

UNIVERSITE DE NANTES

FACULTE DE MEDECINE

Année 2016-2017

THESE

pour le

DIPLOME D'ETAT DE DOCTEUR EN MEDECINE

Spécialité Endocrinologie, Diabétologie et Maladies Métaboliques

par

Elias BARRAK

Né le 17 Septembre 1986 à Zahlé-Liban

Présentée et soutenue publiquement le 26 septembre 2017

Adénomes hypophysaires isolés familiaux :

Stratégie pour l'identification d'une nouvelle cause génétique :

Observation clinique, stratégie de recherche, et revue de la littérature

Président : Monsieur le Professeur Philippe CHANSON

Directeur : Monsieur le Professeur Bertrand CARIOU

فقل لمن يدعي في العلم فلسفة
حفظت شيئا وغابت عنك أشياء

Abû Nuwâs, VIII^e-IX^e siècles

*Dis à celui qui prétend tout connaître en science, que tu as bien retenu des choses mais que tu as oublié tout le reste.

Remerciements

À mon Président de thèse, Monsieur le Professeur Philippe CHANSON,

Je vous remercie de m'avoir fait l'honneur d'accepter la présidence de cette thèse, j'en suis touché et reconnaissant. Je vous remercie de m'avoir fait partager votre passion pour l'endocrinologie pendant tout un semestre. Veuillez recevoir l'expression de ma respectueuse gratitude.

À mon Directeur de thèse, Monsieur le Professeur Bertrand CARIOU,

Merci de m'avoir soutenu et dirigé pendant toutes mes années d'internat et de m'avoir accordé ta confiance tout au long de mon parcours. Merci de m'avoir fait découvrir et aimer l'aspect recherche de notre métier.

À Monsieur le Professeur Stéphane BEZIEAU,

Je vous remercie pour l'intérêt que vous avez immédiatement porté à mon égard. Vous me faites l'honneur de juger mon travail. Veuillez recevoir l'expression de ma respectueuse gratitude.

À Monsieur le Docteur Jean-Jacques SCHOTT,

Tu as si gentiment et spontanément accepté de participer au jury. Merci de ta disponibilité et de ton écoute. Je te suis très reconnaissant. Merci infiniment.

À Madame le Docteur Maëlle LE BRAS,

Merci de m'avoir transmis ta passion pour la médecine et tout particulièrement pour l'Endocrinologie. Sans toi, je n'aurais pas travaillé sur ce sujet. Mille fois Merci.

A mon père...

Abréviations

ADN	Acide Désoxyribonucléique
AIP	Aryl Hydrocarbon Receptor-Interacting Protein
ARN	Acide Ribonucléique
ARNase	Ribonucléase
ADNc	Acide Désoxyribonucléique complémentaire
CDKN1B	Cyclin-Dependent Kinase Inhibitor 1b
CG	Complete Genomics
Chr	Chromosome
CHU	Centre Hospitalo-Universitaire
CNC	Complexe de Carney
CNV	Copy number variation
COL5A2	Collagen, Type V, Alpha-2
CSPG5	Chondroitin Sulfate Proteoglycan 5
CST5	Cystatin 5
dbSNP	database Single Nucleotide polymorphism
DFO	Dysplasie fibreuse des Os
EGF	Epidermal growth factor
EGFR	Epidermal growth factor Receptor
EPS8	Epidermal growth factor receptor Pathway Substrate 8
ExAC	Exome Aggregation Consortium
FIPA	Familial isolated Pituitary Adenoma
FSH	Follicle Stimulating Hormone
GATK	Genome Analysis Tool Kit
GATK HapCaller	Genome Analysis Tool Kit HapCaller
GERP	Genomic Evolutionary Rate Profiling
IFS	Isolated Familial Somatotropinoma
IGF1	Insulin Growth Factor 1
IGV I	Integrative Genome Viewer
KCNH6	Potassium Channel, Voltage-Gated, Subfamily H, Member 6
LH	Hormone Lutéinisante

MAF	« Minor Allele Frequency » (Fréquence de l'allèle mineur)
MEN1	gène de la ménine
MYO3A	Myosin IIIa
N	Norme
NEM1	Néoplasie endocrinienne de type 1
NEM4	Néoplasie endocrinienne de type 4
NEMX	Néoplasie endocrinienne de type X
NFPA	Non Functional pituitary adenoma
NGS	« Next Generation Sequencing » (Séquençage Nouvelle Génération)
pb	paire(s) de bases
PIKFYVE	Phosphatidylinositol 3-Phosphate 5-Kinase, Type 3
PRKAR1A	Proteine kinase cAMP-dépendante de type 1 ; sous unité alpha
PRKCH	Protéine Kinase C
SMA	Syndrome de McCune-Albright
ST	Samtool
T3	Triiodothyronine
T4	Thyroxine
THADA	Thyroid Adenoma-Associated Gene
TNE	Tumeur Neuro-Endocrine
TNP2	Transition Protein 2
TSH	Thyroid Stimulating Hormone
VCF	« Variant Call Format » (Format de fichier référençant les variants identifiés)
VEP	Variant effect predictor
WIPF3	Was/Wasl-Interacting Protein Family, Member 3
WT	Wild type (sauvage)

SOMMAIRE

Introduction

- a. Epidémiologie
- b. Etiologie des adénomes hypophysaires
 - 1. Néoplasie endocrinienne multiple de type 1
 - 2. Néoplasie endocrinienne multiple de type 4
 - 3. Complexe de Carney
 - 4. Syndrome de McCune-Albright
 - 5. FIPA

Observation clinique

Méthodes et stratégies

- a. Stratégie de diagnostic clinique par séquençage ciblé
- b. Stratégie de recherche clinique
 - 1. Prélèvement sanguin, consentement et extraction de l'ADN
 - 2. Extraction de l'ARN
 - 3. Séquençage type EXOME
 - I. Séquençage à haut-débit – Illumina HiSeq 2500
 - II. Alignement et annotations des séquences
 - Alignement
 - Détection de variants («calling»)
 - Annotation des variants
 - III. Filtrage des variants identifiés à partir de l'outil Knime4Bio
 - Filtre des variants fonctionnels
 - Conservation des variants rares
 - Variants identifiés par au moins deux algorithmes chez les individus atteints de la famille
 - IV. Contrôle *in silico* par visualisation des alignements avec l'outil IGV
 - 4. Analyse de ségrégation des variants identifiés
 - 5. Etude de liaison par «IBD » (Identité par descendance)

Résultats

- a. Résultats du séquençage ciblé
 - 1. Résultats de séquençage des gènes connus
 - 2. Analyse de ségrégation des variants identifiés
- b. Résultats de recherche clinique
 - 1. Résultats de séquençage haut débit type exome
 - 2. Analyse de ségrégation des variants
 - I. Analyse des nouveaux variants
 - II. Analyse des variants connus
 - 3. Résultats de l'analyse de liaison par IBD

Discussion

- a. Variants identifiés par le séquençage ciblé
 - 1. Le variant *CDKN1B* : c.326T>G, p.val109Gly
 - 2. Le variant *AIP* c.469-2A>G
 - 3. Le variant *MEN1* exon 9 c.1254C>T, p.Asp418Asp

- b- Variants identifiés par le séquençage type exome
 - 1. Les variants répertoriés dans les bases de données
 - I. Le variant *PRKCH* chr14:61857988.A->T p.Thr137Ser
 - II. Le variant *MYO3A* chr10:26243804.A->C p.Asp57Ala
 - 2. Le nouveau variant *EPS8* chr12 :15784564.TC->CG p.Asp619Arg
- c. Limites de la méthode utilisée
- d. Limites du séquençage d'exome
- e. Limites de l'analyse de ségrégation
- f. Conclusion et perspectives

Bibliographie

Introduction

L'hypophyse, ou glande pituitaire, est une glande endocrine qui se trouve dans une cavité osseuse au sein de l'os sphénoïde : la selle turcique. Elle est reliée à l'hypothalamus par la tige pituitaire. Elle est régulée par des neurohormones émises par l'hypothalamus, et produit des hormones qui stimulent les autres glandes endocrines de l'organisme. L'hypophyse peut être à l'origine de multiples pathologies comme le nanisme hypophysaire, la puberté précoce ou encore le diabète insipide ; elle est également le siège de tumeurs, avec ou sans syndrome d'hypersécrétion hormonale.

L'adénome hypophysaire est une tumeur bénigne développée aux dépens de l'antéhypophyse. Il est monoclonal: une mutation cellulaire unique est suivie d'une expansion clonale. L'initiation se produit grâce à l'activation d'un oncogène ou à l'inactivation d'un gène suppresseur de tumeur (1). La bénignité n'épargne pas de possibles caractéristiques d'agressivité : hormonales, radiologiques ou histologiques. Le taux de récurrence toute forme comprise est estimé à 10-35% (1).

En fonction de la taille de la tumeur, on décrit quatre types : les picoadénomes, les microadénomes, les macroadénomes et les adénomes géants. Les picoadénomes mesurent moins de 3 mm de diamètre. Les microadénomes mesurent moins de 10 mm alors que les macroadénomes mesurent plus de 10mm. Un macroadénome peut rester « enclavé » dans la selle turcique (2/3 des cas) ou être « invasif » (1/3 des cas). Il est dit invasif lorsqu'il a traversé les parois de la selle turcique, perforé la dure-mère, le diaphragme sellaire et parfois le corps du sphénoïde. Enfin, les adénomes géants mesurent plus de 40 mm. Ils envahissent souvent la base du crâne et la fosse cérébrale antérieure, moyenne ou postérieure.

a. Épidémiologie

Les adénomes représentent 10% des tumeurs intra-crâniennes, mais seulement 0,02 à 0,05% d'entre eux sont symptomatiques. La prévalence est de 1/1064 personnes, mais les autopsies révèlent qu'ils seraient en fait beaucoup plus fréquents : 14% de la population (2). Leur découverte se fait habituellement entre 35 et 65 ans, avec une légère prédominance féminine (53%), sauf pour l'adénome somatotrope. Avant l'âge de 20 ans, les adénomes sont rares (85% de prolactinomes et 10% des adénomes corticotropes) (1,3,4).

Les adénomes peuvent être sécrétants (parfois pluri-sécrétants) ou non sécrétants (5). Aujourd'hui, cette terminologie est remplacée par "fonctionnels" ou "non fonctionnels". On distingue les adénomes à prolactine, les adénomes gonadotropes, somatotrope et corticotrope et plus rarement thyrotrope (**Figure 1**) (6). Les adénomes à prolactine représentent 40-60% des adénomes hypophysaires. Les adénomes gonadotropes et adénomes non fonctionnels représentent 30% des adénomes hypophysaires. Les macroadénomes sont souvent volumineux au moment du diagnostic avec en général une importante expansion supra-sellaire et une compression chiasmatique responsable d'un syndrome tumoral. Les adénomes somatotrope (8-16%) sont dans 80% des cas des macroadénomes et envahissent fréquemment le sinus caverneux. Ils peuvent involuer spontanément par infarctissement ou nécrose. Les adénomes corticotrope représentent 5-15% des adénomes hypophysaires. Les autres types d'adénomes, les thyrotrope (qui ne présentent pas de caractéristiques particulières), représentent les 2% restant.

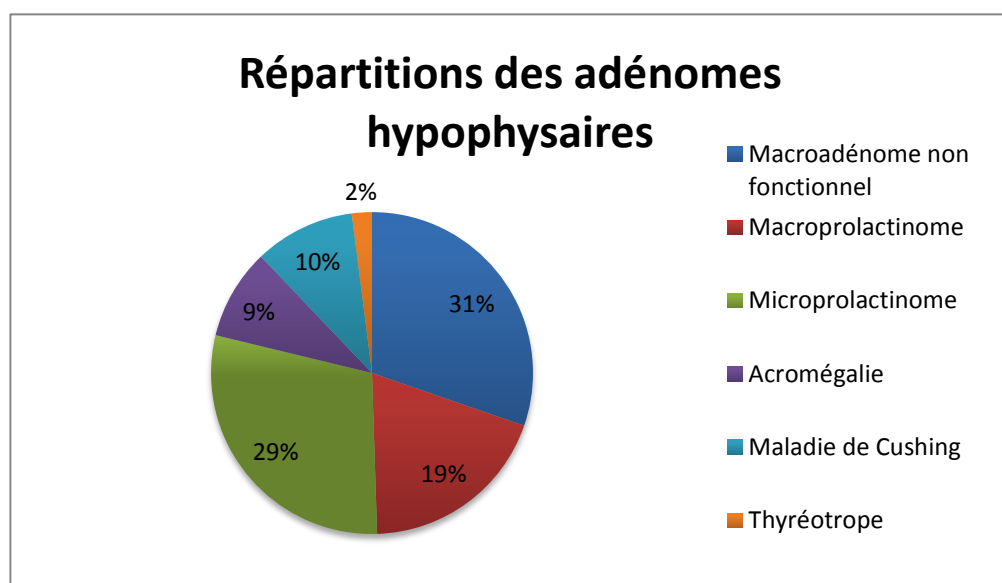


Figure 1 - Répartition des adénomes selon leur sécrétion hormonale.

Par ailleurs, on distingue les adénomes sporadiques des adénomes familiaux : soit dans le cadre de syndromes tumoraux avec anomalies génétiques dont les Néoplasies Endocriniennes Multiples (NEM) de type 1 et 4, le complexe de Carney et le syndrome de McCune Albright et soit isolés dans les formes familiales - Familial Isolated Pituitary Adenoma (FIPA) (7). L'histoire familiale est présente dans 5% des cas d'adénomes hypophysaires (8). La NEM1, décrite en 1954 par Wermer, associe classiquement un adénome hypophysaire, une hyperplasie parathyroïdienne et une tumeur gastro-pancréatique. Les adénomes hypophysaires surviennent dans le cadre d'une NEM1 dans environ 30% des cas. Les mutations du gène codant la ménine dans la NEM1 comptent pour plus de 80 % des cas. Quelques cas de NEM4 sont décrits porteurs de mutations dans le gène *CDKN1B* codant l'Inhibiteur 1B de Kinase Cycline-Dépendante. Des mutations dans le gène *PRKAR1A* qui code la sous-unité régulatrice de la protéine kinase cAMP- dépendante (PKA) sont décrites dans 50% des cas de CNC. Les adénomes hypophysaires familiaux (FIPA) peuvent également impliquer le gène *AIP* (aryl hydrocarbon receptor interacting protein). Les mutations dans ce gène comptent pour 2% à 3 % de tous les adénomes hypophysaires confondus (sporadiques et familiaux) (9)(10). Des mutations dans le gène *AIP* sont retrouvées dans 15 % des cas d'adénomes hypophysaires isolés tout phénotype confondu et dans 50 % des adénomes hypophysaires isolés de type somatotrope (IFS) (11). (**Tableau 1**).

Syndrome	Gènes	Atteintes endocriniennes classiques
Néoplasie endocrinienne de type 1 (NEM 1)	<i>MEN1</i> (>80%)	- Hyperparathyroïdie primaire - Tumeurs neuroendocrines pancréatiques - Adénome hypophysaire
Complexe de Carney (CNC)	<i>PRKAR1A</i> (>70%)	- Surrénalienne - Tumeurs gonadiques - Adénome hypophysaire (acromégalie)
Néoplasie endocrinienne de type 4 (NEM 4)	<i>CDKN1B</i> (>90%)	- Hyperparathyroïdie primaire - Tumeurs neuroendocrines pancréatiques - Adénome hypophysaire
Syndrome de McCune-Albright (SMA)	<i>GNAS</i> (>90%) (mutation somatique)	- Puberté précoce - Adénome hypophysaire (acro-gigantisme)
FIPA	<i>AIP</i> (15%)	- Adénome hypophysaire

Tableau 1 : Les causes génétiques des adénomes hypophysaires chez l'adulte

b. Etiologie des adénomes hypophysaires

1. La Néoplasie Endocrinienne Multiple de type 1 (NEM1)

La NEM1 est un syndrome de prédisposition héréditaire aux tumeurs dont la présentation clinique objective principalement des lésions tumorales de 3 glandes endocrines: parathyroïdes (95% des cas), pancréas (30-80% des cas) et hypophyse antérieure (15-90% cas) (12,13). Il s'agit d'une maladie génétique rare qui touche environ un individu sur 30 000 à un sur 500 000 (14). Elle est transmise sur le mode autosomique dominant, à pénétrance quasi complète (95%) (15). La présentation clinique revêt deux formes : une forme familiale (90%) et une forme sporadique (10%). La forme familiale typique montre chez au moins deux sujets apparentés au premier degré l'existence d'au moins une des lésions cardinales de la NEM1. La forme sporadique est caractérisée par l'existence chez un sujet d'au moins deux des lésions cardinales de la NEM1 (16). En dehors des 3 atteintes cardinales citées précédemment, il existe d'autres lésions moins fréquentes, endocrines ou non endocrines, pouvant compléter le tableau clinique : il s'agit de tumeurs cortico-surréaliennes, de tumeurs carcinoïdes des bronches, du tube digestif ou du thymus, de tumeurs cutanées (lipomes, angiofibromes, collagénomes), et de méningiomes. Une vingtaine d'associations lésionnelles différentes a été décrite, mais la lésion la plus fréquente et la plus constante est celle des parathyroïdes (plus de 95% des sujets sont atteints dans la cinquième décade), classiquement de type hyperplasie multiglandulaire (15).

Le gène *MEN1* responsable de ce syndrome est localisé en 11q13 (17,18); il s'agit d'un gène suppresseur de tumeurs, qui code une protéine de 610 acides aminés, principalement nucléaire, la ménine. Les fonctions de cette protéine sont encore très mal connues, mais elles concernent la régulation transcriptionnelle (différenciation, prolifération, cycle cellulaire, apoptose), la stabilité du génome (réplication, réparation), et s'exercent dans de très nombreux types cellulaires (expression ubiquitaire) (19,20). Il a une taille de 10kb et comporte 10 exons, dont 9 codants. A ce jour plus de 400 mutations différentes ont été décrites, et sont réparties dans toute la partie codante, sans véritable hot spot. Environ 40% de patients atteints d'une NEM1 ont un adénome hypophysaire (21,22). La répartition des phénotypes des adénomes hypophysaires est celle retrouvée dans les cas sporadiques. Il

faut cependant souligner que dans le cas des prolactinomes associé à une NEM1, il s'agit en grande majorité de macroprolactinomes (84% des prolactinomes sont des macroadénomes). Ils sont invasifs et résistants au traitement dopaminergique dans 60% des cas (21). Approximativement 10% des patients atteints de NEM1 ne sont pas porteurs de mutations dans le gène de la menine. D'autres gènes peuvent être impliqués mais à ce jour, aucun n'a été identifié (23).

2. La Néoplasie Endocrinienne Multiple de type 4 (NEM 4)

La NEM4 est une entité syndromique décrite récemment par Pellegata et al. (2012). Ce syndrome a été extrapolé d'un modèle animal : NEMX. En effet, l'équipe de Pellegata a remarqué qu'un modèle de rat de laboratoire présentait une association d'atteintes endocriniennes appartenant à la NEM1 mais également à la néoplasie endocrinienne de type 2 (phéochromocytome, cancer médullaire de la thyroïde...)(24). L'analyse génétique des rats atteints de NEMX a révélé des mutations homozygotes dans le gène *CDKN1B*. Il code pour une protéine de 196 acides amines de type inhibiteur de kinase cycline-dépendante (CK1). Ce gène a été analysé chez les patients ayant un syndrome NEM1 mais ne présentant pas de mutation du gène de la ménine. Trois pour cent des patients atteints du NEM1-like syndrome associant hyperparathyroïdie primaire, adénome hypophysaire, tumeurs neuroendocrines pancréatiques mais aussi des tumeurs surrénaliennes, gonadiques et thyroïdiennes sont porteurs de mutations dans *CDKN1B*. A ce jour, seulement huit mutations hétérozygotes ont été décrites dans la littérature comme étant responsables de NEM4 (24).

3. Le Complexe de Carney (CNC)

Le Complexe de Carney (CNC) est caractérisé par une pigmentation tachetée de la peau (lentigines et naevi bleus), une suractivité endocrinienne qui inclut : un syndrome de Cushing indépendant de l'ACTH dû à une dysplasie surrénalienne pigmentaire micronodulaire primaire (PPNAD), un adénome hypophysaire responsable le plus souvent d'une acromégalie, des tumeurs de la thyroïde et des testicules et des myxomes (coeur,

peau, sein). Il s'agit d'une maladie rare dont la prévalence est inconnue mais 750 cas ont été indentifiés jusqu'à alors (25,26). *PRKAR1A* est l'un des gènes responsable de CNC, il est situé en 17q22-24 et code pour la sous-unité régulatrice (R1A) de la protéine kinase A (27). Cette protéine est une protéine clé de la voie de signalisation de l'AMPc. Des mutations inactivatrices du gène *PRKAR1A* ont été rapportées dans environ 45 à 70% des cas index de CNC. Dans plus de 30 % des cas, la cause génétique reste inconnue. L'adénome hypophysaire n'est présent que dans 12% des CNC (25).

4. Le syndrome de McCune-Albright (SMA)

Le SMA est classiquement défini par une triade clinique comportant une dysplasie fibreuse des os (DFO), des taches cutanées « café au lait » et une puberté précoce (PP). La prévalence est estimée entre 1/100 000 et 1/1 000 000. La DFO peut impliquer un ou plusieurs sites squelettiques et elle se manifeste par une boiterie et/ou une douleur et, parfois, par une fracture pathologique. La scoliose est fréquente et peut être évolutive. En plus de la puberté précoce (saignements vaginaux ou développement précoce du tissu mammaire chez les filles, élargissement des testicules et du pénis et comportement sexuel précoce chez les garçons), d'autres endocrinopathies responsables de syndrome d'hypersécrétion peuvent être présentes comme une hyperthyroïdie, un adénome hypophysaire responsable d'une acromégalie, un syndrome de Cushing, etc ... Les tâches « café au lait » apparaissent en général durant la période néonatale, mais c'est le plus souvent la puberté précoce ou la DFO qui évoque le diagnostic. Une atteinte rénale est observée chez environ 50 % des patients. Le SMA est dû à des mutations somatiques du gène *GNAS* codant la protéine régulatrice de l'AMPc, la protéine Gs alpha. L'étendue de l'atteinte du SMA dépend de la prolifération, de la migration et de la survie des cellules dans lesquelles la mutation a eu lieu spontanément durant le développement embryonnaire.

5. Les adénomes hypophysaires familiaux isolés (FIPA)

Le syndrome FIPA est caractérisé par la présence au sein d'une même famille d'au moins 2 adénomes hypophysaires isolés sans autres types de tumeurs endocrines associées. Ces tumeurs hypophysaires peuvent être hétérogènes associant adénomes à GH, à PRL, à ACTH ou être non fonctionnelles (NFPA). Les familles d'adénomes somatotropes appartiennent au sous-groupe de FIS (Familial Isolated somatotropinomas). La fréquence des différents types d'adénomes dans FIPA sont : les prolactinomes, 41%; les adénomes somatotropes, 37%; les NFPA, 17%, les adénomes corticotropes, 4% et les thyréotropes, 1% (11) (Figure 2).

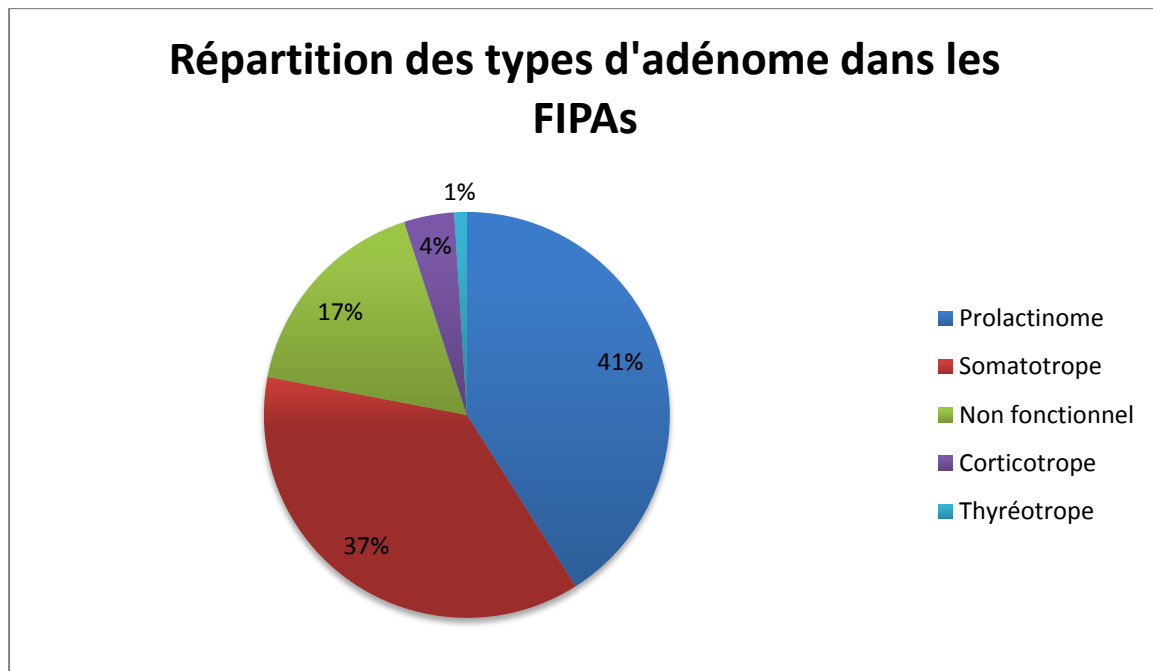


Figure 2 : Répartition des types d'adénome dans les FIPAs (11)

A ce jour, *AIP* est le seul gène connu impliqué dans les FIPAs. Il code la protéine AIP. Il est situé sur le chromosome 11q13. La majorité des mutations rapportées dans le gène *AIP* implique une perte du domaine protéique C-terminal contenant le « Tetratricopeptide Repeat Domain » essentiel pour les interactions avec le récepteur des aryl-hydrocarbones et de la protéine Hsp90 (heatshock protein).

Les mutations du gène *AIP* sont retrouvées dans environ 15% des FIPAs et 50% des IFSS. Les adénomes hypophysaires *AIP* mutés se caractérisent par un âge de survenue plus

précoce que celui observé dans les adénomes sporadiques ou dans les FIPAs sans mutation dans le gène *AIP* (15 à 20% des adénomes hypophysaires mutés *AIP* sont diagnostiqués avant 30 ans, l'âge de survenu est de 12 ans inférieur à celui retrouvé dans les FIPAs). Il s'agit souvent de macroadénomes agressifs, résistants aux traitements. La pénétrance de la maladie est encore mal connue, elle est située entre 33 et 60% (11). La cause génétique demeure inconnue dans près de 80% des cas d'adénomes hypophysaires familiaux.

Ce travail consiste à s'intéresser à une famille, suivie au Centre Hospitalo-Universitaire de Nantes, atteinte d'adénomes hypophysaires sur trois générations afin de mieux les caractériser sur le plan clinique et génétique. Nous avons élaboré une stratégie translationnelle entre la pratique clinique et la recherche scientifique pour trouver une cause génétique à l'origine de cette pathologie.

Observation clinique

Madame G. âgée de 61 ans consulte en ophtalmologie pour la première fois en 2009 pour des troubles visuels: des difficultés à l'accommodation associées à des céphalées. L'examen ophtalmologique met en évidence un ptosis droit, une paralysie oculomotrice complète avec atteinte des nerfs III et IV droits, une diplopie avec baisse de l'acuité. Au fond d'œil, il existe une atrophie péri-papillaire gauche avec excavation et œdème papillaire droit (syndrome de Foster-Kennedy). Le scanner cérébral puis l'IRM hypophysaire réalisée en urgence trouve un macroadénome hypophysaire de 28*20*20 mm avec envahissement du sinus caverneux droit, traversant le plancher de la selle turcique ainsi qu'une protrusion marquée dans le sphénoïde. Le bilan hormonal révèle une hyperprolactinémie à 4899 ng/mL (N < 25), une insuffisance thyroïdienne avec une TSH à 0,01 mUI/L, et T4 à 11,5 pmol/L (patiente déjà sous traitement substitutif par L-Thyroxine suite à une thyroïdectomie pour un goitre multi-hétéronodulaire), une insuffisance gonadotrope (FSH et LH basses chez une femme ménopausée), un cortisol à 8h normal à 188 ng/mL et un taux normal d'IGF1 à 219 ng/mL (N < 256). La patiente est prise en charge initialement au CHU de Nancy pour résection transsphenoidale (résection incomplète compte tenu de l'envahissement du sinus caverneux droit). L'analyse anatomopathologique est en faveur d'un adénome hypophysaire à prolactine avec Ki67 supérieur à 3%. La patiente est traitée médicalement par un agoniste dopaminergique : la cabergoline. Le bilan initial ne retrouve pas d'autres atteintes endocriniennes avec notamment un bilan phosphocalcique normal. La patiente a déménagé à Nantes pour rapprochement familial et elle est actuellement suivie dans le service d'Endocrinologie du CHU de Nantes.

Le conseil familial trouve un antécédent d'adénome hypophysaire chez son père âgé de 93 ans et ses 2 filles ainsi que chez sa cousine germaine (**Figure 3 : arbre généalogique**).

Son père, Monsieur B. (LD0014), âgé de 93 ans, est suivi au CHU de Nantes pour adénome hypophysaire géant mesurant 41*38*42 mm au dernier TDM réalisé en février 2016. Sur le plan hormonal, il existe une insuffisance gonadotrope et somatotrope non substituée compte-tenu de l'âge du patient. Par ailleurs, il présente une hyperparathyroïdie primaire traitée médicalement.

La fille ainée, Madame L. (LD0013) âgée de 47 ans, est également suivie au CHU de Nantes pour un adénome hypophysaire de 9 mm, d'allure non fonctionnel. Il a été diagnostiqué à l'âge de 44ans. Son bilan phosphocalcique est normal.

La seconde fille, Madame V. (LD0003) présente un microadénome de 3 mm de diagnostic récent à l'âge de 44 ans sans syndrome d'hypersécrétion ou d'insuffisance hormonale. Son bilan phosphocalcique est normal.

La cousine germaine (LD0009) présente un microadénome à prolactine de 3 mm à l'âge de 43 ans. Il a été diagnostiqué sur galactorrhée persistance et l'absence de retour de couches à plus de 6 mois post-partum et sans notion d'allaitement. Elle est également suivie au CHU de Nantes dans le service de gastro-entérologie pour une tumeur neuroendocrine bien différenciée de la valvule iléo-caecale. Son bilan phosphocalcique est normal.

La tante, Madame M âgée de 86 ans, n'a pas souhaité effectuer les explorations radiologiques. Le bilan biologique hypophysaire est normal. Pas d'anomalie du bilan phosphocalcique.

D'autres membres de la famille ont été explorés par imagerie hypophysaire (IRM) et par un bilan hypophysaire et phosphocalcique complet. Aucune anomalie n'a été retrouvée. Les différentes atteintes des individus explorés sont résumées dans le **Tableau 2**.

Arbre généalogique de Madame G.

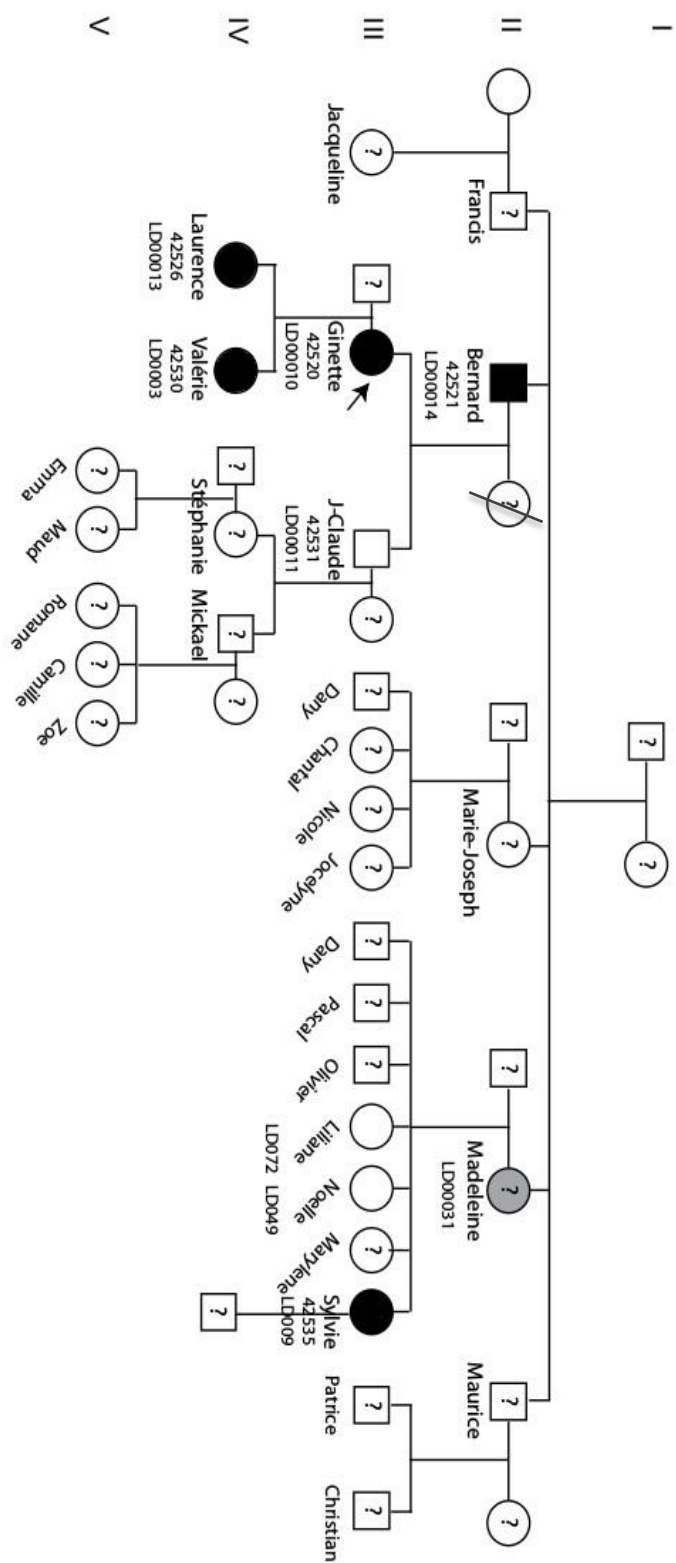


Figure 3 : Arbre généalogique de la famille

Arbre généalogique : le carré désigne un individu de sexe masculin, le rond désigne un individu de sexe féminin, la flèche noire indique le cas index. Les patients affectés sont représentés en noir. En gris, la patiente ayant refusé l'imagerie hypophysaire. Les apparentés désignés par '?' ont un phénotype inconnu

Patient	Age	Phénotype	Profil hormonal	IRM	Autres atteintes
LD0003	44	Microadénome	Aucun	IRM 2016: microadénome de 3 mm	
LD0009	49	Microadénome	Hyperprolactinémie	IRM 2016: microadénome de 3 mm du plancher médian gauche	TNE bien différenciée de la valvule iléo-caecale en 2010 à 43 ans.
LD0010 Cas index	69	Macroadénome à prolactine	Hyperprolactine Déficit gonadotrope	Macroprolactinome de 28mm diagnostiqué à l'âge de 62 ans	
LD0011	66	Non atteint	Aucun	IRM 2016 pas d'anomalie	
LD0013	47	Microadénome	Aucun	IRM 2016 : adénome de 9 mm diagnostiqué à l'âge de 44ans	
LD0014	93	Adénome hypophysaire géant	Déficit gonadotrope et somatotrope	TDM 2016 : macroadenome de 41*38*42 mm	Hyperparathyroïdie primaire
LD0031	86	Porteuse ?	Aucun	Pas d'IRM (refus de la patiente)	
LD0049	59	normal	Aucun	IRM normale 2015	
LD0072	64	normal	Aucun	IRM normale 2015	

Tableau 2 : Description phénotypique des membres de la famille de Madame G.

Compte tenu de la complexité de cette famille, nous avons entrepris dans un premier temps, une analyse génétique par séquençage ciblée des gènes connus comme cela est fait en pratique clinique quotidienne. Dans un second temps, nous avons élaboré une stratégie de recherche avec un séquençage de nouvelle génération de type exome afin d'identifier de nouveaux gènes impliqués dans les adénomes hypophysaires.

Méthodes et stratégies

a. Stratégie de diagnostic clinique

La stratégie initiale a consisté à envoyer des échantillons de sang total (2 tubes EDTA) des patients atteints d'adénome hypophysaire au centre diagnostic des maladies hypophysaires au CHRU de Lille pour la recherche de mutations dans les gènes connus (*MEN1*, *CDKN1B* et *AIP*). Les analyses ont été réalisées par séquençage nouvelle génération (NGS) par la technique Haloplex. Tous les patients ont signé un consentement adapté.

b. Stratégie de recherche clinique

1. Prélèvement sanguin, consentement et extraction de l'ADN

Des échantillons de sang total (2 tubes EDTA) ont été prélevés pour chaque membre de la famille ayant souhaité participer à ce travail. L'extraction d'ADN est faite sur machine automatisée à l'institut du thorax (IRT). Tous les patients ont signé un consentement de recherche clinique de l'institut du thorax.

2. Extraction de l'ARN

Le prélèvement de sang frais total est réalisé sur tube PAXGENE, tube spécifique contenant des inhibiteurs de ARNase afin de préserver au mieux la qualité des ARN du patient. L'extraction de l'ARN a été réalisée dans les 48 heures suivant le prélèvement avec le Kit Macherey Nagel.

3. Séquençage de type EXOME

Dans un second temps, nous avons entamé une démarche de séquençage à haut débit de type exome chez 2 patients (LD0010 et LD0009) appartenant à cette famille afin d'essayer d'identifier une nouvelle cause génétique des FIPAs.

I. Séquençage à haut débit – illumina HiSeq 2500

Le séquençage d'exome consiste à séquencer toutes les régions codantes du génome humain après une étape de capture, de purification puis d'amplification. Pour la capture des

parties codantes, nous avons utilisé le kit SureSelect Human All Exon V5 commercialisé par Agilent Technologies Santa Clara, CA. Il permet de capturer 45 Mb de séquence comprenant 21 522 gènes. Le séquençage est effectué sur l'appareil HiSeq 2500 (Illumina, San Diego, CA, USA) disponible à la plateforme de séquençage de Nantes

II. Alignement et annotations des séquences

Alignement

L'outil BWA-MEM permet d'aligner chaque "read" sur le génome humain de référence (hg19/HGRC37) et va générer des fichiers SAM qui colligent de nombreuses informations relatives à tous les "reads" appariés (identifiant, séquence, scores de qualité, position sur le génome etc...). Ces fichiers SAM sont volumineux et vont donc être compressés en fichier BAM (fichiers binaires). Les "reads" sont ensuite triés et hiérarchisés sur le génome. Les "reads" issus d'un même fragment d'ADN sont ensuite annotés avec Picard/MarkDuplicates. Les "reads" sont ensuite réalignés autour des insertions / délétions spécifiquement dans les régions d'intérêt. Une re-calibration des "reads" BAM est ensuite effectuée grâce à GATK. Cela permet d'affiner les scores qualité.

Détection de variants (« calling »)

La détection des variants a été réalisée grâce à 3 algorithmes : GATK Unified Genotyper (version 2.8) (28), Samtools mpileup (Version 0.1.19) (29) et GATK HapCaller. Le calling permet la genèse de fichiers VCF (« Variant Call Format ») (30) Les fichiers VCF sont des fichiers tabulaires listant l'ensemble des variations détectées par rapport à la séquence de référence du génome humain pour chaque patient séquencé et pour chaque algorithme. Chaque ligne correspond à un variant et chaque variant est annoté avec de nombreuses informations relatives. On retrouve, la position génomique du variant, son identifiant « rs » (si ce variant est référencé dans la base de données), l'allèle de référence, l'allèle alternatif observé, la qualité de « calling » attribuée à ce variant, l'état du variant (hétéro/homozygote) et une grande quantité d'informations relatives à celui-ci (score de conservation, fréquence allélique etc.), toutes ces informations étant regroupées dans une même colonne.

Annotation des variants

Les variants sont ensuite annotés en utilisant l'algorithme VEP (« Variant Effect Predictor »). Cette annotation, basée sur la base de données Ensembl (31), permet de prédire l'effet du variant nucléotidique sur l'ensemble des transcrits référencés. L'utilisation de bases de données publiques, contenant des données de séquençage de nombreux individus contrôles et de patients atteints de pathologies diverses de différentes populations, va permettre d'estimer la fréquence des variants recherchés.

Les bases de données utilisées sont :

- la base de données « 1000 Genomes » (phase 1- 379 individus d'origine européenne)
- la base «Exome Variant Server » (NHLBI GO Exome Sequencing Project (ESP) - 4300 individus d'origine européenne)
- la base de données ExAC «Exome Aggregation Consortium» - 36 923 individus européens non finlandais.

De plus, un score de conservation inter-espèces est attribué à chaque variant nucléotidique. Le score de conservation que nous avons utilisé est le score GERP (« Genomic Evolutionary Rate Profiling ») compris entre -12.3 et 6.17 à partir de comparaisons basées sur les génomes de 29 espèces de mammifères (32,33).

III. Filtrage des variants identifiés à partir de l'outil Knime4Bio

Suite aux étapes successives de séquençage, de traitement des données de séquençage, d'alignement et d'annotation, l'analyse génétique des fichiers VCFs est réalisée grâce à l'outil Knime4bio développé au laboratoire (34). Cet outil permet de contourner la difficulté du code informatique pour analyser ces fichiers contenant de nombreuses informations ne pouvant être gérées par des tableurs classiques. L'ajout de commandes est visualisé à l'aide d'une interface et permet d'appliquer successivement des actions de formatage et de filtrage présentées dans la Figure 8. L'enchaînement de commandes est appelé « workflow ».

Filtre des variants fonctionnels

Dans un premier temps, les VCF générés par Samtools, GATKHapCaller et GATK sont traités et introduits dans le workflow séparément. L'étape suivante consiste à conserver les variants ayant un effet fonctionnel potentiel sur la protéine sur la base des transcrits ENSEMBL et des annotations VEP (35) (**Figure 4-1**). Nous gardons ainsi les variants ayant des annotations non-synonymes et/ou impliqués dans les sites d'épissage des catégories suivantes (<http://www.sequenceontology.org/>) : (splice_donor_variant (SO:0001575), splice_acceptor_variant (SO:0001574), stop_gained (SO:0001587), frameshift_variant (SO:0001589), stop_lost (SO:0001578), initiator_codon_variant (SO:0001582), inframe_insertion (SO:0001821), inframe_deletion (SO:0001822), missense_variant (SO:0001583), transcript_amplification (SO:0001889), splice_region_variant (SO:0001630), transcript_ablation (SO:0001893), incomplete_terminal_codon_variant (SO:0001626), coding_sequence_variant (SO:0001580).

Conservation des variants rares

Les variants sont considérés rares de manière empirique si la fréquence de l'allèle mineur (MAF) est inférieure à 0,1%. Ainsi, nous utilisons successivement les bases de données externes (**Figure 4-2**):

- "1000 Genomes" phase 1 (379 individus européens (<http://www.1000genomes.org/>)).
- "Exome Variant Server" (EVS) (<http://evs.gs.washington.edu/EVS/>) (version ESP95000SI-V2) (4300 exomes d'individus d'origine européenne).

Ensuite nous utilisons un set de 387 exomes issus du projet UK10K (<http://www.uk10k.org/>), un set de 69 génomes de la base de données « Complete Genomics » (<http://www.completegenomics.com/>) puis la base de données « Genome of the Netherlands » (<http://www.nlgenome.nl/>) (499 génomes d'individus d'origine hollandaise) (**Figure 4-3**). Le seuil de fréquence choisi pour ces bases est de 5% car elles contiennent des individus apparentés.

Nous utilisons ensuite des bases de données «internes». Le but est de filtrer nos variants en excluant les erreurs de séquençage fréquentes (faux positifs récurrents). Ainsi cette base de données contient 100 exomes d'individus Nantais séquencés sur la plateforme de séquençage de Barcelone (Centre for Genomic Regulation) et 160 exomes de d'individus

séquencés sur la plateforme génomique de Nantes (<http://www.pf-genomique.univ-nantes.fr/>) (**Figure 4-4**). Le seuil arbitraire de fréquence choisi pour ces deux bases de données est de 20% car elles contiennent des individus apparentés.

Enfin, nous utilisons la base de données de l'Exome Aggregation Consortium (ExAC) du Broad institute (<http://exac.broadinstitute.org>) qui constitue actuellement la plus grande base de données publiques. Cette méta-base regroupe les données de 17 bases de données ré-analysées. Ainsi, elle contient les données de séquençage d'exome de 60 706 individus non apparentés de différentes origines. Nous utilisons des seuils de 0,01% dans la sous-population européenne non finlandaise (33 300 individus) (**Figure 4-5**).

Variants identifiés par au moins 2 algorithmes chez les patients de la famille

Afin de limiter le nombre de faux positifs, la dernière étape consiste à ne conserver que les variants qui sont trouvés par au moins 2 sur les 3 algorithmes. Enfin, dans le cas d'analyses familiales, nous focalisons notre analyse sur les variants partagés par les individus atteints séquencés.

- 1) Chargement et lecture des fichiers VCF, extraction des informations de fréquence et de conservation des variants, impact fonctionnel décrit par Variant effect predictor)
- 2) Conservation des variants ayant des fréquences de l'allèle mineur (MAF) inférieures à 1% dans les bases de données européennes « 1000 génomes » et « Exome Variant Server ».
- 3) Conservation des variants ayant des MAF inférieures à 5% dans les bases de données Agilent, et Complete Genomics.
- 4) Conservation des variants ayant des MAF inférieures à 0.1% dans la base de données de « Exome aggregation consortium » (Européens non-Finlandais).
- 5) Conservation des variants détectés par au moins deux algorithmes de « calling » (gatk, gatkHapCaller et samtools) et partagés par les patients séquencés de la famille.

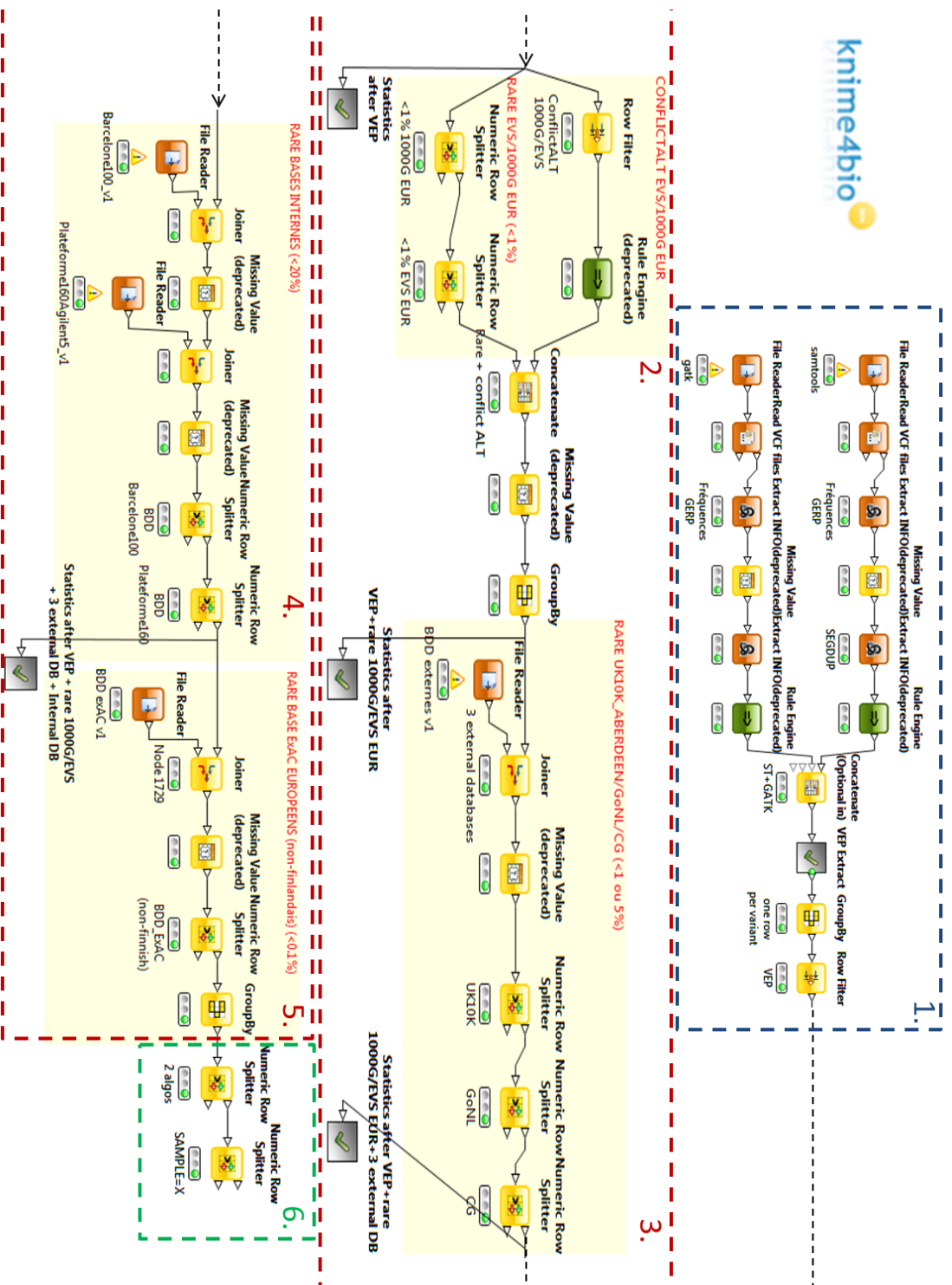


Figure 4: Visualisation du « workflow » d'analyse réalisée avec l'outil Knime4bio afin de filtrer les fichiers VCF

IV. Contrôle *in silico* par visualisation des alignements avec l'outil IGV

Suite aux analyses des données NGS, les variants d'intérêt identifiés ont été vérifiés par alignement des "reads" grâce à l'outil IGV « Integrative Genome Viewer » (Thorvaldsdóttir, Robinson, and Mesirov 2013). Cette étape constitue une vérification visuelle de la qualité et de l'alignement des reads à partir des fichiers BAM, avant le séquençage capillaire qui validera définitivement la présence du variant.

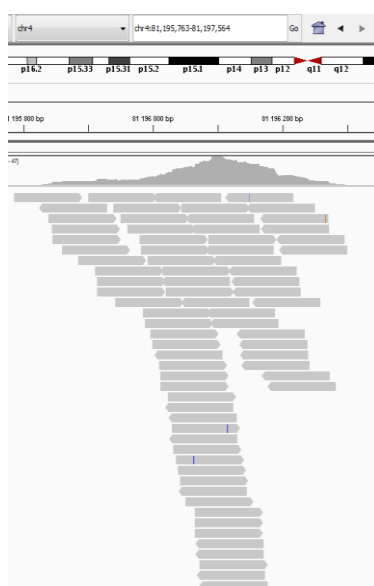


Figure 5: Exemple d'alignement des reads de séquençage à haut-débit grâce à l'outil IGV

4. Analyse de ségrégation des variants par la technique de séquençage SANGER

La validation des variants et les analyses de co-ségrégation ont été réalisées par séquençage capillaire (méthode de Sanger). Le choix des amorces de PCR et de séquençage a été réalisé grâce au logiciel Primer3web version 4.0.0 (<http://primer3.ut.ee/> - Broad Institute of MIT) en se basant sur les séquences génomiques obtenues sur l'outil « Genome Browser » Université de Californie, Santa Cruz (<http://genome.ucsc.edu/>). Le séquençage a été réalisé en utilisant le kit Big Dye Terminator v3.1, puis une migration électrophorétique sur séquenceur automatique 3730 DNA Analyzer 48-capillaires (Life technologies). Les séquences ont été analysées avec le logiciel SeqScape v2.5 (Applied Biosystems) par une inspection visuelle.

5. Etude de liaison par « IBD linkage »

La recherche de régions IBD permet d'identifier les régions génomiques partagées par les individus atteints d'une même famille grâce aux données de génotypage haut-débit sur génome entier. Le séquençage d'exome complet nous permet ensuite d'identifier les variations génétiques retrouvées dans les régions codantes du génome. Ces deux stratégies de recherche sont complémentaires.

Les analyses d'identité par descendance ont pour but d'identifier des blocs haplotypiques partagés par des paires d'individus apparentés hérités d'un ancêtre commun sans événement de recombinaison. Ces analyses ont été réalisées avec l'algorithme « Identity By Descent Linkage Disequilibrium » (Han and Abney 2011). Les probabilités d'IBD sont calculées pour chaque « état d'IBD ». En effet, trois états d'IBD sont définis : - l'état IBD 0 qui correspond à l'absence de partage d'un haplotype, - l'état IBD1 qui correspond au partage d'un haplotype par une paire d'individus et, - l'état IBD2 qui correspond au partage de deux haplotypes par une paire d'individus.

Dans le cadre de notre analyse familiale, nous recherchons des zones haplotypiques partagées entre individus d'une même famille. Ainsi, nous allons rechercher les régions génomiques ayant une très faible probabilité d'être IBD0 (ne partageant aucun haplotype) et donc une plus forte probabilité d'être IBD 1 ou 2. Chaque paire d'individus a été comparée en utilisant un seuil de probabilité IBD0 de 0.2. Ainsi, chaque marqueur génotypé (SNP), est comparé en prenant en considération le déséquilibre de liaison au sein du bloc haplotypique, et le nombre de paires d'individus partageant ces SNPs, est attribué. Les données sont représentées sous forme de graphique représentant en abscisse la position chromosomique et en ordonnée le nombre de paires d'individus partageant ce SNPs.

Ces analyses peuvent être réalisées pour l'identification de régions partagées par les individus atteints de la famille (paires concordantes) mais aussi en comparant les individus sains et atteints d'une même famille (paires discordantes). Si une région haplotypique est partagée par toutes les paires concordantes, mais aussi par les paires discordantes, la région ne semblera pas intéressante vis-à-vis du trait phénotypique étudié (dans le cadre d'une pathologie à pénétrance complète).

L'ensemble de ces analyses a été réalisé au sein de l'équipe « Génétique des maladies héréditaires » par les biostatisticiens (Floriane Simonet et Christian Dina).

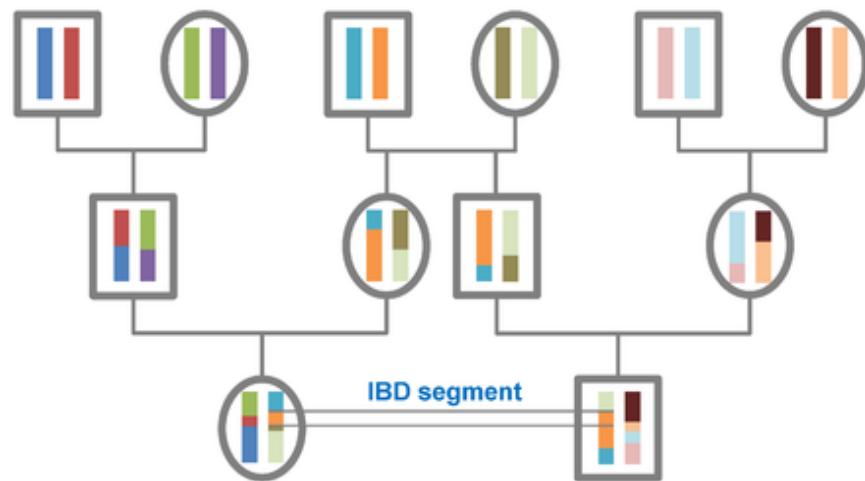


Figure 6 : Schéma des régions chromosomiques partagées au sein d'une même famille.

Résultats

a. Résultat du séquençage ciblé

1. Résultats de séquençage des gènes connus

Les analyses génétiques ont été effectuées dans le laboratoire de biologie moléculaire au CHU de Lille par l'Equipe du Professeur Nicole PORCHET.

Pour la patiente Madame G. (LD0010), le cas index, deux mutations ont été identifiées dans les gènes connus *AIP* et *CDKN1B*

- *AIP* : intron 3 site d'épissage c.469-2A>G
- *CDKN1B* : exon 1 c.326T>G, p.val109Gly

Pour le patient LD0014 (père du cas index), une mutation silencieuse a été trouvée dans *MEN1* exon 9 c.1254C>T, p.Asp418Asp.

Aucune mutation dans les gènes *MEN1*, *CDKN1B* et *AIP* n'a été retrouvée chez les autres patients de la famille (LD0009, LD0003, LD0013) présentant aussi un adénome hypophysaire.

2. Analyse de ségrégation des variants identifiés

L'analyse de ségrégation des variants *MEN1*, *AIP* et *CDKN1B* a été réalisée chez tous les individus de la famille au laboratoire de recherche médicale au sein de l'IRT. Les membres atteints et non atteints restants de la famille (7 membres) ne sont pas porteurs des variants identifiés par le séquençage ciblé.

b. Résultats de recherche clinique

1. Résultats du séquençage d'exome

Afin de déterminer l'origine génétique de FIPA dans cette famille, nous avons entrepris une démarche de séquençage d'exome. Nous avons recherché les variants rares fonctionnels chez 2 apparentés présentant un adénome hypophysaire (LD0010 et LD0009). La préparation des librairies et le séquençage à haut débit sont décrits dans la section « méthodes et stratégies ». Toutes les séquences codantes ont été séquencées pour les 2

patients, avec une profondeur de lecteur moyenne de 65 "reads". Sur l'ensemble des séquences obtenues, 97% des bases des régions ciblées sont couvertes au moins par dix « reads » et 91% sont couvertes par au moins 20 reads. Ces résultats sont cependant suffisants pour garantir la qualité des données issues du séquençage d'exome. L'analyse des données de séquençage permet de conserver les variants fonctionnels, très rares (Fréquence de l'allèle mineur inférieure à 0.01%). Les différentes étapes de l'analyse et des filtres sont décrites dans la section « méthodes et stratégies ».

A l'issue des étapes de filtrage, nous trouvons 21 variants commun aux deux patients dont 7 nouveaux (Novel), jamais décrits dans les bases de données et 14 déjà répertoriés et cela dans 20 gènes différents (**Tableau 3 et 4**).

Chromosome	Position génomique	Allèle de référence	Allèle alternatif	Gènes	Statut
2	43768343	C	T	<i>THADA</i>	Connu
2	179659646	T	C	<i>TTN</i>	Connu
2	189957140	G	A	<i>COL5A2</i>	Novel
2	209182628	A	G	<i>PIKFYVE</i>	Connu
3	36545905	G	A	<i>STAC</i>	Connu
3	47618639	C	G	<i>CSPG5</i>	Novel
4	8588786	C	T	<i>GPR78</i>	Connu
7	29923759	C	T	<i>WIPF3</i>	Novel
10	26243804	A	C	<i>MYO3A</i>	Connu
11	1017958	G	A	<i>MUC6</i>	Connu
12	15784564	T	C	<i>EPS8</i>	Novel
12	15784565	C	G	<i>EPS8</i>	Novel
14	61857988	A	T	<i>PRKCH</i>	Connu
16	11362741	T	C	<i>PRM3, RMI2, TNP2</i>	Connu
16	71660454	G	A	<i>MARVELD3</i>	Connu
16	75269619	C	T	<i>BCAR1</i>	Connu
17	6929789	C	A	<i>BCL6B</i>	Connu
17	40950135	C	T	<i>CNTD1, COA3, WNK4</i>	Connu
17	61611436	C	T	<i>KCNH6</i>	Novel
17	73870977	A	C	<i>TRIM47</i>	Connu
20	23858213	G	A	<i>CST5</i>	Novel

Tableau 3 : Liste des 21 variants identifiés et leur position génomique

Gènes avec un variant novel	Gènes avec un variant connus
<i>KCNH6</i> <i>EPS8</i> <i>COL5A2</i> <i>CST5</i> <i>CSPG5</i> <i>WIPF3</i>	<i>GPR78</i> <i>STAC</i> <i>PIKFYVE</i> <i>TTN</i> <i>THADA</i> <i>TRIM47</i> <i>COA3</i> <i>BCL6B</i> <i>BCAR1</i> <i>MARVELD3</i> <i>TNP2</i> <i>PRKCH</i> <i>MUC6</i> <i>MYO3A</i>

Tableau 4 : Listes des gènes classés selon que le variant qu'ils portent est nouveau ou déjà répertorié dans les bases de données

2. Analyse de ségrégation des variants identifiés

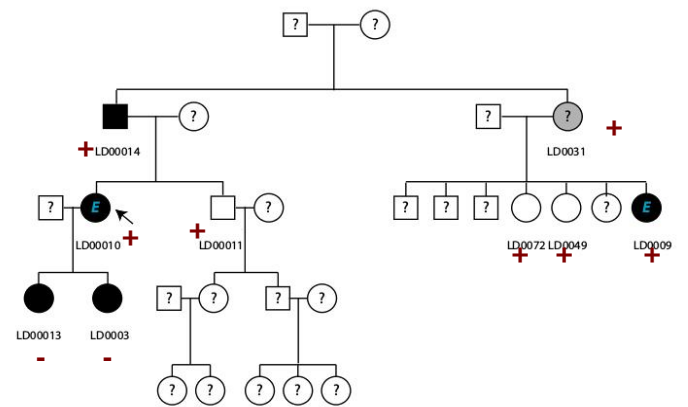
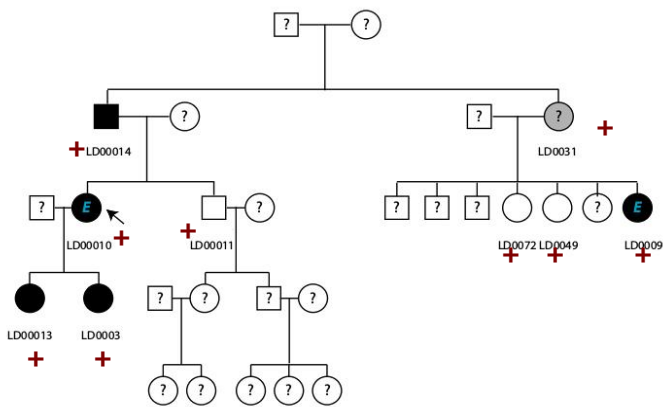
Nous avons entrepris une analyse de ségrégation des 21 variants identifiés sur 9 patients de la famille dont 6 atteints d'adénome hypophysaire et 3 non atteints.

I. Analyse des variants nouveaux

L'analyse de ségrégation des variants nouveaux a été réalisée après avoir choisi des amorces spécifiques pour chaque variant. Après les étapes de PCR, de purification puis de séquençage capillaire, le variant EPS8 est le seul variant qui co-ségrège avec la maladie dans la famille. Tous les individus atteints d'adénome hypophysaire sont porteurs du variant et tous les individus non atteints ne sont pas porteurs du variant testé. Les 5 autres gènes sont exclus car soit des individus atteints d'un adénome hypophysaire ne sont pas porteurs du variant testé soit des individus dont le phénotype est normal sont porteurs du variant (Figure 7).

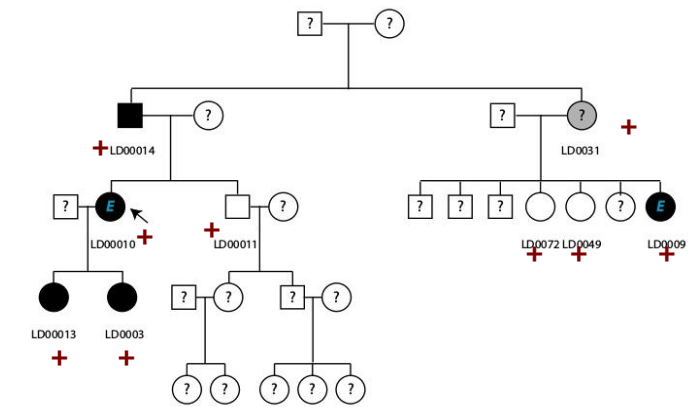
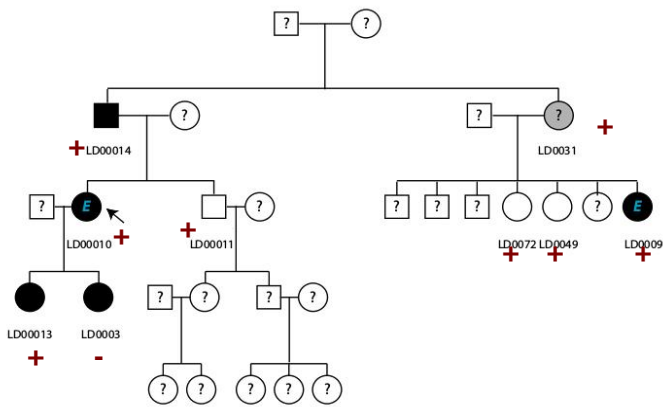
A- COL5A2 chr2:189957140.G->A p.R155C

B- KCNH6 chr17:61611436.C->T p.R289W



C- CSPG5 chr3:47618639 C->G p.D293H

D- CST5 chr20:23858213.G->A p.R92X



E- WIPF3 chr7:29923759.C->T p.P155S

F- EPS8 chr12:15784564.TC->CG p.D619T

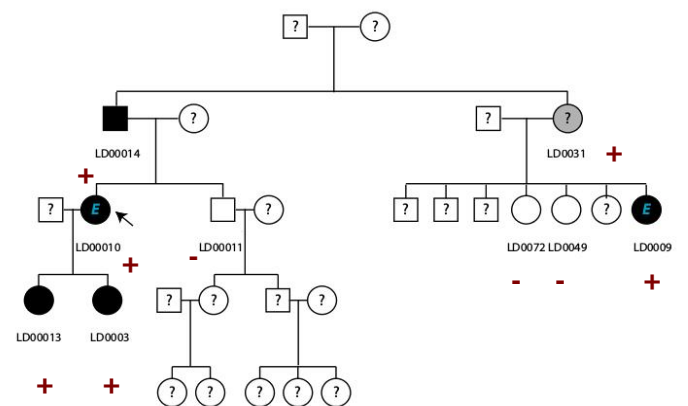
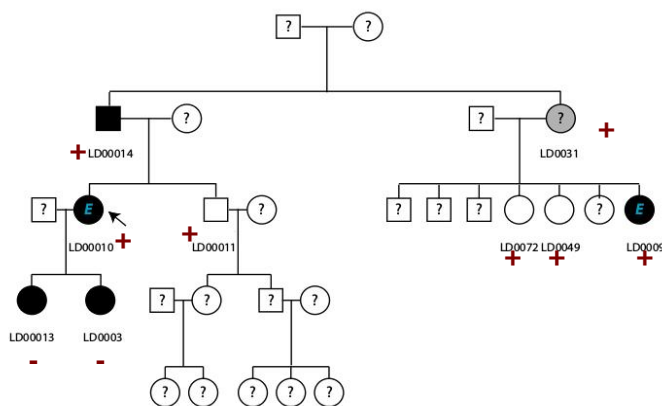


Figure 7 : Résultats de l'analyse de ségrégation pour les variants nouveaux représentés sur un arbre généalogique simplifié de la famille

A- COL5A2 p.R155C ; B- KCNH6 p.R289W ; C- CSPG5 p.D293H ; D- CST5 p.R92X ;
E-WIPF3 p.P155S ; F- EPS8 p.D619T

Arbre généalogique : le carré désigne un individu de sexe masculin, le rond désigne un individu de sexe féminin, la flèche noire indique le cas index. Les patients affectés sont représentés en noir. Les apparentés désignés par '?' ont un phénotype inconnu. Le « E » désigne les patients séquencés par exome

II. Analyse des variants connus

Deux variants sur 14 co-ségrègent avec la maladie. Ces variants sont situés dans les gènes *PRCKH* et *MYO3A*.

3. Résultats de l'analyse de liaison génétique par IBD

Onze régions chromosomiques de 30 Mb environ chacune sont communes aux 5 patients atteints de la famille, ces régions sont notées dans le **Tableau 5**. Ces régions sont bien couvertes par le séquençage d'exome. Cette analyse a été réalisée par l'outil IGV (analyse des fichiers BAM). L'analyse des données d'exome couplée aux résultats de l'IBD ainsi qu'à l'analyse de co-ségrégation permet de restreindre la liste des variants à 3 gènes (**Figure 8**).

Patients atteints	Régions chromosomiques communes		
	chr	start	end
LD00003			
LD00009	1	19210777	30347456
LD00010	2	40476402	66899308
LD00013	2	177924358	220025122
LD00014	4	141851355	175829769
	8	70513855	102081930
	9	36927558	72448508
	9	84204244	104012075
	10	2805026	26808573
	12	12922323	22952745
	14	49806075	77301219
	16	6731521	15881967

Tableau 5 : Position génomique des onze régions partagées par les 5 patients atteints de la famille de Madame G.

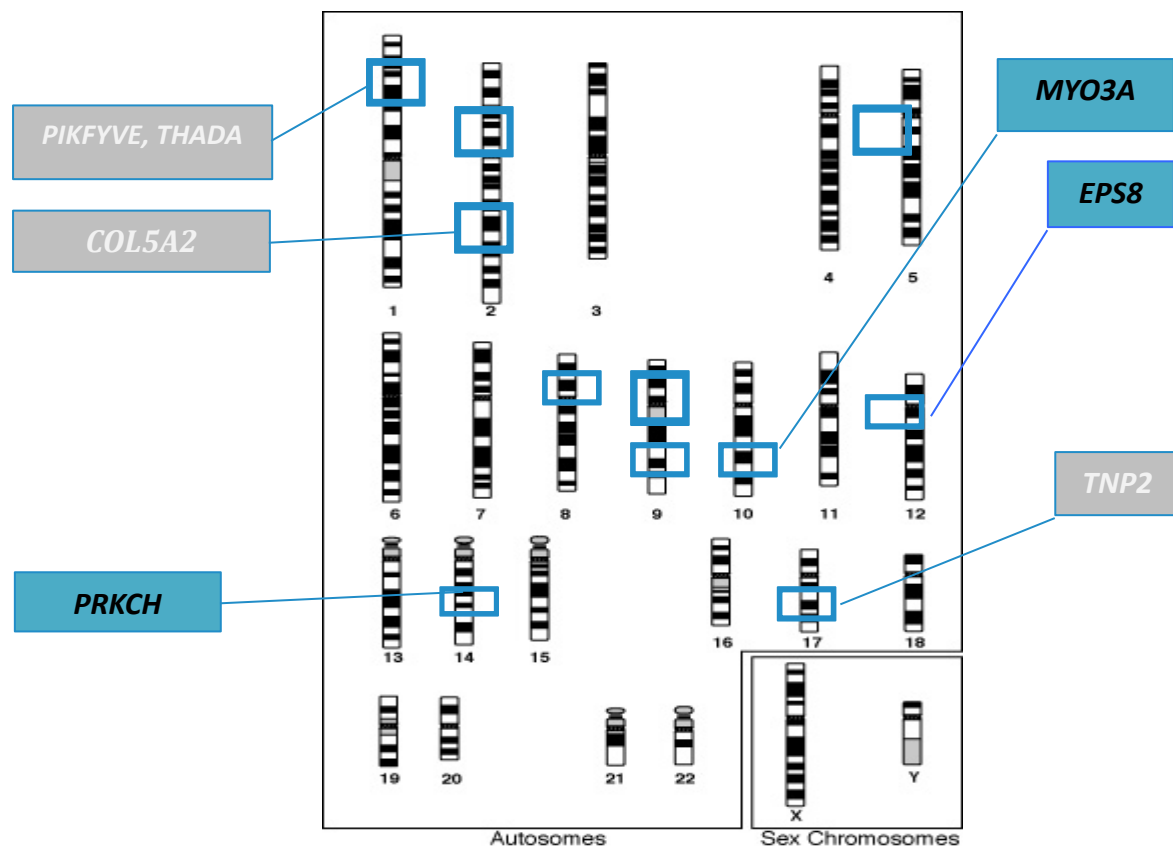


Figure 8 : Synthèse des résultats obtenus par l'analyse IBD et par le séquençage d'exome :
3 variants dans 3 gènes sont retenus (*PRKCH*, *MYO3A* et *EPS8*).
Les autres gènes (*PIKFYVE*, *THADA*, *COL5A2* et *TNP2*) ont été éliminés par l'analyse de ségrégation
Encadrés gris, les noms des gènes éliminés par l'analyse de ségrégation. Encadrés bleus, les noms des 3 gènes candidats

Discussion

a. Variants identifiés par le séquençage ciblé

Madame G. (LD0010), le cas index de la famille, présente 2 variants dans les gènes *CDKN1B* et *AIP*.

1. Le variant *CDKN1B* : c.326T>G, p.val109Gly

Il s'agit d'un variant dans l'exon 1 en position 326 avec un remplacement d'une Thymine par une Guanine responsable d'une mutation faux-sens au niveau protéique induisant le remplacement d'une valine par une glycine en position 109 (c.326T>G, p.Val109Gly). Ce variant n'est pas hérité du père et il n'a pas été transmis à ces deux filles. Ce variant est connu des différentes bases de donnée sous la référence rs2066827. La fréquence allélique de ce variant dans EXAC est supérieure à 1%. La position de l'acide-aminé ne semble pas être préservée entre les espèces contribuant à émettre l'hypothèse d'un variant non causal. Ce variant est déjà décrit dans la littérature et il ne fait pas partie des mutations décrites responsables de NEM4 (24). A ce jour, 8 mutations dans le gène *CDKN1B* sont responsables de NEM4 (24,36). Une étude menée par Longuini et al. a recherché ce variant chez les patients atteints de NEM1 avec une mutation germinale dans le gène *MEN1* et elle n'a pas montré de corrélation entre la sévérité de la maladie et la présence de ce variant (37). *In vitro*, la surexpression de ce variant dans des cellules somatotropes issues de rat ne montre pas de différence avec les cellules sauvages (38). Le variant dans le gène *CDKN1B* : exon 1 c.326T>G, p.val109Gly semble être un polymorphisme génétique sans effet fonctionnel causal d'autant plus qu'aucun autre membre atteint de la famille n'est porteur de ce variant. Il peut s'agir d'un variant apparu *de novo* ou hérité de la mère de la patiente dont le phénotype est inconnu (décès de cause cardio-vasculaire).

2. Le variant *AIP* c.469-2A>G

Il s'agit d'un variant situé dans l'intron 3 du gène *AIP*, au niveau d'un site d'épissage c.469-2A>G. Ce variant intronique peut modifier le site d'épissage selon les logiciels de prédiction (VEP, ClinVar...). Il est décrit une fois dans la littérature en 2007 par l'équipe de Cazabat et al. (CHU de Lille). Il a été retrouvé chez un patient de 40 ans présentant une acromégalie avec un macroadénome hypophysaire de 15mm (39). Ce variant n'est

également pas hérité du père de la patiente et il n'a pas été transmis à ces 2 filles. Malgré ce fait, nous ne pouvons exclure sa pathogénicité. Afin de vérifier l'implication de ce variant dans la pathologie de Madame G., nous avons entamé une étude fonctionnelle sur l'ARNm extrait du sang total. Après RT-PCR à l'aide de 4 amorces différentes, suivi du séquençage capillaire de type Sanger, nous n'avons pas mis en évidence de modification de l'épissage. Il n'y a pas d'insertion ou de délétion des bases nucléotidiques entre l'exon 3 et l'exon 4 du gène *AIP* (**Figure 9**). Le variant *AIP* intron 3 c.469-2A>G ne semble donc pas être impliqué dans la pathologie présentée par Madame G, ni être responsable des cas d'adénomes hypophysaires présentés par la famille (seul le cas index est porteur de ce variant). De plus, le type d'adénome hypophysaire ainsi que l'âge au diagnostic ne corrélaient pas avec la présentation classique des adénomes hypophysaires *AIP* mutés.

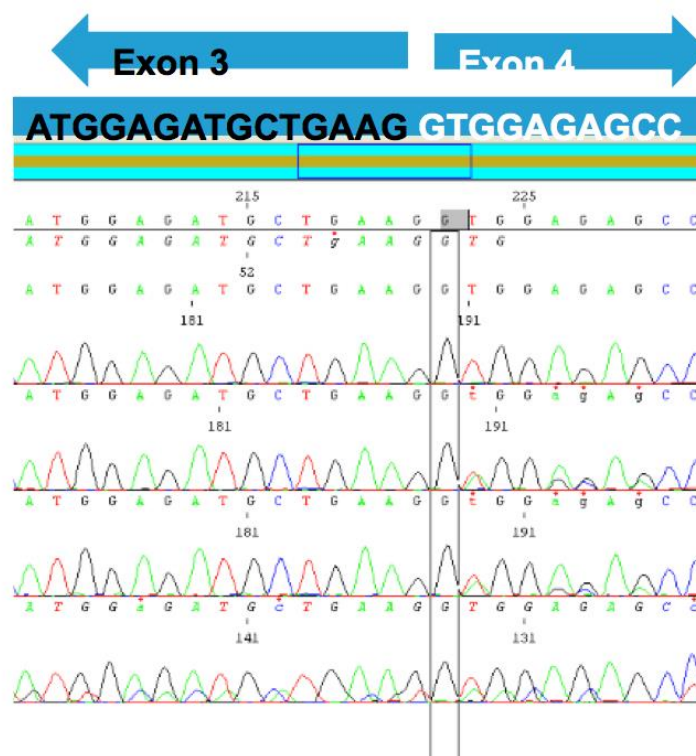


Figure 9 : Chromatogrammes montrant la séquence des exons 3 à 4 de l'ADNc de *AIP* obtenue par la méthode de SANGER à l'aide de 4 amorces différentes. Il ne montrent pas de modification de la séquence de l'ADNc suggérant que le variant *AIP* c.469-2A>G n'entraîne pas de modification de l'épissage.

3. Le variant *MEN1* exon 9 c.1254C>T, p.Asp418Asp

Le père de Madame G. (LD0014) est porteur d'un variant dans le gène *MEN1* exon 9 c.1254C>T, p.Asp418Asp observé après séquençage ciblé du gène *MEN1* devant la forte suspicion de NEM1 (association d'un adénome hypophysaire et d'une hyperparathyroïdie primaire). Ce variant est déjà répertorié dans les bases de données sous la référence rs2071313. Il a une MAF à 0,39% dans la base de données ExAC. Une mutation silencieuse est le plus souvent bénigne sauf si elle induit une modification de l'épissage au moment de l'étape de transcription de l'ADN en ARNm. Les logiciels de prédiction (VEP, Clinvar...) prédisent la création d'un site accepteur d'épissage au sein de l'exon 9 (**Figure 10**) conduisant à la création d'une protéine tronquée.



Figure 10 : Variant *MEN1* exon 9 c.1254C>T, p.Asp418Asp est une mutation silencieuse pouvant créer un site accepteur d'épissage

Les encadrés bleu représentent les introns, les Encadrés rouge représentent les exons. Les nucléotides CT au milieu de l'exon 9 peuvent créer un site accepteur d'épissage créant une insertion ou une délétion de bases nucléotidiques.

Afin de vérifier la causalité du variant *MEN1* exon 9 c.1254C>T, p.Asp418Asp, nous avons entamé une étude sur l'ARN extrait à partir de sang total (tube PAXGENE) comme pour le variant *AIP*. Après RT-PCR et séquençage capillaire par la méthode de SANGER, nous ne trouvons pas de modification de la séquence de l'ADNc. Ce variant ne semble donc pas avoir un effet sur l'épissage et ne semble pas être impliqué dans la pathologie présentée par le père de Madame G. d'autant plus qu'il n'a pas été transmis à cette dernière. De plus,

aucun autre membre de la famille qu'il soit atteint ou non, n'est porteur de ce variant. Il ne semble donc pas être causal.

b. Variants identifiés par le séquençage d'exome

Aucun des variants des gènes connus être impliqués dans les adénomes hypophysaires n'est présent chez les autres membres atteints de la famille. C'est pour cela que nous avons entrepris une démarche de séquençage d'exome afin de rechercher l'implication de nouveau(x) gène(s) dans les FIPAs. Nous avons effectué ce séquençage chez 2 patients de la famille : LD0010 et LD0009. Cette technique a révélé 21 variants dont 7 "novel" et 14 connus dans 20 gènes distincts. Madame M (LD0031) a été exclue de l'analyse de ségrégation car elle a un phénotype incertain (refus de l'imagerie hypophysaire). Trois variants co-ségrégent avec la maladie chez les 5 membres atteints de la famille : *EPS8* chr12 :15784564.TC->CG p.D619T, *PRCKH* chr14:61857988.A->T p.Thr137Ser, *MYO3A* chr10:26243804.A->C p.Asp57Ala. Les 5 autres variants novel ont été exclus car des membres de la famille - dont le statut d'atteint est certain - ne sont pas porteurs. En ce qui concerne les variants connus, l'analyse de liaison par IBD nous a permis d'en éliminer certains. Ainsi, les gènes *PIKFYVE*, *THADA* et *TNP2*, ont été exclus du fait que plusieurs individus sains étaient porteurs de ces variants probablement bénins.

1. Les variants répertoriés dans les bases de données

I. Le variant *PRCKH* chr14:61857988.A->T p.Thr137Ser

Le gène *PRCKH* code une protéine kinase de type C qui régule la différenciation des keratinocytes par l'action de la MAP kinase. Le variant *PRCKH* chr14:61857988.A->T est une mutation faux-sens engendrant le remplacement de la thréonine en position 137 par une sérine (p.Thr137Ser). Il est référencé dans les bases de données sous le numéro rs142236183. Il a une MAF à $4.524e^{-05}$ % dans la population européenne non finlandaise dans la base de données EXAC. Il n'est pas associé à un phénotype donné. Cependant, selon les logiciels de prédiction Polyphen/SIFT, ce variant est considéré comme bénin. Par ailleurs, d'autres polymorphismes du gène *PRCKH* ont été décrits comme associés à une susceptibilité aux accidents vasculaires cérébraux (40)

II. Le variant *MYO3A* chr10:26243804.A->C p.Asp57Ala

Le gène *MYO3A* code une myosine de type 3a. Des mutations dans ce gène ont été décrites en 2002 responsables de surdit  d'origine autosomique r cessive (41). Le variant *MYO3A* chr10:26243804.A->C est  galement une mutation faux-sens avec remplacement de l'acide aspartique en position 57 par une alanine (p.Asp57Ala). Il est r f renc  sous le num ro rs14651180. Il est tr s rare dans la base de donn es EXAC avec une MAF dans la population europ enne non finlandaise   0.0009159%. Les logiciels Polyphen/SIFT pr disent un effet probablement pathog ne. Ce variant est class  de signification inconnue (VUS) par ClinVar. La causalit  de ce g ne ne peut  tre  cart , il peut  tre responsable de FIPA. Un meilleur recrutement au sein de la famille permettra peut- tre de l' liminer.

2. Le variant "novel" *EPS8* chr12 :15784564.TC->CG p.Asp619Arg

Il s'agit de 2 variants contigus, jamais r pertori s dans les bases de donn es. Il s'agit d'une mutation faux-sens avec le remplacement de l'acide aspartique en position 619 par une arginine. Le g ne *EPS8* code une prot ine impliqu e dans la cascade enzymatique du r cepteur   l'EGF (Epidermal growth factor) (**Figure 11**).

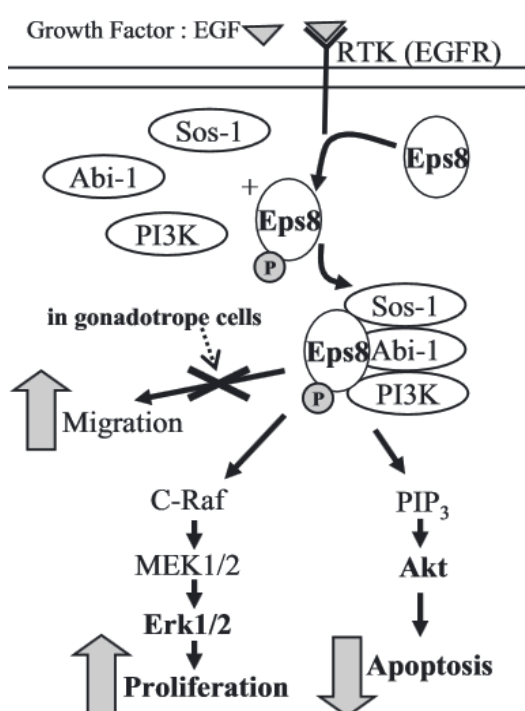


Figure 11 : Cascade enzymatique impliquant la prot ine EPS8 avec le r cepteur   l'EGF. Suite   la liaison de l'EGF sur EGFR, EPS8 est activ e induisant une acc l ration de la prolif ration cellulaire avec une diminution de l'apoptose.

Behlouli et al. (2014) a identifié par séquençage d'exome une mutation homozygote dans l'exon 3 (c.88C->T) d'*EPS8* responsable de l'apparition prématurée d'un codon stop en position protéique 30 (p.gln30X (Q30X)). Cette mutation a été identifiée chez une famille algérienne consanguine atteinte de surdité congénitale (42). A ce jour, aucune autre mutation germinale n'est décrite dans la littérature médicale. Cependant, une étude menée par Xu M. et al. (2009) a comparé la signature moléculaire de 9 adénomes gonadotropes contre 9 hypophyses normales issue d'autopsie. Elle montre une surexpression du gène *EPS8* (x 5,9) en comparaison d'hypophyses saines suggérant l'implication de *EPS8* dans la prolifération cellulaire (43). Une autre étude menée par Jiang et al (2012) a montré une surexpression de 6 gènes dont *EPS8* après une analyse transcriptomique de 7 adénomes hypophysaires en comparaison de 3 tissus hypophysaires sains (44).

Après analyse des données de la littérature et des publications impliquant le rôle de *EPS8* dans la prolifération cellulaire, nous avons émis l'hypothèse que le variant *EPS8* chr12:15784564.TC->CG peut procurer soit une surexpression du gène soit un gain de fonction de la protéine *EPS8*, responsable d'une prolifération cellulaire pouvant être impliquée dans la formation des adénomes hypophysaires dans notre famille.

c. Limites de la méthode utilisée

La difficulté des analyses bio-informatiques réside dans le choix des filtres, principalement des bases de données, afin d'éliminer les variants fréquents. Cela nécessite de connaître leur qualité et leur composition. Il est primordial de consulter des bases de données composées d'individus démographiquement compatibles (contrôles locaux) afin d'éliminer les variants spécifiques de nos régions. A l'inverse, le nombre d'individus composant les bases de données ne prédit pas systématiquement une capacité à filtrer un grand nombre de variants. En effet, la base de données « Complete Genomics » composée de seulement 69 génomes est plus efficace (76.5% des variants éliminés) que les bases de données génomiques de UK10K (72.2% d'élimination) possédant plus de 2000 génomes, malgré un seuil plus élevé. Ceci s'explique par le fait que le seuil, fixé à 0,01% car estimé à partir de la prévalence (très rare) des adénomes hypophysaires, nécessite des bases de données composées au minimum de 1000 allèles. Dans le cas contraire, ce seuil est

augmenté. De plus, les données ne donnent pas accès au phénotype des individus pouvant suggérer la présence de patients avec une pathologie similaire au sein de la base. L'application successive de ces filtres, permet d'éliminer 99.5% des variants identifiés et ainsi de ne conserver que les variants potentiellement causaux. L'ordre des filtres est également un paramètre important dans la sélection des variants. En effet, l'inversion des étapes de sélection des variants par les algorithmes et par leur partage entre les individus séquencés peut éliminer un variant s'il n'a été détecté que par l'un des 3 algorithmes alors qu'il était commun à tous les individus.

d. Limites du séquençage d'exome

La couverture de séquençage constitue une autre limite de l'analyse des exomes. En moyenne, 2 à 3% des bases des exons ne sont pas couvertes au minimum dix fois selon les résultats du contrôle qualité (régions difficiles à séquencer). On ne peut donc exclure que le variant causal se situe parmi les bases non couvertes. La solution pour palier à cette limite est le séquençage du génome entier. Il présente l'avantage d'une couverture homogène de séquençage en raison de l'absence de capture, pour un coût encore supérieur au séquençage d'exome mais qui ne cesse de baisser. De plus, il permet d'explorer, en complément des régions codantes, les variations du nombre de copies (CNV) ainsi que toutes les régions non codantes. Il apparaît indispensable d'étudier les zones non codantes du génome dont les éléments régulateurs et les régions promotrices. Cette approche reste toutefois limitée par la disponibilité, à ce jour, des bases de données répertoriant tous les variants du génome humain, y compris des régions non codantes et des difficultés d'interprétation que cela engendre.

e. Limites de l'analyse de ségrégation

Nous n'avons pas pu éliminer tous les variants identifiés par séquençage d'exome après analyse de ségrégation. Un nombre plus important d'individus appartenant à la famille pourrait permettre une analyse plus fine. La difficulté réside dans le recrutement des autres membres de la famille. Cela implique la réalisation d'une imagerie hypophysaire de type IRM et un bilan hormonal à tous les membres de la famille souhaitant participer à ce travail.

f. Conclusion et perspectives

Le gène *EPS8* semble être le gène candidat principal pouvant être impliqué dans les adénomes hypophysaires et responsable de FIPA dans notre famille. Nous avons fait l'hypothèse que le variant *EPS8* chr12:15784564.TC->CG p.D56T par le biais d'une surexpression ou gain de fonction peut être responsable de la survenue des adénomes hypophysaires chez les membres atteints de la famille de Madame G.

Deux démarches complémentaires peuvent nous aider à démontrer la causalité d'*EPS8* dans les FIPAs. Premièrement, nous envisageons de tester notre gène candidat sur une cohorte plus large de familles atteintes de FIPA voire de cas sporadiques atteints d'adénome hypophysaire, ne présentant pas de mutations dans les gènes connus (*MEN1*, *CDKN1B*, *AIP*) à ce jour. Ce travail sera réalisé en collaboration avec le Pr CHANSON de l'hôpital Kremlin-Bicêtre ainsi que le Dr CASTINETTI et le Pr BARLIER du CHU de Marseille. Parallèlement, Nous avons entamé des études fonctionnelles spécifiques du variant *EPS8* découvert dans cette famille. Nous avons généré un plasmide porteur du gène *EPS8* sauvage (WT) et par le biais d'une mutagenèse dirigée, nous avons généré un plasmide porteur du variant *EPS8* chr12:15784564.TC->CG trouvé dans notre famille atteinte de FIPA. Dans un second temps, nous envisageons la transfection des plasmides dans des cellules neutres de type 293 et des cellules hypophysaires spécialisées de type gonadotrope afin d'étudier la réponse signalétique du récepteur de l'EGF après stimulation par de l'EGF. Enfin, l'on peut imaginer poursuivre les études fonctionnelles du rôle d'*EPS8* par la reprogrammation de cellules IPS à partir de cellules épithéliales urinaires et leur différenciation en cellule ante-hypophysaire. Ce travail sera poursuivi au sein de l'équipe du Pr Bertrand CARIOU à l'institut du thorax.

Bibliographie

1. Faglia G. Epidemiology and pathogenesis of pituitary adenomas. *Acta Endocrinol (Copenh)*. 1993 Jul;129 Suppl 1:1–5.
2. Daly AF, Rixhon M, Adam C, Dempegioti A, Tichomirowa MA, Beckers A. High prevalence of pituitary adenomas: a cross-sectional study in the province of Liege, Belgium. *J Clin Endocrinol Metab*. 2006 Dec;91(12):4769–75.
3. Mindermann T, Wilson CB. Age-related and gender-related occurrence of pituitary adenomas. *Clin Endocrinol (Oxf)*. 1994 Sep;41(3):359–64.
4. Gold EB. Epidemiology of pituitary adenomas. *Epidemiol Rev*. 1981;3:163–83.
5. Asa SL. The pathology of pituitary tumors. *Endocrinol Metab Clin North Am*. 1999 Mar;28(1):13–43, v – vi.
6. Fontana E, Gaillard R. [Epidemiology of pituitary adenoma: results of the first Swiss study]. *Rev Med Suisse*. 2009 Oct 28;5(223):2172–4.
7. Beckers A, Aaltonen LA, Daly AF, Karhu A. Familial isolated pituitary adenomas (FIPA) and the pituitary adenoma predisposition due to mutations in the aryl hydrocarbon receptor interacting protein (AIP) gene. *Endocr Rev*. 2013 Apr;34(2):239–77.
8. Daly AF, Tichomirowa MA, Beckers A. The epidemiology and genetics of pituitary adenomas. *Best Pract Res Clin Endocrinol Metab*. 2009 Oct;23(5):543–54.
9. Daly AF, Jaffrain-Rea M-L, Ciccarelli A, Valdes-Socin H, Rohmer V, Tamburrano G, et al. Clinical characterization of familial isolated pituitary adenomas. *J Clin Endocrinol Metab*. 2006 Sep;91(9):3316–23.
10. Vasilev V, Daly A, Naves L, Zacharieva S, Beckers A. Clinical and genetic aspects of familial isolated pituitary adenomas. *Clin Sao Paulo Braz*. 2012;67 Suppl 1:37–41.
11. Burlacu MC, Tichomirowa M, Daly A, Beckers A. Adénomes hypophysaires familiaux. *Presse Médicale*. 2009 Jan;38(1):112–6.
12. Wermer P. ENDOCRINE ADENOMATOSIS AND PEPTIC ULCER IN A LARGE KINDRED. INHERITED MULTIPLE TUMORS AND MOSAIC PLEIOTROPISM IN MAN. *Am J Med*. 1963 Aug;35:205–12.
13. Trump D, Farren B, Wooding C, Pang JT, Besser GM, Buchanan KD, et al. Clinical studies of multiple endocrine neoplasia type 1 (MEN1). *QJM Mon J Assoc Physicians*. 1996 Sep;89(9):653–69.
14. Agarwal SK. Multiple endocrine neoplasia type 1. *Front Horm Res*. 2013;41:1–15.
15. Machens A, Schaaf L, Karges W, Frank-Raue K, Bartsch DK, Rothmund M, et al. Age-related penetrance of endocrine tumours in multiple endocrine neoplasia type 1 (MEN1): a multicentre study of 258 gene carriers. *Clin Endocrinol (Oxf)*. 2007 Oct;67(4):613–22.
16. Falchetti A. Genetics of multiple endocrine neoplasia type 1 syndrome: what's new and what's old. *F1000Research* [Internet]. 2017 Jan 24 [cited 2017 Jul 7];6. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5288685/>
17. Larsson C, Skogseid B, Oberg K, Nakamura Y, Nordenskjöld M. Multiple endocrine neoplasia type 1 gene maps to chromosome 11 and is lost in insulinoma. *Nature*. 1988 Mar 3;332(6159):85–7.
18. Emmert-Buck MR, Lubensky IA, Dong Q, Manickam P, Guru SC, Kester MB, et al. Localization of the multiple endocrine neoplasia type I (MEN1) gene based on tumor loss of heterozygosity analysis. *Cancer Res*. 1997 May 15;57(10):1855–8.
19. Chandrasekharappa SC, Guru SC, Manickam P, Olufemi SE, Collins FS, Emmert-Buck MR, et al. Positional cloning of the gene for multiple endocrine neoplasia-type 1. *Science*. 1997 Apr 18;276(5311):404–7.
20. Wautot V, Khodaei S, Frappart L, Buisson N, Baro E, Lenoir GM, et al. Expression analysis of endogenous menin, the product of the multiple endocrine neoplasia type 1 gene,

in cell lines and human tissues. *Int J Cancer*. 2000 Mar 15;85(6):877–81.

21. Vergès B, Boureille F, Goudet P, Murat A, Beckers A, Sassolas G, et al. Pituitary disease in MEN type 1 (MEN1): data from the France-Belgium MEN1 multicenter study. *J Clin Endocrinol Metab*. 2002 Feb;87(2):457–65.
22. Kouvaraki MA, Lee JE, Shapiro SE, Gagel RF, Sherman SI, Sellin RV, et al. Genotype-phenotype analysis in multiple endocrine neoplasia type 1. *Arch Surg Chic Ill 1960*. 2002 Jun;137(6):641–7.
23. Lemos MC, Thakker RV. Multiple endocrine neoplasia type 1 (MEN1): analysis of 1336 mutations reported in the first decade following identification of the gene. *Hum Mutat*. 2008 Jan;29(1):22–32.
24. Pellegata NS, Quintanilla-Martinez L, Siggelkow H, Samson E, Bink K, Höfler H, et al. Germ-line mutations in p27Kip1 cause a multiple endocrine neoplasia syndrome in rats and humans. *Proc Natl Acad Sci U S A*. 2006 Oct 17;103(42):15558–63.
25. Pellegata NS. MENX and MEN4. *Clin Sao Paulo Braz*. 2012;67 Suppl 1:13–8.
26. Bertherat J, Horvath A, Groussin L, Grabar S, Boikos S, Cazabat L, et al. Mutations in regulatory subunit type 1A of cyclic adenosine 5'-monophosphate-dependent protein kinase (PRKAR1A): phenotype analysis in 353 patients and 80 different genotypes. *J Clin Endocrinol Metab*. 2009 Jun;94(6):2085–91.
27. Stratakis CA, Kirschner LS, Carney JA. Clinical and molecular features of the Carney complex: diagnostic criteria and recommendations for patient evaluation. *J Clin Endocrinol Metab*. 2001 Sep;86(9):4041–6.
28. Kirschner LS, Sandrini F, Monbo J, Lin JP, Carney JA, Stratakis CA. Genetic heterogeneity and spectrum of mutations of the PRKAR1A gene in patients with the carney complex. *Hum Mol Genet*. 2000 Dec 12;9(20):3037–46.
29. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010 Sep;20(9):1297–303.
30. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinforma Oxf Engl*. 2009 Aug 15;25(16):2078–9.
31. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. *Bioinforma Oxf Engl*. 2011 Aug 1;27(15):2156–8.
32. Cunningham F, Amode MR, Barrell D, Beal K, Billis K, Brent S, et al. Ensembl 2015. *Nucleic Acids Res*. 2015 Jan;43(Database issue):D662–9.
33. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput Biol*. 2010;6(12):e1001025.
34. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res*. 2010 Jan;20(1):110–21.
35. Lindenbaum P, Le Scouarnec S, Portero V, Redon R. Knime4Bio: a set of custom nodes for the interpretation of next-generation sequencing data with KNIME. *Bioinforma Oxf Engl*. 2011 Nov 15;27(22):3200–1.
36. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, et al. The Ensembl Variant Effect Predictor. *Genome Biol* [Internet]. 2016 Jun 6 [cited 2016 Aug 29];17. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4893825/>
37. Longuini VC, Lourenço DM, Sekiya T, Meirelles O, Goncalves TD, Coutinho FL, et al. Association between the p27 rs2066827 variant and tumor multiplicity in patients harboring MEN1 germline mutations. *Eur J Endocrinol Eur Fed Endocr Soc*. 2014 Sep;171(3):335–42.
38. Sekiya T, Bronstein MD, Benfini K, Longuini VC, Jallad RS, Machado MC, et al. p27 variant and corticotropinoma susceptibility: a genetic and in vitro study. *Endocr Relat*

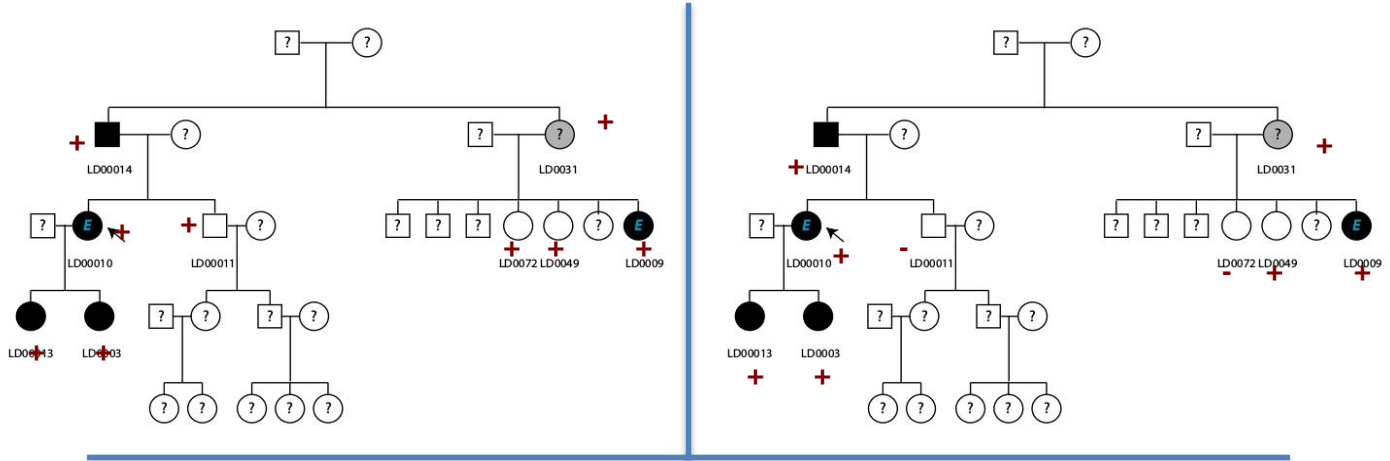
Cancer. 2014 Jun;21(3):395–404.

39. Cazabat L, Libè R, Perlemoine K, René-Corail F, Burnichon N, Gimenez-Roqueplo A-P, et al. Germline inactivating mutations of the aryl hydrocarbon receptor-interacting protein gene in a large cohort of sporadic acromegaly: mutations are found in a subset of young patients with macroadenomas. *Eur J Endocrinol Eur Fed Endocr Soc.* 2007 Jul;157(1):1–8.
40. Kubo M, Hata J, Ninomiya T, Matsuda K, Yonemoto K, Nakano T, et al. A nonsynonymous SNP in PRKCH (protein kinase C ϵ) increases the risk of cerebral infarction. *Nat Genet.* 2007 Feb;39(2):212–7.
41. Walsh VL, Raviv D, Dror AA, Shahin H, Walsh T, Kanaan MN, et al. A mouse model for human hearing loss DFNB30 due to loss of function of myosin IIIA. *Mamm Genome Off J Int Mamm Genome Soc.* 2011 Apr;22(3-4):170–7.
42. Behlouli A, Bonnet C, Abdi S, Bouaita A, Lelli A, Hardelin J-P, et al. EPS8, encoding an actin-binding protein of cochlear hair cell stereocilia, is a new causal gene for autosomal recessive profound deafness. *Orphanet J Rare Dis.* 2014 Apr 17;9:55.
43. Xu M, Shorts-Cary L, Knox AJ, Kleinsmidt-DeMasters B, Lillehei K, Wierman ME. Epidermal growth factor receptor pathway substrate 8 is overexpressed in human pituitary tumors: role in proliferation and survival. *Endocrinology.* 2009 May;150(5):2064–71.
44. Jiang Z, Gui S, Zhang Y. Analysis of differential gene expression in plurihormonal pituitary adenomas using bead-based fiber-optic arrays. *J Neurooncol.* 2012 Jul;108(3):341–8.

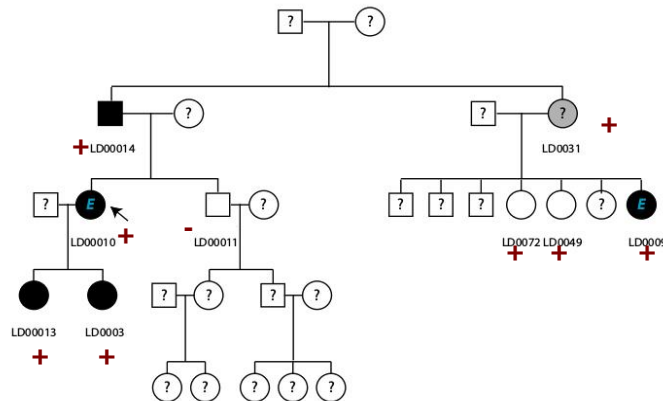
Annexe

A- PIKFYVE chr2 :209182628 A->G p.N682S

B- THADA chr2 :43768343 C->T p.M783I



C- TNP2 chr16 :11362741 T->C p.K127E

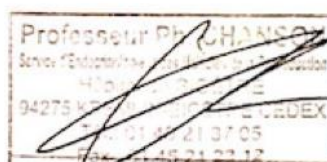


Annexe 1: Résultats de l'analyse de ségrégation pour les 3 variants connus retrouvés par exome couplé à l'analyse « IBD linkage ». Ils sont représentés sur un arbre généalogique simplifié de la famille

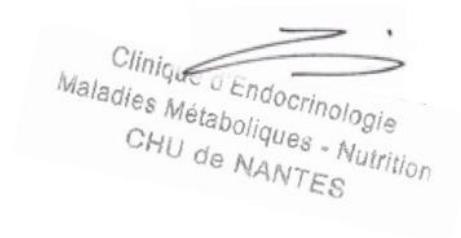
A- PIKFYVE p.N682S ; B- THADA chr2 :43768343 C->T p.M783I ; C- TNP2 p.K127E

Arbre généalogique : le carré désigne un individu de sexe masculin, le rond désigne un individu de sexe féminin, la flèche noire indique le cas index. Les patients affectés sont représentés en noir. Les apparentés désignés par '?' ont un phénotype inconnu. Le « E » désigne les patients séquencés par exome.

Vu, le Président du Jury,
(tampon et signature)



Vu, le Directeur de Thèse,
(tampon et signature)



Vu, le Doyen de la Faculté,
(tampon et signature)

Adénomes hypophysaires isolés familiaux :

Stratégie pour l'identification d'une nouvelle cause génétique : Observation clinique, stratégie de recherche, et revue de la littérature

Résumé

Les adénomes hypophysaires peuvent être syndromiques dans le contexte de néoplasies endocriniennes multiples (NEM1 ou NEM4) ou isolés (adénomes hypophysaires isolés familiaux ou FIPA). Les FIPA est dû à des mutations d'*AIP* dans 15% des cas mais dans plus de 80% des cas, la cause génétique demeure inconnue.

Nous avons étudié une famille nantaise atteinte de FIPA. L'analyse ciblée des gènes candidats a identifié trois variants dans les gènes *MEN1*, *CDKN1B* et *AIP*, qui ne seraient pas causals. Le séquençage d'exome chez deux patients a identifié 21 variants. Après les analyses de ségrégation et de liaison génétique d'identité par descendance dans la famille, trois gènes, *MYO3A*, *PRKCH* et *EPS8* s'avèrent des candidats potentiels.

EPS8 est particulièrement attractif, étant donné son implication dans la voie de signalisation de l'EGF. Une surexpression d'*EPS8* ou un gain de fonction peut s'accompagner d'une prolifération cellulaire et d'une diminution de l'apoptose et a été associé à des adénomes hypophysaires. Afin de montrer sa responsabilité dans les FIPA, il sera nécessaire de trouver d'autres familles mutées et de montrer *in vitro* l'impact fonctionnel de ses mutations.

L'ensemble des résultats conduira à la découverte d'un nouveau gène responsable de FIPA ouvrant de nouveaux espoirs pour le développement de thérapies ciblées.

Mots clés : adénome, hypophyse, *AIP*, exome, *EPS8*