and combined respiratory chain deficiency (MIM #613076)	chr16:g.2034438delC	NM_005262.2:c.219delC	p.(Cys74Alafs*76)	Frameshift	1	Probably pathogenic
	PED1906.1		Myopathy with cataract and combined respiratory chain deficiency (MIM #613076)	chr16:g.2034722_2034723delCA	NM_005262.2:c.259-26_259-25delCA		Alternative transcript		Probably pathogenic
Year 1	PED2118.1	*THOC6*	Beaulieu-Boycott-Innes syndrome (MIM #613680)	chr16:g.3075804C>A	NM_001142350.1:c.135C>A	p.(Tyr45*)	Nonsense	2	Probably pathogenic
	PED2118.1		Beaulieu-Boycott-Innes syndrome (MIM #613680)	chr16:g.3076765G>A	NM_001142350.1:c.569G>A	p.(Gly190Glu)	Missense	8	Unknown significance
Year 1	PED2645.1	*AP3B2*	PMID: 27889060	chr15:g.83348481C>T	NM_001278511.1:c.1086G>A	p.(=)	Silencing	9	Probably pathogenic
	PED2645.1		PMID: 27889060	chr15:g.83348926C>G	NM_001278511.1:c.1014+1G>C	Intron 8/24	Splice site		Probably pathogenic
X-linked de novo									
Year 2	PED2404.1	*ATRX*	X-linked alpha-thalassemia/mental retardation syndrome (MIM #301040)	chrX:g.76972632G>A	NM_000489.3:c.109C>T	p.(Arg37*)	Nonsense	2	Pathogenic
Year 3	PED2895.1	*MED12*	Opitz-Kaveggia syndrome (MIM #305450)	chrX:g.70348993G>T	NM_005120.2:c.3505G>T	p.(Ala1169Ser)	Missense	25	Probably pathogenic
Year 3	PED3044.1	*KDM5C*	Claes-Jensen type syndromic X-linked mental retardation (MIM #300534)	chrX:g.53246393dupG	NM_001146702.1:c.388dupC	p.(Leu130Profs*23)	Frameshift	3	Probably pathogenic
Year 3	PED3650.1	*CDKL5*	Early infantile epileptic encephalopathy-2 (MIM #300672)	chrX:g.18593504G>A	NM_003159.2:c.176G>A	p.(Arg59Gln)	Missense	12	Probably pathogenic
	Reanalysis								
Year 2	PED0055.1	*KDM5C*	Claes-Jensen type syndromic X-linked mental retardation (MIM #300534)	chrX:g.53231036C>A	NM_004187.3:c.1866G>T	p.(Trp622Cys)	Missense	13	Unknown significance
Year 1	PED1342.1	*NAA10*	PMID: 26522270	chrX:g.153197528A>T	NM_001256120.1:c.364T>A	p.(Phe122Ile)	Missense	6	Probably pathogenic

Year 2	PED2489.1	*DDX3X*	X-linked mental retardation-10 (MIM #300958)	chrX:g.41203348_41203349dupT	NM_001193416.1:c.831_832dup	p.(Leu278Cysfs*44)	Frameshift	9	Probably pathogenic
Year 2	PED2601.1	*DDX3X*	X-linked mental retardation-10 (MIM #300958)	chrX:g.41203514G>C	NM_001193416.1:c.887G>C	p.(Arg296Pro)	Missense	10	Probably pathogenic
X-linked Dominant									
Year 3	PED3010.1	*DDX3X*	X-linked mental retardation-10 (MIM #300958)	chrX:g.41203342_41203343delGA	NM_001356.3:c.830_831delAG	p.(Glu277Valfs*17)	Frameshift	9	Probably pathogenic
Year 3	PED3300.1	*NHS*	Nance-Horan syndrome (MIM #302350)	chrX:g.17742490C>T	NM_198270.2:c.1117C>T	p.(Arg373*)	Nonsense	5	Pathogenic
Year 1	PED1338.1	*SLC16A2*	Allan-Herndon-Dudley syndrome (MIM #300523)	chrX:g.73740833G>A	NM_006517.4:c.439G>A	p.(Gly147Arg)	Missense	2	Pathogenic
X-linked Recessive									
Year 1	PED1314.1	*CUL4B*	Cabezas type of X-linked syndromic mental retardation (MIM #300354)	chrX:g.119675504G>A	NM_001079872.1:c.1396C>T	p.(Arg466*)	Nonsense	7	Probably pathogenic
Year 3	PED2911.1	*ARHGEF9*	Early infantile epileptic encephalopathy 8 (MIM #300607)	chrX:g.62885872G>A	NM_015185.2:c.950C>T	p.(Ser317Leu)	Missense	7	Unknown significance
Year 3	PED3376.1	*TRAPPC2*	Spondyloepiphyseal dysplasia tarda (MIM #313400)	chrX:g.13752240C>T	NM_001128835.2:c.12G>A	p.(Trp4*)	Nonsense	2	Probably pathogenic
Year 3	PED3454.1	*PHF8*	Siderius type X-linked syndromic mental retardation (MIM #300263)	chrX:g.53989374T>C	NM_001184896.1:c.2552-2A>G	Intron 18/21	Splice site		Probably pathogenic
Carrier									
Year 2	PED2255.1	*TTPA*	Ataxia with vitamin E deficiency (MIM #277460)	chr8:g.63973904delT	NM_000370.2:c.744delA	p.(Glu249Asnfs*15)	Frameshift	5	Pathogenic
CNVs									
Autosomal dominant de novo									
Year 3	PED2391.1			chr1:59922631-72058653			Deletion		Pathogenic
Year 3	PED2625		Marfan syndrome (MIM #154700)	chr15:48051994-48830007			Deletion		Pathogenic
Autosomal dominant inherited									
Year 2	PED1674	*CLN3*	Neuronal ceroid lipofuscinosis-3 (MIM #204200)	chr16:28497282-28498403			Deletion		Pathogenic
Autosomal dominant undetermined (one parent unavailable or deceased)									
Year 3	PED2773	*TCF4*	Pitt-Hopkins syndrome (MIM #610954)	chr18:52921549-52922186			Deletion	16	Probably pathogenic
Autosomal recessive Homozygous									
Year 3	PED2770	*PPT1*	Neuronal ceroid lipofuscinosis-1 (MIM #256730)	chr1:40558255-40562842			Deletion	Intron 1	Pathogenic

Supplemental table 2: List of results of unknown significance

SNVs									
Year	*Patient's ID*	Gene	Pathology	Genomic annotation	cDNA	Protein	Mutation type	Exon	Classification
Autosomal dominant de novo									
Year 1	*PED0464.1*	*PACS2*	PMID: 22488736	chr14:g.105834449G>A	NM_001100913.2:c.625G>A	p.(Glu209Lys)	Missense	6	Unknown significance
Year 3	*PED0732.1*	*STAG1*	PMID : 23020937, 24896178	chr3:g.136162242T>G	NM_005862.2:c.1433A>C	p.(His478Pro)	Missense	15	Unknown significance
Year 1	*PED1185.1*	*TCF4*	Pitt-Hopkins syndrome (MIM #610954)	chr18:g.52896176A>T	NM_001083962.1:c.1781T>A	p.(Met594Lys)	Missense	18	Unknown significance
Year 1	*PED1309.1*	*KCNMA1*	Generalized epilepsy and paroxysmal dyskinesia (MIM #609446)	chr10:g.78869939C>T	NM_001014797.2:c.1123G>A	p.(Gly375Arg)	Missense	8	Unknown significance
Year 3	*PED3006.1*	*KCNMA1*	Generalized epilepsy and paroxysmal dyskinesia (MIM #609446)	chr10:g.78869939C>T	NM_001014797.2:c.1123G>A	p.(Gly375Arg)	Missense	8	Unknown significance
Year 3	*PED2628.1*	*ANKRD11*	KBG syndrome (MIM #148050)	chr16:g.89345578T>C	NM_001256182.1:c.7372A>G	p.(Ile2458Val)	Missense	10	Unknown significance
Year 3	*PED2903.1*	*CACNA1G*	Spinocerebellar ataxia-42 (MIM #616795)	chr17:g.48685266A>G	NM_018896.4:c.4591A>G	p.(Met1531Val)	Missense	25	Unknown significance
Year 2	*PED2234.1*	*PACS2*	PMID: 22488736	chr14:g.105834449G>A	NM_001100913.2:c.625G>A	p.(Glu209Lys)	Missense	6	Unknown significance
Year 2	*PED2291.1*	*CHRNA7*	Chromosome 15q13.3 microdeletion syndrome (MIM #612001)	chr15:g.32450712A>G	NM_000746.3:c.698A>G	p.(Tyr233Cys)	Missense	7	Unknown significance
	Reanalysis								
Year 1	*PED1310.1*	*SMARCC2*	PMID: 27620904	chr12:g.56566216A>G	NM_001130420.1:c.1922T>C	p.(Leu641Pro)	Missense	20	Unknown significance
Autosomal dominant inherited									
Year 1	PED1186.1	*ZFPM2*	Diaphragmatic hernia-3 (MIM #610187); Tetralogy of Fallot (MIM #187500)	chr8:g.106801034delA	NM_012082.3:c.621delA	p.(Ala208Leufs*5)	Frameshift	6	Unknown significance
Year 3	PED1658.1	*PITX1*	Congenital clubfoot with or without deficiency of long bones and/or mirror-image polydactyly (MIM #119800)	chr5:g.134365002T>G	NM_002653.4:c.412A>C	p.(Lys138Gln)	Missense	3	Unknown significance

Year 3	PED2658.1	*CIC*	PMID: 21076407	chr19:g.42796325C>T	NM_015125.3:c.2974C>T	p.(Gln992*)	Nonsense	12	Unknown significance
Year 3	PED2672.1	*YWHAE*	PMID: 23901910	chr17:g.1268169C>T	NM_006761.4:c.248G>A	p.(Arg83Gln)	Missense	2	Unknown significance
Year 3	PED2759.1	*TNNT3*	Distal arthrogryposis type 2B (MIM #601680)	chr11:g.1953744T>C	NM_001042780.2:c.147T>C	p.(=)	Silencing	9	Unknown significance
Year 3	PED3305.1	*PRKG1*	Thoracic aortic aneurysm-8 (MIM #615436)	chr10:g.52913043G>C	NM_006258.3:c.431G>C	p.(Cys144Ser)	Missense	2	Unknown significance
Year 3	PED3507.1	*TCF12*	Craniosynostosis-3 (MIM #615314)	chr15:g.57544630_57544631delTG	NM_207037.1:c.1200_1201delTG	p.(Glu401Alafs*19)	Frameshift	15	Unknown significance
Year 3	PED2081.1	*PLCB4*	Auriculocondylar syndrome-2 (MIM #614669)	chr20:g.9389801A>T	NM_000933.3:c.1936A>T	p.(Met646Leu)	Missense	20	Unknown significance
Year 3	PED3341.1	*PPP2R5D*	Autosomal dominant mental retardation-35 (MIM #616355)	chr6:g.42975739A>C	NM_001270476.1:c.340A>C	p.(Ile114Leu)	Missense	7	Unknown significance
Autosomal dominant undetermined (one parent unavailable or deceased)									
Year 2	PED1139.1	*SMARCB1*	Coffin-Siris syndrome-3 (MIM #614608)	chr22:g.24175842C>G	NM_003073.3:c.1070C>G	p.(Thr357Arg)	Missense	8	Unknown significance
Year 3	PED3297.1	*DYNC1H1*	Mental retardation-13 (MIM #614563)	chr14:g.102445718G>A	NM_001376.4:c.407G>A	p.(Arg136Gln)	Missense	3	Unknown significance
	Reanalysis								
Year 2	PED2244.1	*KIRREL3*	Autosomal dominant mental retardation-4 (MIM #612581)	chr11:g.126310436_126310437insG	NM_032531.3:c.1260_1261insC	p.(Ile421Hisfs*90)	Frameshift	11	Unknown significance
Autosomal recessive	Homozygous								
Year 1	PED1177.1	*CADPS*		chr3:g.62543079C>A	NM_003716.3:c.1753+1G>T	Intron 10/29	Splice site		Unknown significance
Year 1	PED1333.1	*SCN10A*	Familial episodic pain syndrome-2 (MIM #615551)	chr3:g.38739976G>A	NM_006514.2:c.4735C>T	p.(Arg1579*)	Nonsense	27	Unknown significance
Year 3	PED3249.1	*PEX6*	Heimler syndrome-2 (MIM #616617)	chr6:g.42946450C>T	NM_000287.3:c.439G>A	p.(Gly147Arg)	Missense	1	Unknown significance
Autosomal recessive	Compound heterozygous								
Year 1	PED1514.1	*UNC45A*	PMID: 16478993	chr15:g.91485808C>T	NM_001039675.1:c.784C>T	p.(Arg262*)	Nonsense	10	Unknown significance
		UNC45A	PMID: 16478993	chr15:g.91489957T>A	NM_001039675.1:c.1268T>A	p.(Val423Asp)	Missense	13	Unknown significance
Year 2	PED2230.1	*SACS*	Spastic ataxia of the Charlevoix-Saguenay (MIM #270550)	chr13:g.23912112G>A	NM_014363.5:c.5903C>T	p.(Ala1968Val)	Missense	10	Unknown significance
		SACS	Spastic ataxia of the Charlevoix-Saguenay (MIM #270550)	chr13:g.23914588G>T	NM_014363.5:c.3427C>A	p.(Gln1143Lys)	Missense	10	Unknown significance

Year 2	PED2297.1	*ARGFX*	PMID: 27412763	chr3:g.121303834T>G	NM_001012659.1:c.291T>G	p.(Phe97Leu)	Missense	4	Unknown significance
		ARGFX	PMID: 27412763	chr3:g.121305258_121305259delTC	NM_001012659.1:c.763_764delTC	p.(Ser255Thrfs*38)	Frameshift	4	Unknown significance
Year 2	PED2298.1	*RAB3GAP2*	Martsolf syndrome (MIM #212720)	chr1:g.220335609G>A	NM_012414.3:c.3156C>T	p.(=)	Silencing	28	Unknown significance
		RAB3GAP2	Martsolf syndrome (MIM #212720)	chr1:g.220363770G>A	NM_012414.3:c.1580C>T	p.(Pro527Leu)	Missense	15	Unknown significance
Year 2	PED2329.1	*MRPS22*	Combined oxidative phosphorylation deficiency-5 (MIM #611719)	chr3:g.139069822A>G	NM_020191.2:c.652A>G	p.(Met218Val)	Missense	5	Unknown significance
		MRPS22	Combined oxidative phosphorylation deficiency-5 (MIM #611719)	chr3:g.139069844T>C	NM_020191.2:c.674T>C	p.(Val225Ala)	Missense	5	Unknown significance
Year 2	PED2348.1	*KIAA1033*	Autosomal recessive mental retardation-43 (MIM #615817)	chr12:g.105534127A>G	NM_015275.1:c.1508A>G	p.(His503Arg)	Missense	7	Unknown significance
		KIAA1033	Autosomal recessive mental retardation-43 (MIM #615817)	chr12:g.105557967A>G	NM_015275.1:c.3236A>G	p.(Lys1079Arg)	Missense	13	Unknown significance
Year 3	PED2859.1	*SLC30A10*	Hypermanganesemia with dystonia-1 (MIM #613280)	chr1:g.220088979G>A	NM_018713.2:c.1270C>T	p.(His424Tyr)	Missense	4	Unknown significance
		SLC30A10	Hypermanganesemia with dystonia-1 (MIM #613280)	chr1:g.220101611C>T	NM_018713.2:c.172G>A	p.(Gly58Ser)	Missense	1	Unknown significance
Year 3	PED2688.2	*SLC5A7*	Autosomal dominant distal hereditary motor neuronopathy type VIIa (MIM #158580)	chr2:g.108604734_108604737delCATC	NM_021815.2:c.123_126del	p.(Ile42*)	Frameshift	2	Probably pathogenic
		SLC5A7	Autosomal dominant distal hereditary motor neuronopathy type VIIa (MIM #158580)	chr2:g.108625107G>A	NM_021815.2:c.1082G>A	p.(Arg361Gln)	Missense	8	Unknown significance
Year 3	PED3459.1	*COL6A1*	Bethlem myopathy 1 (MIM #158810); Ullrich congenital muscular dystrophy 1 (MIM #254090)	chr21:g.47406990A>G	NM_001848.2:c.717+4A>G	Intron 5/34	Splice site	5	Pathogenic
		COL6A1	Bethlem myopathy 1 (MIM #158810); Ullrich congenital muscular dystrophy 1 (MIM #254090)	chr21:g.47410687G>A	NM_001848.2:c.1003G>A	p.(Gly335Arg)	Missense	14	Unknown significance
Autosomic recessive 2nd variant missing									
Year 3	PED3289.1	*ABCA12*	Autosomal recessive ichthyosis-4B (harlequin) (MIM #242500); Autosomal recessive congenital ichthyosis-4A (MIM #601277)	chr2:g.215802332G>A	NM_015657.3:c.6490C>T	p.(Arg2164*)	Nonsense	43	Pathogenic

Year 2	PED2603.1 *DOCK7*	Early infantile epileptic encephalopathy-23 (MIM #615859)	chr1:g.63024804C>A	NM_001271999.1:c.2287G>T	p.(Glu763*)	Nonsense	20	Unknown significance
X-linked	De novo							
Year 1	PED1907.1 *MSL3*		chrX:g.11790375G>T	NM_001193270.1:c.1345+1G>T	Intron 11/12	Splice site	11	Unknown significance
X-linked Dominant								
Year 2	PED2065.1 *BRWD3*	X-linked mental retardation 93 (MIM #300659)	chrX:g.79959013C>G	NM_153252.4:c.2801G>C	p.(Arg934Pro)	Missense	24	Unknown significance
Year 2	PED2085.2 *HUWE1*	X-linked syndromic mental retardation, Turner type (MIM #300706)	chrX:g.53561162T>C	NM_031407.4:c.12832-4A>G	Intron 82/83	Splice site		Unknown significance
Year 2	PED2245.1 *SHROOM4*	Stocco dos Santos X-linked mental retardation syndrome (MIM #300434)	chrX:g.50378235G>A	NM_020717.3:c.838C>T	p.(Arg280Trp)	Missense	4	Unknown significance
Year 3	PED3285.1 *TMLHE*	Susceptibility to X-linked autism-6 (MIM #300872)	chrX:g.154754189delCinsGGTGAGTCTTAGG	NM_018196.3:c.286delinsCCTAAGACTCACC	p.(Asp96delinsProLysThrHisHis)	In/del	3	Unknown significance
Year 3	PED3322.1 *FAM58A*	STAR syndrome (MIM #300707)	chrX:g.152864503C>A	NM_152274.3:c.27G>T	p.(=)	Silencing	1	Unknown significance
Year 3	PED3354.1 *ATRX*	X-linked mental retardation-hypotonic facies syndrome (MIM #309580)	chrX:g.76938121G>A	NM_000489.3:c.2627C>T	p.(Ser876Phe)	Missense	9	Unknown significance
Year 3	PED3686.1 *NHS*	Nance-Horan syndrome (MIM #302350)	chrX:g.17394202G>A	NM_198270.2:c.322G>A	p.(Glu108Lys)	Missense	1	Unknown significance
X-linked undetermined (one parent unavailable or deceased)								
Year 3	PED2477.1 *PHF6*	Borjeson-Forssman-Lehmann syndrome (MIM #301900)	chrX:g.133547979G>A	NM_001015877.1:c.712G>A	p.(Ala238Thr)	Missense	7	Unknown significance

Supplemental table 2 : list of incidental findings

SNVs								
Year	Patient's ID	Gene	Pathology	Genomic annotation	cDNA	Protein	Mutation type	Exon	Classification
Autosomal dominant de novo									
Year 3	PED3384.1	*BRCA2*	Susceptibility to familial breast-ovarian cancer-2 (MIM #612555)	chr13:g.32914103_32914107delAGTAA	NM_000059.3:c.5616_5620delAGTAA	p.(Lys1872Asnfs*2)	Frameshift	11	Probably pathogenic
Autosomal dominant inherited									
Year 1	PED1318.1	*BRCA2*	Susceptibility to familial breast-ovarian cancer-2 (MIM #612555)	chr13:g.32910661_32910662insA	NM_000059.3:c.2169_2170insA	p.(Val726Serfs*25)	Frameshift	11	Pathogenic
Year 3	PED2568.1	*BRCA2*	Susceptibility to familial breast-ovarian cancer-2 (MIM #612555)	chr13:g.32968863C>A	NM_000059.3:c.9294C>A	p.(Tyr3098*)	Nonsense	25	Pathogenic
Year 3	PED3295.1	*SCN5A*	Susceptibility to long QT syndrome 3 (MIM #608567)	chr3:g.38592648G>A	NM_000335.4:c.5212C>T	p.(Arg1738Trp)	Missense	28	Pathogenic
Year 3	PED3407.1	*KCNQ1*	Susceptibility to long QT syndrome 1 (MIM #192500)	chr11:g.2869078G>A	NM_000218.2:c.1876G>A	p.(Gly626Ser)	Missense	16	Pathogenic
Year 3	PED3646.1	*BRCA1*	Susceptibility to familial breast-ovarian cancer-1 (MIM #604370)	chr17:g.41244214_41244217delCTTG	NM_007294.3:c.3331_3334delCAAG	p.(Gln1111Asnfs*5)	Frameshift	10	Pathogenic

TITRE DE LA THESE : Le séquençage d'exome dans le diagnostic des maladies rares avec anomalies du développement: intérêt de la réanalyse des données annuelle prospective

AUTEUR : NAMBOT SOPHIE

RESUME :

Les anomalies du développement et la déficience intellectuelle représentent un véritable défi diagnostique en génétique clinique actuelle, sachant que la moitié des patients restent sans base moléculaire identifiée après les évaluations traditionnelles, réalisées au cours d'une longue « odyssée » diagnostique. Pour améliorer le rendement diagnostique de notre centre de consultation génétique généraliste, nous avons introduit le séquençage d'exome en solo dans notre pratique clinique courante.

Afin d'évaluer l'intérêt et la rentabilité de cette stratégie, nous rapportons içi les résultats cliniques et moléculaires de 418 patients atteints d'anomalies du développement et/ou de déficience intellectuelle, ayant bénéficié d'une analyse d'exome en solo, associée à une réanalyse annuelle systématique des données pour les cas négatifs ou non-concluants.
Le rendement diagnostique global obtenu était de 31%. Le rendement diagnostique moyen par réanalyse prospective annuelle était de 7,5%, principalement lié au partage international de données permettant l'identification de 5 nouveaux gènes associés à une pathologie humaine. Cette stratégie a permis de diminuer le nombre moyen des examens complémentaires à visée étiologique, entraînant une réduction significative des coûts.

Nos résultats montrent la faisabilité et l'intérêt majeur de l'implantation du séquençage d'exome en solo dans la démarche diagnostique des anomalies du développement et de la déficience intellectuelle, y compris à l'échelle d'un petit centre clinique.

MOTS-CLES : séquençage d'exome, anomalies du développement, déficience intellectuelle, partage de données, réanalyse