



**HAL**  
open science

# Du génome humain au modèle murin : compréhension des ciliopathies

Clarisse Delvallee

► **To cite this version:**

Clarisse Delvallee. Du génome humain au modèle murin : compréhension des ciliopathies. Génétique. Université de Strasbourg, 2020. Français. NNT : 2020STRAJ039 . tel-03662407

**HAL Id: tel-03662407**

**<https://theses.hal.science/tel-03662407>**

Submitted on 9 May 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Thèse présentée pour l'obtention du grade de :  
**Docteur de l'Université de Strasbourg**

Discipline : Science de la vie et de la santé (ED 414)

Spécialité : Aspects moléculaires et cellulaires de la Biologie

par

**Clarisse DELVALLEE**

## **DU GENOME HUMAIN AU MODELE MURIN : COMPREHENSION DES CILIOPATHIES**

Laboratoire de Génétique médicale, INSERM UMR-S 1112, Strasbourg

Soutenue publiquement le 1<sup>er</sup> octobre 2020 devant le jury :

Directrice de thèse	Mme le Professeur Hélène DOLLFUS, Strasbourg
Rapporteur externe	Mme le Docteur Valérie BUEE-SCHERRER, Lille
Rapporteur externe	M. le Docteur Jean-Michel ROZET, Paris
Examineur interne	M. le Docteur Yvon TROTTIER, Strasbourg

Contrat doctoral financé par une bourse de la Fondation pour la recherche médicale (FRM)  
« Programme espoir de la recherche 2017 » ECO20170637509

*A toi + moi + eux + tous ceux qui le veulent...*

## *Remerciements*

---

Avant d'exposer le pan scientifique de ce manuscrit, il me tient à cœur de remercier les personnes qui m'ont accompagnée pendant ces 3 dernières années.

Pour commencer, je tiens à remercier ma directrice de thèse Mme Hélène Dollfus pour m'avoir donné l'opportunité de rejoindre le laboratoire en 2016 pour mon stage de M1. Hélène je vous remercie pour votre soutien. Malgré votre charge de travail, vous trouvez toujours le temps de nous écouter, nous orientez et prendre les décisions qui s'imposent.

Ensuite, je tiens à remercier mes 2 rapporteurs externes, Mme Valérie Buée-Scherrer et M. Jean-Michel Rozet d'avoir accepté de relire, de corriger mon manuscrit de thèse et de participer à ma soutenance. Je tiens également à remercier mon examinateur externe, M. Yvon Trottier de m'avoir suivi en participant l'an dernier à mon comité de mi-thèse et en revenant cette année pour ma soutenance. J'espère que vous apprécierez tous les 3 mon travail de thèse et ma présentation lors de la soutenance.

Je tiens particulièrement à remercier M. Jean Muller pour son soutien pendant mes 5 années au laboratoire, pour ses conseils sur les projets scientifiques, pour la rédaction des articles et du manuscrit, pour les demandes de financements... Jean, j'ai adoré travailler avec toi mais pas seulement. Ce fût un grand plaisir de rire et discuter avec toi : les déjeuners étaient bien plus animés en ta présence. J'ai beaucoup aimé te tenir compagnie au téléphone lors de tes trajets en voiture.

Je tiens à remercier Mme Corinne Stoetzel qui a pris le risque de m'accueillir au laboratoire malgré mes demandes de stage très insistantes. Elle m'a appris les bases des techniques de biologies moléculaires et cellulaires et la rigueur du travail en laboratoire. Corinne, je te considère comme ma mère du laboratoire, j'ai adoré évoluer à tes cotés aussi bien scientifiquement que personnellement.

Je tiens à remercier M. Vincent Marion pour son encadrement. Malgré certains désaccords sur le plan scientifique, vous savez challenger les gens tout en aménageant des périodes de défoulements à intervalles de temps réguliers.

Je tiens à remercier tous nos collaborateurs et leurs équipes respectives pour leur implication dans l'ensemble des projets rapportés dans ce manuscrit.

Merci à Véronique pour les discussions matinales que nous avons eu tous les jours, tu me permettait d'évacuer la tension et de commencer les journées avec le sourire. Merci à Cathy, parce que même si tu es débordée tu trouves toujours le moyen de répondre à nos questions, de nous donner un coup de main ou de chanter « ragoût toutou ».

Je souhaite à présent dire un grand merci à mes collègues qui sont devenus des amis au fil du temps. Merci à Daniel de préparer le café tous les matins. Merci à Amandine et Laurie qui sont des filles tops



## *Remerciements*

---

et qui apportent de l'énergie et du sang neuf au laboratoire. Merci aux personnes que j'ai pu côtoyer au cours de ma thèse et qui sont parties depuis. Merci à Katia pour sa joie de vivre et son humour. Merci à Aline pour les « bonjours » à longueur de journée et les expériences du week-end en belle compagnie. Merci à Catherine, mamie Tiphon, Agnès, Ariane. Merci à Jérémie, Céleste, Chiar et Géraldine, mes amis de l'équipe Mensah.

Je veux transmettre « mille mercis » et bien plus à mon entourage car ils ont toujours été présents pour moi pendant cette aventure. Merci aux 4 filles du mont-Rabié : Marie, Mélanie, Estelle et Emilie. Merci à ma famille, mes parents Sylvie et Thierry et mon frère Benjamin parce que vous m'avez supporté ! Vous m'avez écouté parler de manip' alors que vous ne me comprenez pas toujours ce que je dis. Et surtout, vous m'avez toujours soutenu tout au long de mon cursus malgré les moments difficiles. Merci à l'homme qui partage ma vie. Tu es le meilleur !! Tu m'as laissé investir ton appartement pour l'écriture et en plus, tu es parvenu à me supporter : Bel exploit ! Merci à nos animaux de compagnies, les 2 pépettes et surtout à Naggy, dit Gytoune, dit la boule, dit grosse tête, pour les promenades où on ne sait pas qui promène qui mais qui permettent d'aérer l'esprit.

## Abréviations

---

AA : Acide-Aminé	KI : Knock-In
AC : Anticorps	KIT : Institut de Technologie de Karlsruhe
ADN : Acide DésoxyriboNucléique	KO : Knock-Out
ALMS : Syndrome d'Alström	LCA : Amaurose Congénitale de Leber
APP : Amyloid Precursor Protein	LGM : Laboratoire de Génétique Médicale
ARN : Acide Ribonucléique	LINE : Long Interspersted Element
BBS : Syndrome de Bardet Biedl	ME : Microscopie Electronique
CA : Corne d'Amon	MEI : Insertion d'éléments Mobiles
CNV : Variation du Nombre de Copie	MET : Microscopie Electronique à Transmission
Co-IP : Co-Immunoprécipitation	MKS : Syndrome de Meckel-Gruber
CQ : Chloroquine	MT : Microtubules
DNF : Dégénérescence neurofibrillaire	Mu : Murin
Domaine PH : Domaine homologue à la Pleckstrine	NGS : Séquençages de Nouvelles Générations
ERG : Electrorétinogramme	NHC : Nouvel Hôpital Civil
EVC : Ellis Van Creveld	NPHP : Néphronophtise
GCL : Couche des Cellules Ganglionnaires	ONL : Couche Nucléaire Externe
GFP : Green Fluorescence Protein	OPL : Couche Plexiforme Externe
Glc : Glucose	OS : Segment Externe des Photorécepteurs
H&E : Coloration Hématoxyline-Éosine	PAP : Polydactylie Post Axiale
Hh : Hedghog signaling	PCD : Dyskinésie Ciliaire Primitive
Hu : Humain	PCR : Réaction de Polymérisation en Chaîne
HUS : Services Hospitalo-Universitaires	PFA : Paraformaldéhyde
I2 : Nourriture spéciale des Souris	PI3K : Phosphatidyl-Inositol 3 Kinase
IBMC : Institut De Biologie Moléculaire Et Cellulaire	Pi3P : Phosphatidyl-Inositol 3 Phosphate
ICS : Institut Clinique de la Souris	PiP : Phosphatidyl-Inositol Phosphate
IF : Immunofluorescence	PM : Poids Moléculaire
IFT : Transport Intraflagellaire	qRTPCR : Rétrotranscription PCR quantitative
IGBMC : Institut de Génétique, Biologie Moléculaire et Cellulaire	RE : Réticulum Endoplasmique
INL : Couche Nucléaire Interne	RP : Rétinopathie Pigmentaire
IPL : Couche Plexiforme Interne	SAG : SMO Agonist
IS : Segment Interne des Photorécepteurs	SHH : Sonic Hedgehog
JBTS : Syndrome De Joubert	SINE : Short Interspersted Element
	SMO : Smoothened receptor
	SVA : SINE-VNTR-Alu Retrotransposon

## *Abréviations*

---

SVF : Sérum de Veau Fœtal

TES : Target Exome Sequencing

TSD : Target Site Duplication

VNTR : Variable Number Of Tandem Repeat

VUS/VSİ : Variant de Signification Inconnue

WB : Western Blot

WES : Whole Exome Sequencing

WGS : Whole Genome Sequencing

WT : Wild Type

<b>I. Introduction.....</b>	<b>1</b>
<b>I. Les cils .....</b>	<b>1</b>
1. La chronologie des découvertes .....	1
2. Les 2 types de cils .....	2
A) Le cil motile « 9+2 ».....	2
B) Le cil primaire « 9+0 » .....	3
<b>II. Le fonctionnement du cil primaire .....</b>	<b>5</b>
1. Le transport intraflagellaire (IFT) .....	5
A) Le complexe A .....	7
B) Le complexe B .....	7
2. Le BBSome .....	7
3. Les voies de signalisations médiées par le cil .....	9
A) La voie de signalisation Sonic-Hedgehog (SHH) .....	9
➤ L'EVC zone (Ellis Van Creveld) .....	11
B) L'autophagie.....	12
<b>III. Les pathologies associées au fonctionnement du cil .....</b>	<b>15</b>
1. Les ciliopathies.....	15
2. Le syndrome de Bardet-Biedl (BBS).....	16
A) Contribution de chaque gène BBS dans le syndrome .....	17
B) Présentation de certains gènes BBS .....	17
➤ <i>BBS1</i> .....	17
➤ <i>BBS5</i> .....	18
➤ <i>IFT27/BBS19</i> .....	19
3. Les autres ciliopathies et les gènes associés .....	20
A) <i>IQCE</i> .....	20
B) <i>VPS15</i> .....	21
➤ Le rôle de <i>VPS15</i> dans le processus d'autophagie .....	21
➤ Le rôle de <i>VPS15</i> dans le cil primaire .....	22
4. Les maladies neurodégénératives .....	23
A) Les amyloïdoses .....	24
B) La maladie d'Alzheimer : .....	24
➤ Les plaques amyloïdes .....	25
➤ Les protéines TAU phosphorylées et les DNF .....	26
➤ L'impact de l'autophagie.....	26
<b>IV. La recherche translationnelle.....</b>	<b>28</b>
1. Le diagnostic clinique .....	28

A) Les difficultés associées au diagnostic clinique .....	29
2. Le diagnostic génétique .....	30
A) Les difficultés associées au diagnostic génétique .....	30
B) L'apport des techniques NGS .....	31
➤ Le séquençage ciblé ou panel de gènes (TES) .....	35
➤ Le séquençage complet de l'exome (WES) .....	35
➤ Le séquençage du génome complet (WGS) .....	36
<b>V. Les modèles d'étude pour les ciliopathies .....</b>	<b>37</b>
1. Les modèles cellulaires .....	37
2. Les modèles poisson-zèbres .....	38
3. Les modèles murins .....	39
A) Comparaison des tissus humains et murins .....	39
➤ L'œil .....	39
➤ Le rein .....	41
➤ Le cerveau .....	42
B) Les modèles murins <i>Vps15</i> .....	44
➤ Le modèle KO <i>Vps15</i> de Nemazany et al. 2013 .....	44
➤ Le modèle <i>Vps15</i> de Gstrein et al. 2018. ....	45
<b>VI. Les problématiques .....</b>	<b>47</b>
<b>II. Matériel et méthodes.....</b>	<b>49</b>
<b>I. Le matériel.....</b>	<b>49</b>
1. Tableau 2 : Listes des amorces .....	49
2. Tableau 3 : Liste des anticorps (AC).....	51
3. Tableau 4 : Liste des buffers .....	52
4. Tableau 5 : Liste des nourritures pour souris .....	52
5. Tableau 6 : Liste des logiciels.....	53
<b>II. Les méthodes .....</b>	<b>54</b>
1. La culture cellulaire de fibroblastes humains .....	54
A) Induction de la ciliogénèse.....	54
B) Stimulation de la voie Sonic Hedgehog (SHH) .....	54
2. Les explants : culture primaire de fibroblastes murins .....	54
A) Induction de la ciliogénèse.....	55
B) Stimulation/blocage du processus d'autophagie .....	55
3. Les techniques de biologie moléculaire.....	55
A) Design des amorces.....	55
B) PCR et électrophorèse, élution du fragment depuis le gel.....	55

C) Séquençage Sanger .....	56
D) Extraction d'ARN et synthèse d'ADNc.....	56
E) qRTPCR .....	57
4. Les techniques de biochimie .....	57
A) Extractions protéiques .....	57
B) Quantification protéique (test de Bradford) .....	58
C) Poudre d'organes et test ELISA (partie 2 uniquement) .....	58
D) Co-immunoprécipitation (Co-IP) .....	59
E) Western Blot .....	60
F) Immunocytofluorescence et Immunohistofluorescence .....	60
5. Les techniques dédiées aux souris.....	61
A) La génération de la lignée par l'ICS .....	61
B) Maintien de la lignée.....	63
C) Génotypage de l'ADNg et de l'ADNc .....	64
D) Les analyses fonctionnelles .....	65
➤ Electrorétinogramme (ERG) .....	65
➤ Prélèvements sanguins au niveau du sinus rétro-orbital.....	66
➤ Y-maze et frottis vaginaux .....	66
➤ Prélèvements et dissections .....	67
(a) Microdissection des yeux .....	67
(b) Découpe des cerveaux .....	67
(c) Perfusion au PFA 4%.....	68
<b>III. Les méthodes réalisées en collaboration .....</b>	<b>69</b>
1. Analyse cérébrale : Yalcin Binnaz (IGBMC, Illkirch) .....	69
2. Etude histologique : Cristina Antal (Faculté de médecine, Strasbourg) .....	69
3. Microscopie électronique à transmission (MET) : Nadia Messadecq (IGBMC, Illkirch).....	70
4. Spectrométrie de masse : Philippe Hamann (IBMC, Strasbourg) .....	70
5. Etudes portant sur les poissons zèbres : Uwe Strahle et Christelle Etard (KIT, Karlsruhe).....	72
 <b>III. <u>Partie 1</u> : La validation de l'impact fonctionnel de mutations identifiées dans des gènes de ciliopathies.....</b>	 <b>73</b>
<b>Le projet I portant sur <i>IFT27</i>.....</b>	<b>73</b>
<b>I. Le synopsis.....</b>	<b>73</b>
<b>II. Les données supplémentaires.....</b>	<b>74</b>
<b>III. La conclusion et les perspectives.....</b>	<b>75</b>

IV. L'article.....	75
<b>Le projet II portant sur <i>BBS1</i>.....</b>	<b>86</b>
I. Les rétrotransposons.....	86
II. Le synopsis.....	88
III. Les données supplémentaires.....	92
IV. La conclusion et les perspectives.....	94
V. Le manuscrit .....	96
<b>Le projet III portant sur <i>IQCE</i>.....</b>	<b>146</b>
I. Le synopsis.....	146
II. Les données supplémentaires.....	149
III. La conclusion et les perspectives.....	150
IV. L'article.....	151
<b>Le projet IV portant sur <i>BBS5</i>.....</b>	<b>185</b>
I. Le synopsis .....	185
II. Les résultats.....	188
1. Les expériences portant sur la délétion.....	188
A) Détection par PCR et séquençage Sanger .....	188
B) Exploration de la cohorte par PCR duplex.....	189
C) L'étude de l'expression de l'ARNm et protéique pour <i>BBS5</i> .....	189
2. Les expériences portant sur la duplication .....	191
A) Confirmation de la co-ségrégation familiale .....	191
B) Etude de l'interaction entre les sous-unités du BBSome .....	191
C) L'étude du fonctionnement de la voie SHH .....	193
III. La discussion .....	194
IV. La conclusion et les perspectives.....	199
V. Les données supplémentaires.....	200

## IV. Partie 2 : La caractérisation phénotypique du nouveau modèle murin

### **Vps15<sup>R998Q</sup>..... 201**

#### **I. Le synopsis.....201**

1. Remise dans le contexte : ..... 201
2. La démarche scientifique..... 201
3. Le changement de fond génétique..... 202

#### **II. Les résultats .....204**

1. Présentation de la lignée ..... 204
2. Analyse portant sur les yeux : ..... 205
  - A) Etude fonctionnelle par ERG ..... 206
  - B) Etude histologique des yeux ..... 207
  - C) Etude de l'interactome de Vps15..... 208
2. Analyse portant sur le rein : ..... 210
  - A) Prélèvement rétro-orbital et analyses biochimiques ..... 210
  - B) Etude histologique du rein ..... 212
  - C) Etude de l'interactome de Vps15..... 214
3. Analyse portant sur les cellules des souris ..... 216
  - A) Marquage de LC3 et P62 avec/sans Chloroquine..... 216
4. Analyse portant sur le cerveau ..... 218
  - A) Etude préliminaire..... 218
    - Les 1<sup>ères</sup> conclusions ..... 219
  - B) Etude de la mémoire de travail ..... 220
  - C) Analyse de l'ultrastructure des régions hippocampiques ..... 222
  - D) Déterminer la nature des plaques..... 222
  - E) Etude de la protéine Tau ..... 225
    - 2<sup>ème</sup> conclusion partielle sur le phénotype cérébral ..... 225
5. Analyses exploratoires..... 226
  - A) Dosages protéiques cérébrales ..... 226
  - B) Etude protéique (*en cours*)..... 226
  - C) Etude de l'interactome de Vps15 (*en cours*) ..... 227
  - D) Analyses de l'ultrastructure du pancréas..... 228
    - Dernière conclusion sur la 3<sup>ème</sup> vague ..... 228

#### **III. La discussion et les perspectives .....229**

1. Est-ce un bon modèle d'étude des ciliopathies ? ..... 229
  - A) L'interactome de Vps15 en fonction des organes..... 229
  - B) Explications potentielles à l'absence de phénotype ..... 232



2.	Est-ce un modèle pour d'autres maladies plus communes ? .....	233
A)	Refaire et valider les conclusions .....	233
B)	Modèle d'étude pour la maladie d'Alzheimer .....	234
C)	Modèle d'étude pour les amyloïdoses .....	236
<b>IV.</b>	<b>Les conclusions.....</b>	<b>236</b>
 <b>V. Discussion générale.....</b>		 <b>237</b>
<b>I.</b>	<b>Le diagnostic génétique .....</b>	<b>237</b>
1.	Les nouveaux gènes .....	238
2.	Plusieurs gènes = plusieurs signes cliniques .....	239
3.	Les gènes connus .....	239
A)	De nouveaux mécanismes .....	239
B)	Qui s'occupe des VUS ? .....	240
C)	La difficulté d'obtenir du matériel biologique.....	240
D)	Le WGS .....	241
<b>II.</b>	<b>Les validations fonctionnelles .....</b>	<b>242</b>
1.	Les modèles cellulaires .....	242
A)	Lequel choisir ?.....	242
B)	Quelles conditions de culture ?.....	243
2.	Le modèle du poisson-zèbre.....	245
3.	Les modèles murins .....	245
A)	Elargir l'exploration phénotypique des organes .....	246
B)	Elargir l'exploration phénotypique d'autres modèles murins.....	246
4.	Le modèle KI <i>Vps15</i> <sup>R998Q/R998Q</sup> .....	247
 <b>VI. Conclusions et perspectives.....</b>		 <b>252</b>
 <b>VII. Références.....</b>		 <b>254</b>

Figure 1 : L’Historique des découvertes du cil..	2
Figure 2 : Les 2 types de cils et leurs localisations	4
Figure 3 : Le cil primaire et certaines voies associées.	6
Figure 4 : Les IFT.	7
Figure 5 : Le complexe octamérique du BBSome.....	8
Figure 6 : La voie SHH.....	10
Figure 7 : L’EVC zone.	11
Figure 8 : Le processus d’autophagie.....	14
Figure 9 : Les protéines BBS.	16
Figure 10 : Contribution de chaque gène dans le BBS.	17
Figure 11 : Les domaines protéiques de BBS1 et BBS5.	18
Figure 12 : Le complexe IFT27-IFT25.....	20
Figure 13 : Les complexes 1 et 2 de la PI3K de classe III.	22
Figure 14 : Schéma bilan tiré de Stoetzel et al. 2016.	23
Figure 15 : Les amyloïdoses.....	24
Figure 16 : Les voies non amyloïdogéniques et amyloïdogéniques	25
Figure 17 : Tau, phospho-Tau et DNF.....	26
Figure 18 : Autophagie et maladie de d’Alzheimer.....	27
Figure 19 : La recherche translationnelle.....	28
Figure 20 : Abaque des ciliopathies.....	30
Figure 21 : L’hétéogénéité génétique des ciliopathies.....	32
Figure 22 : L’apport des techniques NGS dans le diagnostic génétique.	33
Figure 23 : La filtration de variants.....	34
Figure 24 : Les conditions de culture des fibroblastes.	38
Figure 25 : Représentation schématique de l’œil et de la rétine.....	40
Figure 26 : Coupes transversales de reins.....	42
Figure 27 : Représentation schématique de cerveau.....	44
Figure 28 : Photo d’un explant.	54
Figure 29 : Génération des souris Vps15.....	63
Figure 30 : Electrorétinogramme.	65
Figure 31 : Illustration du prélèvement rétro-orbital.....	66
Figure 32 : Le dispositif du Y-maze	67
Figure 33 : Découpe d’un cerveau de souris.	68
Figure 34 : Découpe au microtome.....	69
Figure 35 : Logigramme.....	74

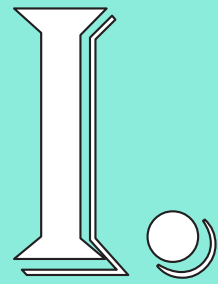
Figure 36 : Comparaison des résultats de PCR sur l'ADNc selon le tissu..	74
Figure 37 : La rétrotransposition dépendante de LINE-1.	88
Figure 38 : Les « soft clips ».	89
Figure 39 : Composition des SVA-F de référence.	90
Figure 40 : Logigramme.	91
Figure 41 : La séquence Sanger totale du SVA-F.	93
Figure 42 : Logigramme.	148
Figure 43 : Volcano plot.	149
Figure 44 : Expériences menées sur la voie SHH.	150
Figure 45 : Description de la famille et des mutations identifiées.	186
Figure 46 : Localisation des mutations.	186
Figure 47 : Logigramme.	187
Figure 48 : Détection des deux mutations par PCR et séquençage Sanger.	188
Figure 49 : Screening de la cohorte de patients BBS non résolus (LGM).	189
Figure 50 : Un seul allèle est exprimé.	190
Figure 51 : Les interactions entre BBS5 et les autres BBS.	192
Figure 52 : Résultats des expériences menées sur la voie SHH.	194
Figure 53 : Détection des mutations en fonction de la technologie NGS utilisée.	195
Figure 54 : Logigramme.	203
Figure 55 : Caractérisations de bases de la lignée.	205
Figure 56 : Résultats d'ERG.	206
Figure 57 : Etude histologique de la rétine.	207
Figure 58 : Résultats de spectrométrie de masse pour la rétine.	209
Figure 59 : Analyses d'enrichissements et validations d'interactions pour la rétine.	209
Figure 60 : Analyses biochimiques plasmatiques.	211
Figure 61 : Etude histologique du rein.	212
Figure 62 : IHC-P sur les coupes rénales de souris WT.	213
Figure 63 : Résultats de spectrométrie de masse pour le rein.	215
Figure 64 : Analyses d'enrichissements et validations d'interactions pour le rein.	215
Figure 65 : Travail portant sur les fibroblastes murins.	217
Figure 66 : Travaux réalisés en collaboration avec l'équipe de Y. Binnaz.	218
Figure 67 : Etude de la mémoire par Y-maze, Cohorte 1.	220
Figure 68 : Etude de la mémoire par Y-maze, Cohorte 2.	221
Figure 69 : Etude de la mémoire par Y-maze, Cohorte 3.	222
Figure 70 : Marquages protéiques sur des cryocoupes de cerveaux.	223

Figure 71 : Observation de l'hippocampe par MET.....	224
Figure 72 : Tau et ses sérines phosphorylées.....	225
Figure 73 : Observation du pancréas par MET. ....	226
Figure 74 : Dosages protéiques. ....	226
Figure 75 : Mise au point de Co-ip anti <i>Vps15</i> dans le cerveau.. ....	227
Figure 76 : L'activation de PKA. ....	231
Figure 77 : Molécules thérapeutiques anti-plaques.....	235

## *Liste des tableaux*

---

Tableau 1 : Présentation de certains modèles murins BBS.....	46
Tableau 2 : Listes des amorces .....	49
Tableau 3 : Liste des anticorps (AC).....	51
Tableau 4 : Liste des buffers.....	52
Tableau 5 : Liste des nourritures pour souris.....	52
Tableau 6 : Liste des logiciels .....	53
Tableau 7 : Liste des tests Elisa .....	59
Tableau 8 : Listes des mutations de BBS5 répertoriées. ....	200
Tableau 9 : Principaux modèles murins de la maladie d'Alzheimer.....	248



# Introduction



## I. Les cils

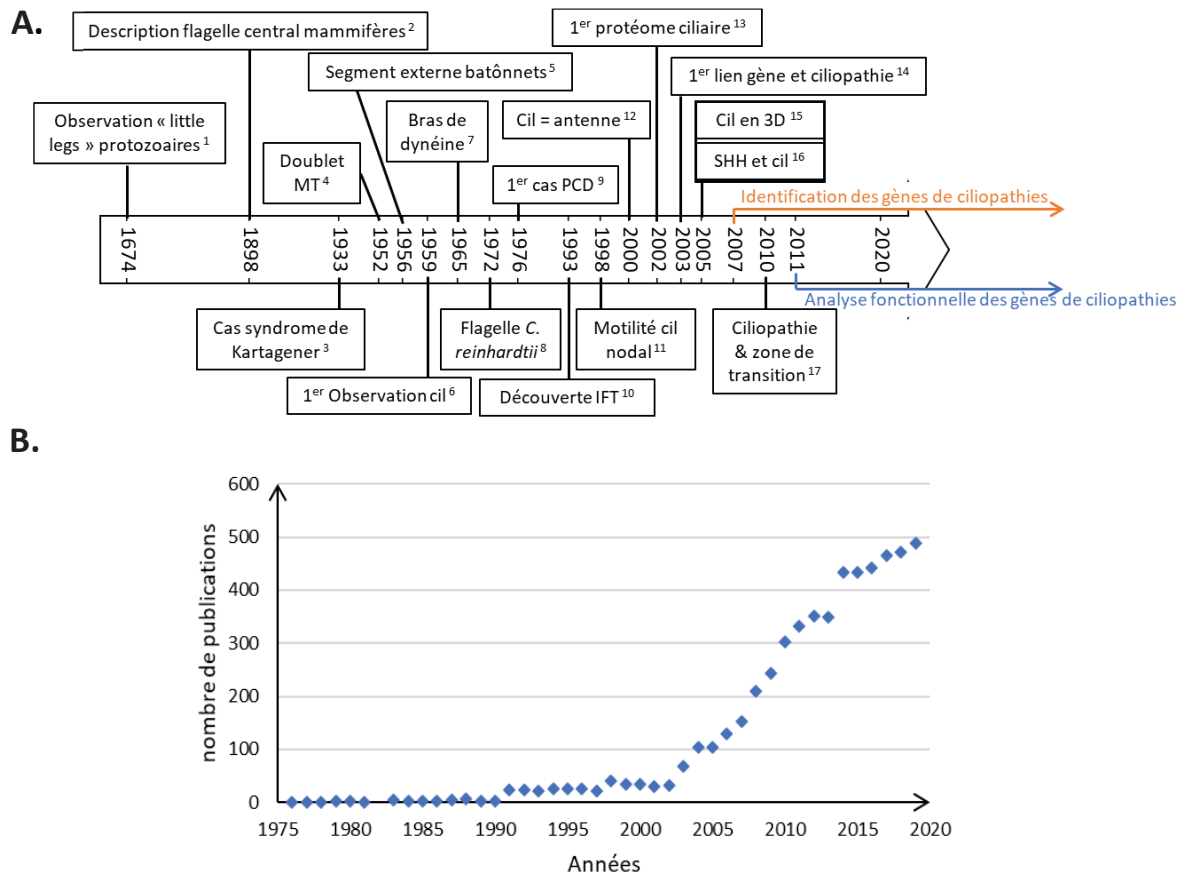
Il existe à la surface de la quasi-totalité des cellules de l'organisme une protubérance de la membrane plasmique appelée « cil ». Il existe 2 types de cil en fonction de leur capacité ou non de mouvement et de la composition de leur squelette interne appelé axonème, composé de 9 doublets périphériques de microtubule (MT). C'est la présence d'un doublet central de MT qui permet de distinguer le cil motile (9+2) du cil primaire où le doublet central est absent (9+0) (Figure 2A).

### 1. La chronologie des découvertes

La première description du cil a été réalisée en 1674 par le hollandais Antoni Van Leeuwenhoek lors de son observation d'échantillons d'eau de pluie où il a mis en évidence des protozoaires pourvus d'un cil leur permettant de se déplacer (Leeuwenhoek 1677). Presque 300 ans après la première observation, dans les années 1950, le perfectionnement de la microscopie électronique (ME) et des méthodes de fixation ont permis des avancées importantes pour l'étude du cil. Après la première observation du doublet de microtubules en 1952 par Manton, Bjorn Afzelius observe 7 ans plus tard, le prolongement ciliaire formé à partir du corps basal (B. Afzelius 1959). Le premier lien entre le cil motile et une maladie date de 1976 avec Afzelius qui découvre l'absence de bras de dynéine au niveau du flagelle des spermatozoïdes d'hommes infertiles (B. A. Afzelius 1976) et explique finalement les premiers cas de dyskinésies ciliaires primitives (PCD).

Dans le cas du cil primaire plus particulièrement, sa première description date de 1898 par Zimmermann qui décrit un flagelle central, présent en unique exemplaire à la surface de cellules qui se développe à partir d'un des 2 centrosomes. Il se distingue de la multitude de cils présents à la surface des cellules dites « multiciliées ». Cependant, l'intérêt porté à cette organelle reste très limité aux observations microscopiques précédemment mentionnées. D'ailleurs, l'état de l'art en 1960 le décrit comme une structure « 9+0 », non motile, associée à une paire de centrosomes, distribuée à la surface de la plupart des cellules des vertébrés (Bloodgood 2009). Le terme « cil primaire » est utilisé pour la première fois par Sergei Sorokin en 1968, il remplace le terme « cil rudimentaire » mais le rôle de cette organelle reste très peu étudié. Il faudra attendre la découverte de la machinerie de transport intraflagellaire (IFT) en 1992 par Kozminski pour voir une évolution majeure concernant l'étude du cil primaire (Figure 1A). Bien qu'il ait été longtemps considéré comme un vestige cellulaire sans rôle essentiel, il apparaît alors comme un compartiment cellulaire à part entière nécessaire au fonctionnement normal des cellules. D'ailleurs, depuis les années 2000, de nombreuses équipes cherchent à déterminer le rôle du cil primaire, à élucider les causes moléculaires des maladies causées par des atteintes des cils. Ces maladies sont regroupées sous le terme de « ciliopathies ». Les premiers liens entre gènes et ciliopathies datent de 2003 : le gène *INVS* est rapporté comme responsable de néphronophtise par Otto et al. 2003 (Figure 1).





**Figure 1 : L'historique des découvertes sur le cil.** **A.** Frise chronologique représentant les date marquantes concernant le champ de recherche portant sur les cils (motiles et primaires) (inspirée de celle de Brown et Witman 2014) **B.** Graphique représentant la cinétique des publications abordant le cil primaire entre 1975 et 2019. Informations tirées de la base de données Web of Science pour les mots clefs « primary cilium » et « primary cilia » (inspiré de Bloodgood 2009). <sup>1</sup> Leeuwenhoek 1677, <sup>2</sup> Zimmermann 1898, <sup>3</sup> Kartagener 1933, <sup>4</sup> Manton 1952, <sup>5</sup> De Roberti 1956, <sup>6</sup> B. Afzelius 1959, <sup>7</sup> Gibbons et Rowe 1965, <sup>8</sup> Witman et al. 1972, <sup>9</sup> B. A. Afzelius 1976, <sup>10</sup> Kozminski et al. 1993, <sup>11</sup> Nonaka et al. 1998, <sup>12</sup> Pazour et al. 2000, <sup>13</sup> Ostrowski et al. 2002, <sup>14</sup> E. A. Otto et al. 2003, <sup>15</sup> Huangfu et Anderson 2005, <sup>16</sup> Nicastro, McIntosh, et Baumeister 2005, <sup>17</sup> Craige et al. 2010.

## 2. Les 2 types de cils

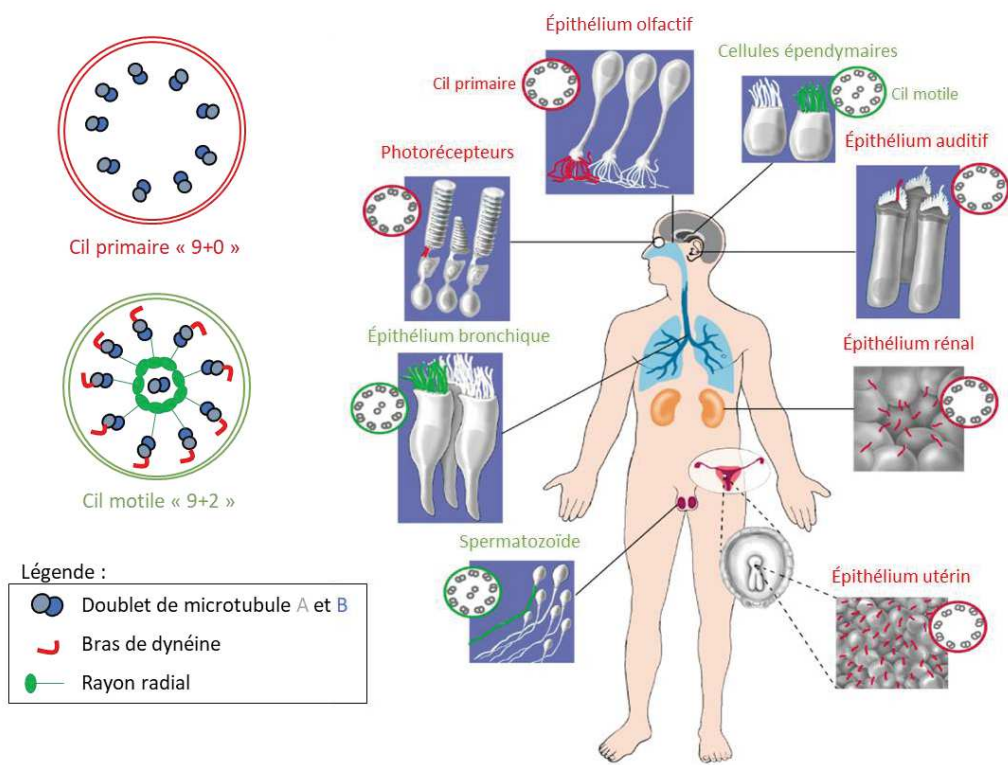
### A) Le cil motile « 9+2 »

Les cils motiles peuvent être unique à surface cellulaire dans le cas des cellules mono-ciliées ou multiples dans le cas de cellules multiciliées qui tapissent la lumière des organes (Mitchison et Valente 2017a). Absents du cil primaire, les bras de dynéines et les rayons radiaux sont des structures ciliaires importantes pour le battement du cil motile qui permettent la mise en place d'un flux. Par exemple, les cellules épithéliales de la trachée sont pourvues à leur pôle apical de cils motiles impliqués dans la mise en œuvre de l'ascenseur muco-cilié : les battements ciliaires participent à la remontée du mucus vers le pharynx. Les spermatozoïdes sont constitués d'un flagelle (considéré comme un cil motile spécialisé) dont les battements sont nécessaires à leur mobilité dans les voies génitales féminines pour aboutir à une fécondation (Figure 2B).

B) Le cil primaire « 9+0 »

Le cil primaire est toujours présent, en un seul exemplaire, à la surface de la quasi-totalité des cellules de l'organisme à un moment de leur vie. L'axonème se forme à partir du corps basal lui-même dérivé du centriole mère nécessaire pour la mise en place du fuseau mitotique lors de la division cellulaire : la ciliogénèse se déroule en phase de quiescence, en phase G0 du cycle cellulaire. *In vitro*, le cil primaire se forme, à la surface de cellules cultivées dans un milieu déprivé en sérum (-SVF).

Le cil primaire forme une sorte « d'antenne » ayant des rôles de mécanosenseur et de chémosenseur. La membrane ciliaire se distingue de la membrane plasmique par sa forte concentration en récepteurs captant les signaux extra-cellulaires. La cascade de signalisation générée recrute des protéines ciliaires qui vont transmettre ces messages aux autres compartiments cellulaires afin de générer finalement une réponse cellulaire (activation transcriptionnelle, prolifération...). Le cil se trouve au centre de nombreuses voies de signalisations cellulaires telles que SHH (Sonic Hedgehog), WNT, BMP (Fry, Leaper, et Bayliss 2014). C'est donc un acteur majeur dans la transmission des signaux et de la communication cellulaire. Le cil primaire est impliqué dans la différenciation des cellules dans le cas des pré-adipocytes, par exemple (Marion, Mockel, et al. 2012). Il permet également de maintenir les fonctions normales de cellules matures post-mitotiques. Par exemple, au niveau du rein, même si le cil primaire n'est pas capable de battement, il se courbe sous l'effet du flux urinaire présent au niveau de la lumière des néphrons. Cette action génère des vagues calciques au niveau cytoplasmique qui participent à maintenir un fonctionnement rénal normal. Dans d'autres organes, la morphologie du cil primaire évolue pour s'adapter au mieux à la fonction qui lui incombe. Par exemple, le segment externe des photorécepteurs (OS) correspond à un cil primaire spécialisé. Il est composé de nombreux repliements de la membrane plasmique où les récepteurs sensibles à la lumière s'accumulent et d'un cil connecteur permettant de relier et d'acheminer les protéines entre le segment interne (IS) et le segment externe (OS).



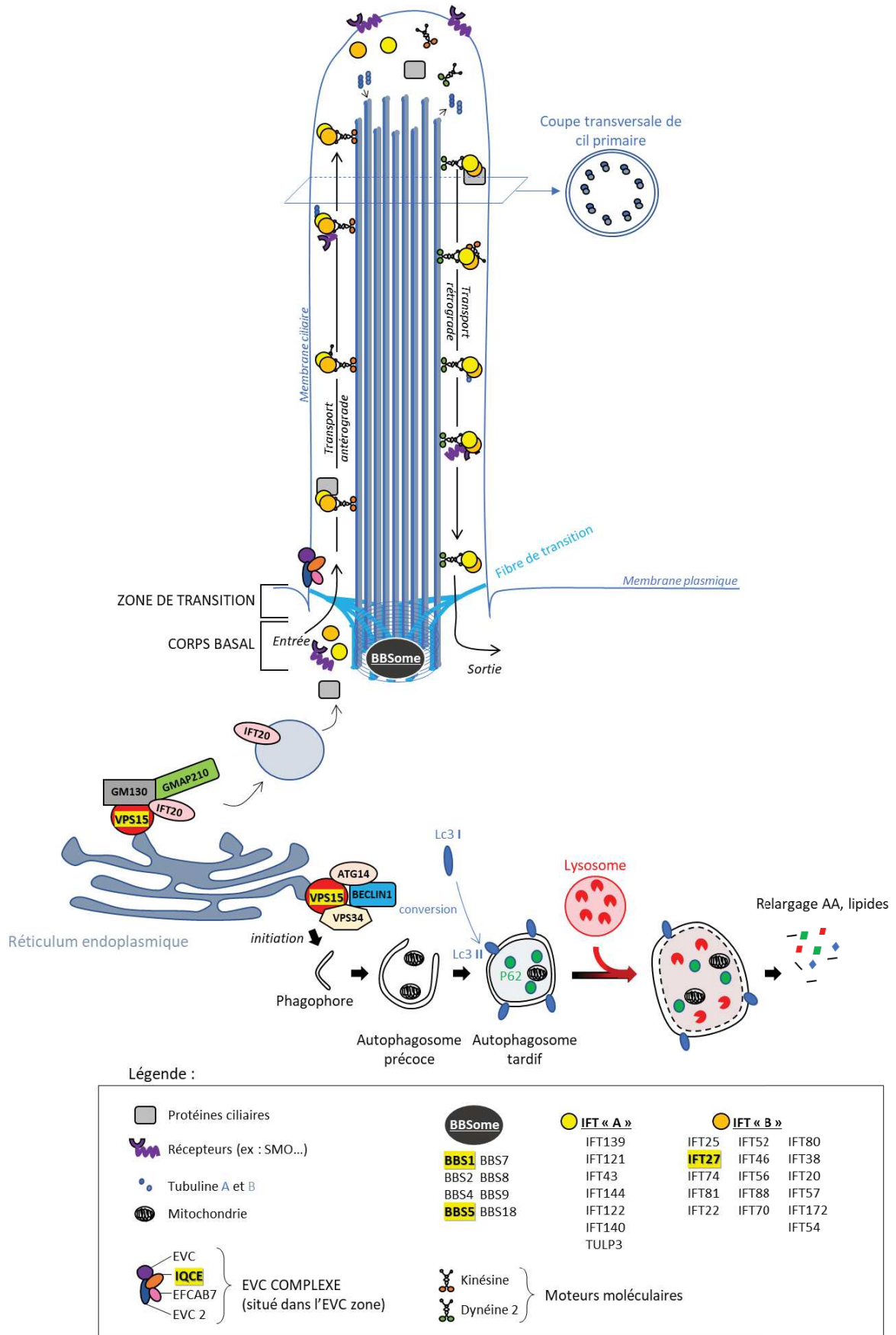
**Figure 2 : Les 2 types de cils et leurs localisations** A. Représentation schématique de coupes transversales de cils primaires et motiles. B. Localisation du type de cil dans les différents organes du corps humain. Le cil primaire est représenté en rouge et le cil motile en vert.

## II. Le fonctionnement du cil primaire

Le cil primaire est dépendant de la machinerie cellulaire pour la synthèse des 1500 protéines impliquées dans sa formation, sa maintenance et son fonctionnement (Whewey, Nazlamova, et Hancock 2018). La synthèse protéique se réalise dans le réticulum endoplasmique (RE). Les modifications post traductionnelles se déroulent au niveau de l'appareil de Golgi. A ce niveau, les protéines sont intégrées dans des vésicules de transport jusqu'à la base du cil. Comme pour l'adressage de protéines aux autres compartiments cellulaires (Figure 3), il existe un peptide signal présent dans la séquence des protéines ciliaires appelé CTS (ciliary targeting signal). Pour l'instant, des dizaines de CTS ont été rapportées en fonction des protéines considérées et aucune séquence consensus n'a pu être identifiée (Madugula et Lu 2016). Les cargos protéiques adressés au cil peuvent suivre 2 voies différentes : soit les protéines transportées dans les vésicules fusionnent directement avec la membrane périciliaire, soit les vésicules de transport fusionnent avec la membrane plasmique à proximité de la base du cil. Par exemple, dans le cas de la voie Sonic hedgehog (SHH), le récepteur SMO (Smoothed) attend la survenue d'un signal extracellulaire d'activation de la signalisation sous-jacente qui provoquera le transport latéral de SMO depuis la membrane plasmique vers l'intérieur du cil via la barrière de diffusion. Il s'agit de la portion membranaire de la zone de transition où le BBSome (impliqué dans le syndrome de Bardet Biedl BBS), le complexe MKS (associé au syndrome de Meckel-Gruber) et le complexe NPHP (associé à la néphronophtose) sont localisés. Les protéines du transport intraflagellaire (IFT) et ces 3 complexes sont des « assistants » impliqués dans la sélection et l'entrée des protéines ciliaires dans le cil, à travers la barrière de diffusion. En cas d'anomalie au niveau de cette barrière, la croissance du cil est inhibée, la réponse cellulaire médiée par des signaux extracellulaires est réduite, le transport protéique est perturbé et la composition lipidique de la membrane ciliaire est anormale (Long et Huang 2020).

### 1. Le transport intraflagellaire (IFT)

Les protéines IFT ont été mises en évidence pour la première fois chez *Chlamydomonas reinhardtii* avec l'observation de « trains d'IFT » par Kozminski et al. en 1993. Elles forment un véritable système de transport grâce à leur interaction avec les moteurs moléculaires que sont les dynéines et les kinésines. L'ensemble peut être assimilé à un transport ferroviaire permettant d'acheminer les protéines ciliaires depuis la base du cil vers l'apex, au cours de la ciliogénèse (antérograde) et dans le sens inverse lors de sa résorption (rétrograde). En effet, la circulation des protéines (passagers) est opérée, de manière bidirectionnelle, par les protéines du transport intraflagellaire (sièges) qui interagissent avec les moteurs moléculaires (wagons) qui vont progresser le long des MT (rail) (Figure 3). Les protéines IFT ont été réparties en 2 complexes : IFT-A et IFT-B selon l'orientation du transport.



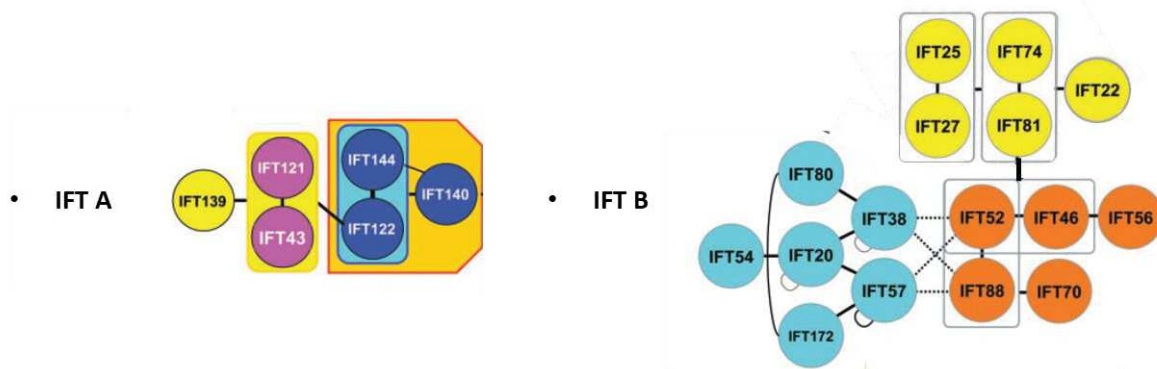
**Figure 3:** Le cil primaire et certaines voies associées. Représentation schématique du cil primaire avec le positionnement des différentes protéines des voies de signalisation. Les protéines soulignées en **jaune** feront l'objet d'études au cours de cette thèse.

A) Le complexe A

Il se compose de 6 sous unités (IFT139, 121, 43, 144, 122, 140), divisées en 2 sous modules associant IFT122-140-144 d'une part et IFT43-121-139 d'autre part (Figure 4). Ils interagissent avec la dynéine-2 et vont participer au transport rétrograde de protéines. Ce système permet l'évacuation et le recyclage des protéines hors du cil, lors de la résorption du cil (évacuation des dimères de tubuline  $\alpha\beta$ ) ou de l'arrêt de la stimulation d'une voie de signalisation spécifique (recyclage des récepteurs hors du cil) (Nakayama et Katoh 2018).

B) Le complexe B

Il regroupe 16 sous unités, subdivisées en protéines « périphériques » ou « centrales » (Figure 4) qui interagissent avec la kynésine-2 et participent au transport antérograde de cargos protéiques. Lors de la ciliogénèse, des interactions entre les dimères de tubuline  $\alpha\beta$  et ces protéines IFT ont été mises en évidence montrant le rôle de ce transport dans l'allongement de l'axonème (Nakayama et Katoh 2018). Les composants ciliaires sont acheminés depuis le cytoplasme ou l'appareil de Golgi, vers la base du cil grâce à la protéine IFT20. C'est la seule protéine IFT du complexe B qui a été localisée au niveau du cil et au niveau de l'appareil de Golgi où elle interagit avec GMAP210 (Long et Huang 2020) (Figure 3).



**Figure 4 : Les IFT.** Représentation des 2 complexes IFT-A et IFT-B. Les protéines périphériques sont représentées en **bleu** et les centrales en **jaune** et **orange** (inspirée de Nakayama et Katoh 2018).

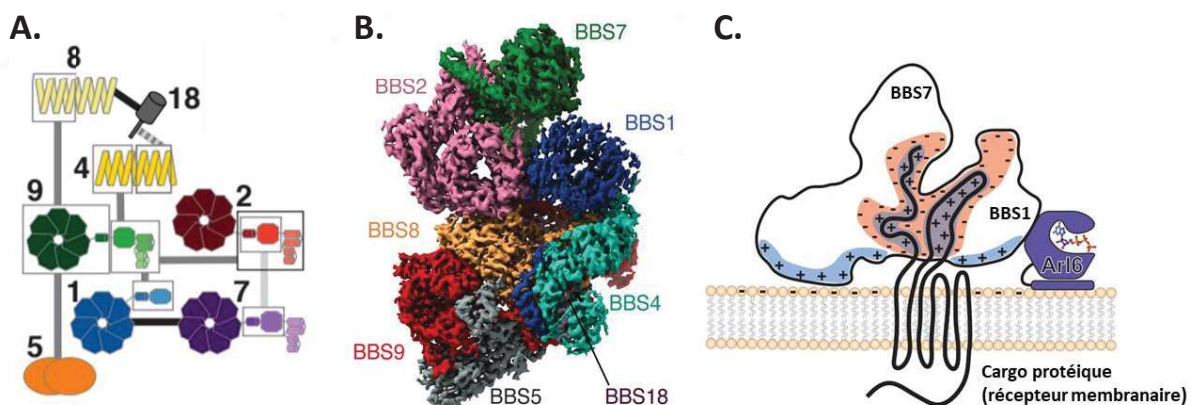
## 2. Le BBSome

Le BBSome est un complexe octamérique de 490 kDa composé des protéines BBS 1, 2, 4, 5, 7, 8, 9, 18 : chaque sous unité est présente en un seul exemplaire. Il est localisé à la base du cil et fait partie des complexes protéiques essentiels au bon fonctionnement du cil primaire. Il a été découvert en 2007 par Nachury et al. 2007a. Plus récemment sa structure 3D a été résolue et les interactions entre les sous-unités ont été précisées (Chou et al. 2019a; S. K. Singh et al. 2020a). A ce jour, son rôle n'est pas totalement élucidé. Il s'agit d'un complexe qui n'est pas nécessaire à la ciliogénèse (Wingfield, Lechtreck, et Lorentzen 2018). En revanche, il est essentiel à la signalisation ciliaire car il permet la sélection des cargos de protéines transmembranaires transportés au cil. C'est un régulateur clef du



protéome ciliaire (S. K. Singh et al. 2020a). Il est impliqué dans l'assemblage et la stabilisation du complexe IFT : il interagit directement avec IFT38 (Nozaki et al. 2019). L'assemblage du BBSome est un processus ordonné qui requiert 2 complexes : le BBS-chaperonines-like (BBS6-BBS10-BBS12) et le complexe CCT-TRiC (Klink et al. 2020a) notamment dans la première étape où la sous unité BBS7 est stabilisée. Le reste de l'assemblage ne nécessite pas le complexe de chaperonnes et implique seulement les interactions protéine-protéine qui s'établissent entre les sous unités du BBSome (Álvarez-Satta, Castro-Sánchez, et Valverde 2017).

Le BBSome se compose d'un noyau central composé de BBS1, 2, 7, 8. Le sous ensemble BBS8, 4 et 18 est en interaction avec le noyau via la liaison entre BBS8 et BBS9. BBS5 est plus à la périphérie et établit une liaison avec BBS9 (Figure 5). L'entrée et l'activation du BBSome au niveau du cil primaire sont directement régulées par une autre protéine BBS : ARL6/BBS3 (une petite GTPase). Lorsque l'interaction n'est pas établie : le BBSome est dans un état auto-inhibé. Quand l'interaction est établie avec ARL6 via la sous unité BBS1, il y a un changement de conformation du complexe, un attachement à la membrane plasmique et une activation fonctionnelle permettant le passage des protéines au niveau de la zone de transition (Nozaki et al. 2019). L'interaction entre le BBSome et la membrane plasmique se localise au niveau d'une zone chargée positivement du complexe. A l'inverse, l'interaction entre le BBSome et les cargos protéiques se situe au niveau d'une zone chargée négativement (Klink et al. 2020a) (Figure 5C). A l'heure actuelle, 89 mutations au niveau d'une des 8 sous unités du BBSome semblent altérer la structure globale du BBSome. Finalement, un défaut de l'une des 11 protéines BBS mentionnées précédemment conduit à la déstabilisation du BBSome et à un dysfonctionnement ciliaire (Chou et al. 2019a).



**Figure 5 :** Le complexe octamérique du BBSome. **A.** Représentation schématique des interactions entre les 8 sous unités (Chou et al. 2019a). **B.** Représentation 3D. **C.** L'interaction du BBSome avec ARL6/BBS3, les cargos protéiques (ex : récepteur transmembranaire) et la membrane (Klink et al. 2020).

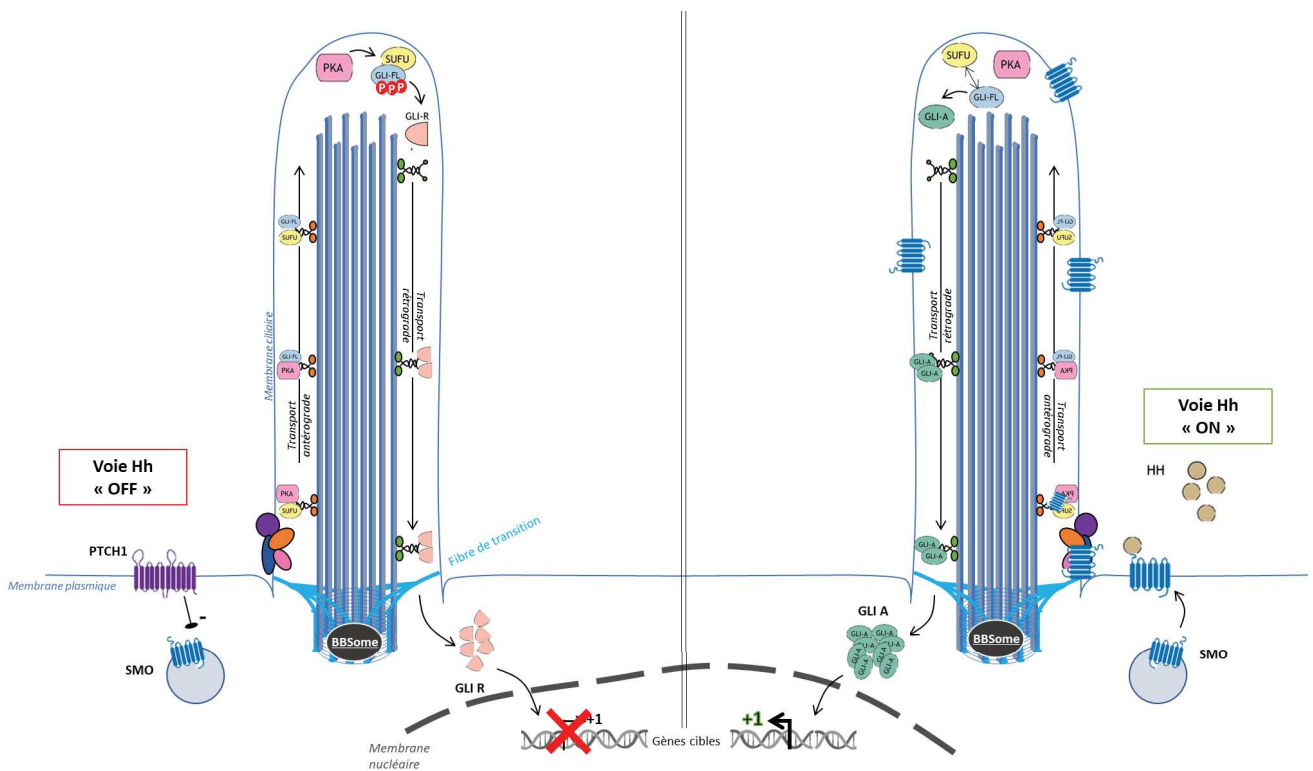
### 3. Les voies de signalisations médiées par le cil

La membrane ciliaire est une zone de concentration de récepteurs qui captent les signaux extracellulaires, provoque l'activation d'une cascade de signalisation intracellulaire aboutissant à la génération d'une réponse de la part de la cellule telles que la prolifération ou la différenciation cellulaire... Ainsi, le cil est impliqué dans l'activation de nombreuses voies de signalisations majeures telles que la voie WNT, BMP, SHH qui sont indispensables à la mise en place des organes durant l'embryogénèse et au maintien de leur fonctionnement normal une fois formés (Whewey, Nazlamova, et Hancock 2018).

#### A) La voie de signalisation Sonic-Hedgehog (SHH)

Il s'agit d'une voie de signalisation majeure du développement embryonnaire qui permet la différenciation et la mise en place d'une multitude d'organes tels que le cerveau et les membres. Même si elle a été décrite pour la première fois dans des cellules non ciliées de *Drosophila melanogaster* (Riddle et al. 1993), tous les acteurs primordiaux de la voie sont concentrés au niveau du cil primaire qui a d'ailleurs été appelé « la machine de transduction du signal Hh » (Goetz, Ocbina, et Anderson 2009). Les protéines qui peuvent activer la voie Hedgehog (Hh) sont des facteurs de signalisation paracrine qui vont permettre la communication entre les cellules. Parmi ceux-ci, Sonic (SHH) est le plus étudié car son expression est ubiquitaire, contrairement à Desert (DHH) dont l'expression est limitée aux testicules et à Indian (IHH) qui est exprimé dans le cartilage et le tube digestif. SHH est un morphogène, c'est à dire une molécule sécrétée par un centre organisateur et dispersée en gradient pour activer des programmes d'expressions géniques précis. Il détermine le destin des progéniteurs environnants en fonction de la concentration de SHH reçue (Kong, Siebold, et Rohatgi 2019). La voie SHH est très importante pour la formation des membres et une dérégulation peut être responsable d'anomalies telle que la polydactylie. Dans ce cas, SHH est produit au niveau de la ZPA (zone polarizing activity) et va agir comme un morphogène pour établir le nombre et l'identité des doigts selon l'axe antéro-postérieur (Petit et al. 2016). La voie SHH est très largement étudiée et le rôle des différents acteurs dans la cascade doit encore être déterminé. La Figure 6 montre que lorsque le ligand SHH ou un agoniste de la voie tel que le SAG (Smo AGonist) est absent de l'environnement proche de la cellule, la voie est « OFF ». Dans ce cas, le récepteur à 12 passages transmembranaires PTCH1 (Patched), inhibe le récepteur à 7 passages transmembranaires SMO (Smoothened) qui reste inactif. Les 2 composants inhibiteurs SUFU (suppressor of fused) et PKA (protéine kinase A) sont recrutés et provoquent la protéolyse partielle de GLI FL (Glioma-associated oncogene homolog, full length) par phosphorylation. Les formes résultantes sont des facteurs de transcription GLI R (répresseurs), notamment GLI3, qui migrent vers le noyau et répriment l'activité transcriptionnelle des gènes cibles de la voie.



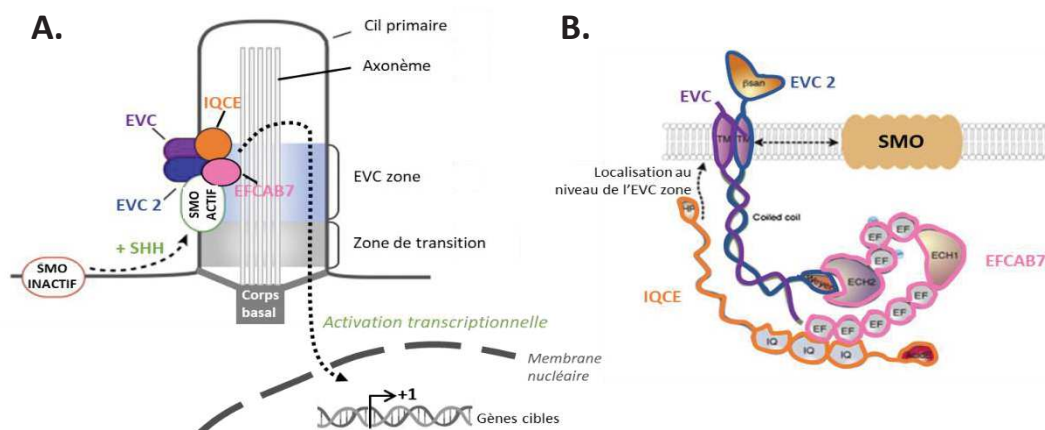


**Figure 6 : La voie SHH.** Représentation schématique des différents acteurs de la voie Hh quand elle est inactive « OFF » en comparant à l'état activé « ON » par un agoniste (SHH ou SAG).

A l'inverse, en présence de SHH ou d'un agoniste (SAG), la voie est « ON ». Dans ce cas, l'inhibition du récepteur SMO par le récepteur PTCH1 est levée. SMO migre alors à l'intérieur du cil (au niveau de la zone de transition), interagit avec EVC2 au niveau de l'EVC zone et est transporté le long de l'axonème grâce aux protéines IFT. Lorsque la voie est « ON », les 2 régulateurs négatifs sont inhibés et les GLI FL ne sont plus séquestrés par SUFU. Dans ce cas, des GLI A (activateurs) sont libérés, notamment GLI2 qui migre depuis le cytoplasme vers le noyau : on parle de translocation nucléaire. GLI2 est un facteur de transcription qui se fixe sur des régions régulatrices du génome et favorise l'expression de gènes cibles de la voie (Kong, Siebold, et Rohatgi 2019). Il est possible de déterminer le niveau d'activation de la voie SHH à plusieurs niveaux : en analysant le niveau transcriptionnel des gènes cibles de la voie (par qRTPCR), en marquant le récepteur SMO au niveau du cil (par immunofluorescence IF) ou en établissant le ratio de GLI2 (cytoplasmique) vs GLI2 (nucléaire) par Western blot WB (Maiti et al. 2017).

➤ L'EVC zone (Ellis Van Creveld)

Il s'agit d'une zone située à la base du cil où les protéines EVC et EVC2 sont en interaction avec 2 autres protéines : EFCAB7 (EF-Hand Calcium Binding Domain 7) et IQCE (IQ domain containing protéin E). L'ensemble a été rapporté comme nécessaire pour la régulation positive de la voie SHH. En effet, EVC2 établit une interaction avec le récepteur SMO lors de l'activation de la voie et conditionne la suite de la cascade de signalisation (Figure 7). Le complexe EFCAB7-IQCE est nécessaire pour l'ancrage et la localisation de EVC2-EVC au niveau de cette zone. Une mauvaise localisation de ce complexe cause un problème d'activation lié à GLI2 en aval de la voie (Pusapati et al. 2014).



**Figure 7 : L'EVC zone.** A. Représentation schématique de la localisation du complexe au niveau de l'EVC zone et de son rôle dans la régulation de la voie SHH (Dorn, Hughes, et Rohatgi 2012). B. Les domaines protéiques de chacun des 4 acteurs et de leurs interactions (adapté de Pusapati et al. 2014)

Depuis ces dernières années, plusieurs articles ont rapporté une relation réciproque entre le fonctionnement ciliaire et le processus d'autophagie. Pampliega et al. 2013 a montré que l'activation de la voie SHH stimule le processus d'autophagie en aval. Il existe un recrutement des protéines nécessaires à l'autophagie à la base du cil dépendant de l'IFT : le cil est un lieu de formation d'autophagosomes. Ces résultats suggèrent ainsi un programme d'autophagie médié par le cil.

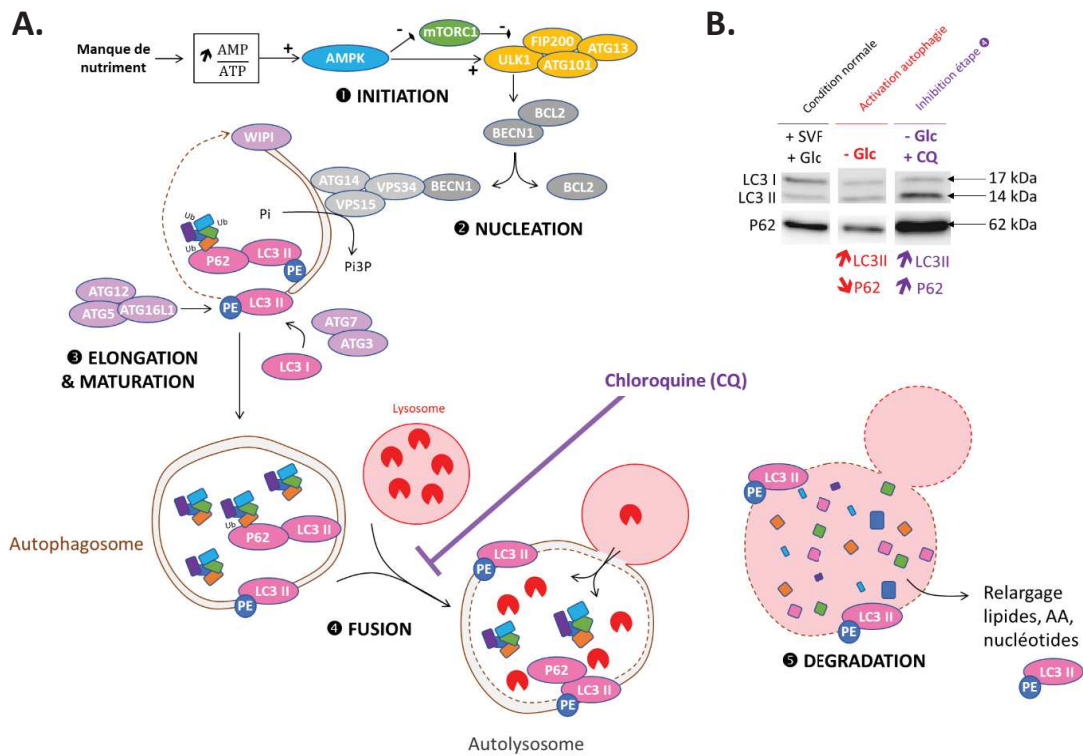
B) L'autophagie

Il s'agit d'un processus ubiquitaire très conservé au cours de l'évolution, depuis les levures jusqu'aux mammifères (Hurley et Young 2017). Ce processus permet de dégrader des protéines et des organelles en grande quantité. Cela permet de maintenir un fonctionnement cellulaire normal (homéostasie physiologique) en fournissant des lipides, des AA, des nucléotides. Ce processus basal est largement favorisé quand la cellule fait face à un environnement moins propice, appauvri en nutriments. (Hurley et Young 2017). Il existe 3 types d'autophagie : la microautophagie, l'autophagie médiée par des chaperonnes (CMA) et la macroautophagie. La première permet la dégradation de protéines par l'invagination ou la protrusion de la membrane lysosomale. La deuxième permet la dégradation de protéines très précises porteuses d'un pentapeptide : il est reconnu par des protéines chaperonnes qui vont permettre l'intégration des protéines captées directement dans le lysosome et leur dégradation. La dernière correspond au processus d'autophagie communément admis que je vais aborder par la suite. Les cargos protéiques sont séquestrés dans un autophagosome avant la fusion et la dégradation par les lysosomes (Parzych et Klionsky 2014).

Comme le résume la Figure 8A, ce processus peut être provoqué par l'inhibition de la voie mTORC1 (Target Of Rapamycin Complex 1) qui est une voie régulant la croissance et le métabolisme cellulaire. Elle est normalement activée en présence de facteurs de croissance. La voie mTORC1 est inhibée par une voie antagoniste, celle de l'AMPK (AMP activated protein kinase) dont l'activation est provoquée par l'augmentation de la quantité d'AMP (adénosine monophosphate) cytoplasmique. L'activation de l'AMPK provoque l'activation du complexe comprenant la kinase ULK1 : cette étape d'initiation (1) marque le début du processus d'autophagie suivi de 4 autres étapes, chacune recourant à de multiples protéines dont les ATG (AuTophagy related Gene). La nucléation (2) regroupe la translocation du complexe ULK1 (Unc-51 autophagy activating kinase) à l'endroit où l'autophagosome se forme (la membrane du RE ou la membrane plasmique) et la phosphorylation de Beclin-1. Cette modification provoque la dissociation du complexe Beclin-1-Bcl2 et permet l'interaction de Beclin-1 avec le complexe PI3K composé de VPS15 (PIK3R4), VPS34 (PIK3C3) et ATG14. L'ensemble contribue à la synthèse de phosphatidylinositol-3-phosphate (Pi3P), un constituant majeur des membranes dont celle de l'autophagosome, en cours d'élongation (3). Les autophagosomes se distinguent des vésicules de transport par la présence d'une double membrane. La protéine LC3-I est une protéine marqueur du

flux autophagique. Elle subit une modification post traductionnelle par l'ajout d'un phosphoéthanolamine (PE) permettant son ancrage dans la membrane de l'autophagosome. La seconde protéine marqueur est SQSTM1/P62. Elle reconnaît des protéines ubiquitinylées et les intègre au cœur de l'autophagosome par interaction avec LC3-II qui agit comme un adaptateur. Lorsque le processus d'autophagie est activé, la quantité de LC3-II généré devient supérieure à celle de LC3-I, la quantité de P62 est également augmentée (Figure 8B). Les autophagosomes subissent une maturation puis fusionnent (4) avec les lysosomes contenant des enzymes hydrolytiques activées dans un environnement très acide. Dans ce cas, la membrane externe de l'autophagosome mature fusionne avec le lysosome : il y a formation d'un autolysosome au sein duquel il y a dégradation (5) du contenu et de la membrane interne par les hydrolases. Lors de la dégradation, la protéine P62 est également digérée contrairement à LC3-II qui est relarguée. Finalement les quantités des 2 protéines marqueurs évoluent de manière inversement proportionnelle. En déterminant les quantités des protéines LC3-II et P62, par WB par exemple, il est possible de mettre en évidence des cas où le processus d'autophagie est altéré et où contrairement à la situation normale, la quantité de P62 augmente (Zhang et al. 2013) (Figure 8B). Pour reproduire et étudier ces situations *in vitro*, il est possible de dépriver le milieu de culture des cellules en glucose (+SVF, -Glc) afin d'activer le processus d'autophagie. La Chloroquine (CQ) peut être ajoutée au milieu pour empêcher la fusion (4) entre le lysosome et l'autophagosome. Dans ces conditions, il n'y a pas de dégradation (5) et les autophagosomes s'accumulent dans le cytoplasme. Il est également possible de visualiser cette accumulation par marquage de ces protéines en réalisant une IF.

Le processus d'autophagie peut également être utilisé pour provoquer la mort cellulaire de type 2, dépendante de l'autophagie (Ávalos et al. 2017) par l'activation de la kinase DAPK1 (Death Associated protein kinase 1). Cette sérine-thréonine kinase est dépendante du calcium pour son activation et a été rapportée pour ses rôles dans la régulation de l'apoptose via l'activation d'un programme de mort cellulaire lié à la protéine p53 (Wu et al. 2020), dans la pathogénèse de la maladie d'Alzheimer via la phosphorylation de certaines sérines de la protéine Tau et dans le processus d'autophagie. Dans ce cas, DAPK1 agit de 2 manières sur le complexe PI3K : (1) via l'activation de PKD qui active VPS34 par phosphorylation, (2) via la phosphorylation directe de Beclin-1 qui permet sa dissociation de Bcl-2 et permet l'activation du processus d'autophagie comme précédemment expliqué (P. Singh, Ravanan, et Talwar 2016).



**Figure 8 : Le processus d'autophagie.** A. Représentation schématique des différentes protéines participant aux 5 étapes du processus d'autophagie. B. Exemple de résultats de WB attendus dans des conditions de cultures normales (+SVF +Glc), dans des conditions où l'autophagie est activée (+SVF -Glc) et en présence de Chloroquine (+SVF -Glc + CQ) (adaptée de Galluzzi et Green 2019; Parzych et Klionsky 2014). SVF : Sérum de veau fœtal, Glc : Glucose, CQ : Chloroquine.

Comme précédemment mentionné, le processus d'autophagie et le fonctionnement ciliaire sont liés par une relation réciproque. D'ailleurs, certaines protéines sont impliquées dans les 2 processus. Par exemple, la protéine INPP5-E est nécessaire à la fusion entre les autophagosomes et les lysosomes. Elle a aussi été localisée à la base du cil et des mutations dans ce gène ont été rapportées dans le cas d'une ciliopathie : le syndrome de Joubert (Morleo et Franco 2019).

### III. Les pathologies associées au fonctionnement du cil

Dans cette partie de l'introduction, je vais approfondir certaines connaissances qui seront nécessaires pour une meilleure compréhension du travail réalisé dans les différents projets présentés par la suite. Parmi les multiples ciliopathies existantes, je vais d'abord m'attarder plus précisément sur le Syndrome de Bardet-Biedl (BBS) et 3 des 24 gènes BBS (*BBS1*, *BBS5*, *BBS19/IFT27*) qui font l'objet de 3 des 5 projets. Je rapporterais également plus en détails les connaissances actuelles portant sur *IQCE* : le gène étudié dans le cadre du 4<sup>ème</sup> projet. Je remettrais dans le contexte les connaissances relatives à *VPS15* qui fait l'objet du dernier projet développé dans ce manuscrit : ces différents rôles seront présentés en détail. Enfin, il paraissait essentiel, face aux résultats inattendus obtenus pour ce projet, d'introduire certaines notions portant les maladies neurodégénératives, notamment les amyloïdoses et la maladie d'Alzheimer.

#### 1. Les ciliopathies

Les pathologies causées par une anomalie du cil consécutive à des mutations dans les gènes codants pour les protéines ciliaires sont regroupées sous le terme de « ciliopathies ». Il s'agit de maladies génétiques rares dont les prévalences peuvent varier de 1/1000 (pour les plus fréquentes comme les maladies polykystiques autosomales dominantes, PKD) à 1/1000000 (dans le cas des ciliopathies ultra-rares comme le syndrome d'Alström) (Grochowsky et Gunay-Aygun, s. d.). En 2017, Reiter et Leroux mentionne 187 gènes avérés et 241 gènes candidats pour un total de 37 ciliopathies rapportées. Il existe 2 types de ciliopathies en fonction du type de cil altéré.

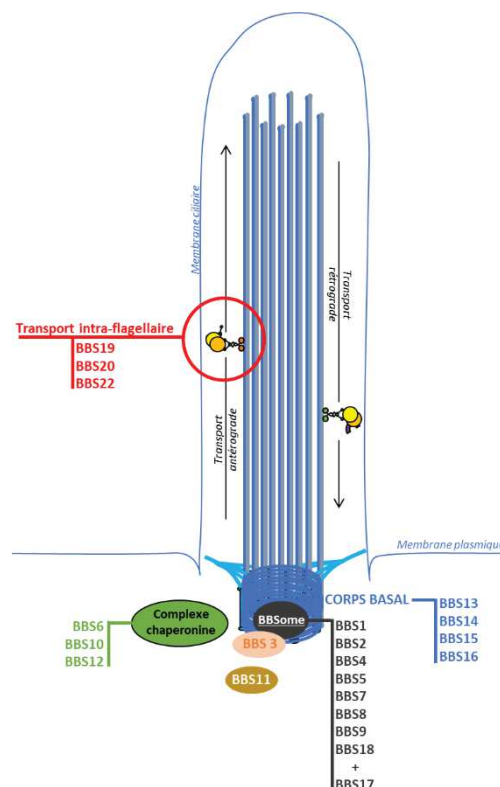
Les ciliopathies du cil motile ou dyskinesies ciliaires primitives (PCD) sont caractérisées par l'altération d'organes pour lesquels la motilité du cil est indispensable. Les patients présentent des *situs inversus*, des bronchites et sinusites chroniques, des problèmes de fertilité. A l'heure actuelle, 37 gènes ont été associés au PCD : la plupart d'entre eux code pour des protéines en lien avec les bras de dynéines.

Les autres ciliopathies sont des maladies du cil primaire dont la présence est quasi-ubiquitaire dans le corps humain, à un moment donné du développement. Un grand nombre d'organes sont concernés en cas d'anomalies ciliaires (la rétine, le rein, le cerveau...) (Figure 2). Ce type de ciliopathies est donc caractérisé par des tableaux cliniques hétérogènes : l'atteinte d'un organe isolé dans le cas d'une ciliopathie non syndromique (une rétinopathie pigmentaire isolée (RP) dans le cas de l'amaurose congénitale de Leber (LCA), une atteinte rénale isolée dans le cas des néphronophtises (NPHP)) à des atteintes multi-organes dans le cas d'une ciliopathie syndromique (le syndrome de Bardet-Biedl (BBS), le syndrome d'Alström (ALMS), le syndrome de Joubert (JBST)...) (Figure 20).

Le travail du laboratoire porte sur les ciliopathies du cil primaire.

## 2. Le syndrome de Bardet-Biedl (BBS)

Il s'agit d'une ciliopathie emblématique dont la prévalence varie entre 1/13500, dans le cas de populations consanguines ou de régions géographiquement isolées (Gouronc et al. 2020) et 1/160000, dans les populations originaires du nord de l'Europe (Waters et Beales 2011). Cette ciliopathie est caractérisée par un mode de transmission autosomique récessif. Elle a été décrite pour la première fois en 1860 par les docteurs Laurence et Moon (LAURENCE 1866). La polydactylie a été ajoutée à la première description clinique en 1920, par le docteur Bardet (BARDET 1920), puis en 1922 par le docteur Biedl (BIEDL 1922). Le diagnostic clinique se base sur la présence d'au moins 4 critères considérés comme « majeurs » (rétinopathie pigmentaire, obésité, polydactylie, anomalie des organes génitaux, anomalie rénale, lenteur d'idéation) ou bien par l'association de 3 critères majeurs et d'au moins 2 critères considérés comme « mineurs » (anomalie hépatique, diabète sucré, retard de croissance, retard psychomoteur, trouble de la parole, retard de développement, strabisme, troubles dentaires, perte auditive, malformations cardiaques) (P. L. Beales et al. 1999a). Parmi les 24 gènes identifiés comme responsables de cette ciliopathie : 8 gènes codent pour les sous unités du BBSome (*BBS1, 2, 4, 5, 7, 8, 9, 18*), 3 gènes codent pour le complexe « chaperonine-like » (*BBS 6, 10, 12*), *BBS3* code pour une GTPase, *BBS11* code pour une ubiquitine ligase, 4 gènes codent pour des protéines du corps basal (*BBS13, 14, 15, 16*), *BBS17* code pour une protéine en interaction avec le BBSome. 3 gènes (*BBS19, 20* et *22*) codent pour des protéines IFT et les rôles de *BBS21, BBS23, BBS24* doivent encore être déterminés (Figure 9).

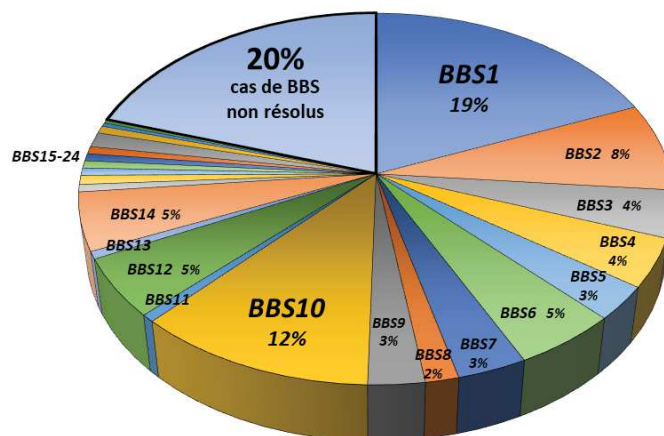


**Figure 9 : Les protéines BBS.** Représentation schématique d'un cil primaire avec la localisation de 21 des 24 protéines BBS (inspirée de Suspitsin et Imyanitov 2016).



A) Contribution de chaque gène BBS dans le syndrome

Comme le résume la Figure 10, 80% des cas de BBS ont été résolus par l'identification et la validation de mutations pathogènes parmi les 24 gènes BBS. *BBS1* et *BBS10* sont des contributeurs majeurs du BBS puisqu'ils sont mutés dans respectivement 19% et 12% des patients. Ils sont tous les 2 concernés par des mutations associées à un effet fondateur, à savoir l'apparition d'une mutation à un moment donné de l'évolution dans une population restreinte, pouvant se disperser à l'échelle de la population générale. 73% des patients mutés dans *BBS1* sont porteurs de la mutation c.1169T>G, p.M390R (Mykytyn et al. 2002a). Avant nos travaux portant sur le gène *BBS1*, certains patients BBS de la cohorte non résolue de Strasbourg étaient porteurs à l'état hétérozygote de la mutation M390R mais la 2<sup>ème</sup> mutation n'avait pas été mise en évidence (Muller et al. 2010a). Concernant *BBS10*, 48% des patients sont porteurs de la mutation c.271dupT, p.C91Lfs\*95 également associée à un effet fondateur. Le reste des gènes *BBS* sont mutés dans un nombre restreint de cas : ce sont des contributeurs mineurs au BBS. Toutefois, même si le BBS est la ciliopathie avec le taux de résolution le plus élevé, 20% des cas restent irrésolus (Mitchison et Valente 2017a).



**Figure 10 : Contribution de chaque gène dans le BBS.** Camembert récapitulant le pourcentage de patients mutés pour chaque gène *BBS* parmi les 80% de patients ayant un diagnostic génétique (adapté de Forsyth et Gunay-Aygun 1993, mise à jour le 23/07/2020).

B) Présentation de certains gènes BBS

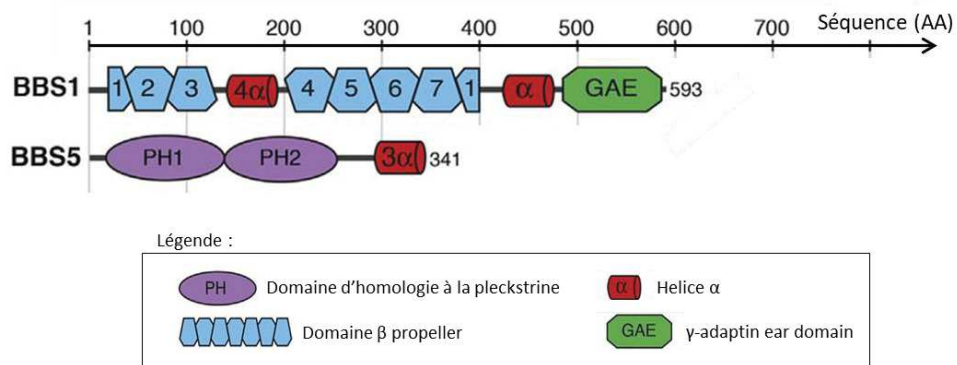
➤ *BBS1*

Le gène *BBS1* situé sur le chromosome 11 (11q13.2), se compose de 17 exons codants pour un transcrit de 3368 pb et une protéine de 593 AA (65 kDa) qui représente une sous unité majeure du BBSome. *BBS1* est le gène retrouvé muté dans la majorité des cas de BBS résolus (~20%). La protéine est en interaction directe avec la GTPase *BBS3/ARL6* au sein du BBSome et permet ainsi la localisation ciliaire, l'activation et l'attache membranaire du complexe (Figure 5). D'ailleurs, en cas de mutation de *BBS1*, le BBSome est mal localisé et non fonctionnel entraînant un dysfonctionnement ciliaire. Dans plus de



70% des cas de patients BBS1 rapportés, une mutation faux-sens (p.M390R) a été identifiée. Cette mutation concerne un AA très conservé au cours de l'évolution (Mykytyn et al. 2002a) situé au niveau du domaine  $\beta$  propeller (d'interaction protéine-protéine) (Davis et al. 2007) (Figure 11). L'AA M390 n'est pas impliqué directement dans l'interaction de BBS1 avec ARL6. Par contre, il a été montré que la mutation M390R provoque un mauvais repliement de la structure du  $\beta$  propeller qui semble adopter une structure en hélice  $\alpha$ . Ce changement a pour conséquence l'absence de formation du complexe BBS1-ARL6. Finalement, quand BBS1 est porteur de la M390R, le transport des protéines au cil est altéré car ARL6 ne recrute pas le BBSome à la membrane (Mourão et al. 2014).

Des souris Knock-in (KI) *BBS1*<sup>M390R/M390R</sup> porteuses de cette mutation présentent un phénotype similaire à celui des patients : atteinte rétinienne, infertilité masculine, obésité, troubles olfactifs, atteintes neuroanatomiques (Davis et al. 2007). La thèse de Yen, s. d. rapporte le modèle poisson zèbre *Bbs1*<sup>M390R</sup> présentant un phénotype typique de poisson-zèbres BBS provoqué par l'utilisation de morpholinos anti-Bbs1 réexprimant *Bbs1*<sup>M390R</sup>. Le phénotype est rétabli avec la réexpression d'un *Bbs1*<sup>WT</sup> confirmant la pathogénicité de la mutation M390R. La taille et le niveau d'expression de la protéine Bbs1 porteuse de la mutation est similaire à la protéine normale. Par contre, la protéine mutée est localisée au niveau du cytoplasme contrairement à la protéine Bbs1 normalement localisée au niveau du corps basal et du cytoplasme.



**Figure 11 :** Les domaines protéiques de BBS1 et BBS5. Représentation schématique des domaines protéiques pour les protéines BBS1 et BBS5 (adapté de Chou et al. 2019a).

### ➤ BBS5

Le gène *BBS5*, situé sur le chromosome 2 (2q31.1) contient 12 exons codants pour un transcrit de 3475 pb et une protéine de 341 AA (36 kDa). Il s'agit d'un contributeur mineur au BBS puisque seulement 2% des familles de diverses origines ethniques présentent des mutations du gène *BBS5*. Jusqu'à présent, plusieurs mutations dans *BBS5* ont été rapportées comme responsables du BBS et sont référencées dans le Tableau 8.

BBS5 est une sous-unité périphérique du BBSome faisant partie de la base large du BBSome en interaction directe avec BBS9 (Figure 5). Deux groupes d'auteurs ont montré, en 1993, la présence de

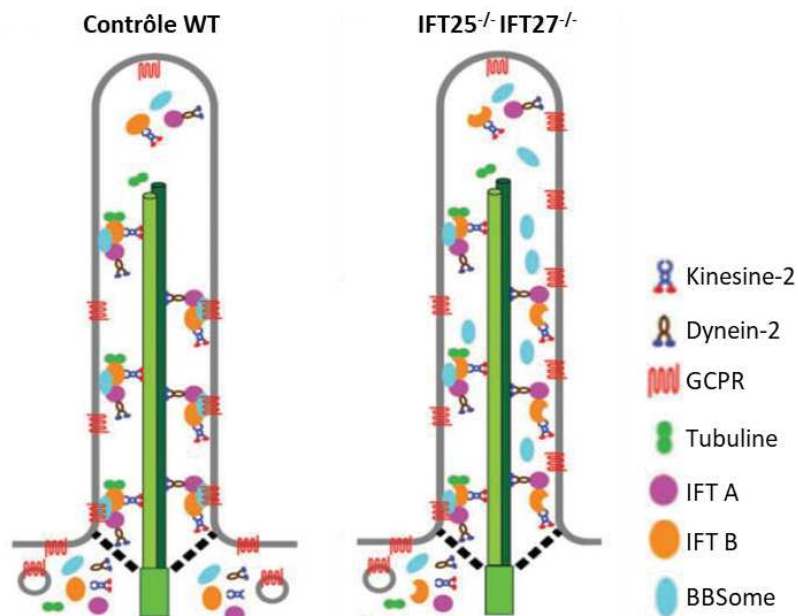
motifs analogues à ceux de la pleckstrine (segments homologues d'environ 100 AA chacun, dénommés domaines homologues à la pleckstrine (PH)) dans un nombre important de protéines dont BBS5, n'ayant apparemment aucun lien fonctionnel commun (Haslam, Koide, et Hemmings 1993 ; Mayer et al. 1993) (Figure 11). L'extension de ces recherches a conduit à retrouver des domaines PH dans plus de 90 protéines (Saraste and Hyvönen, 1995) et la liste n'est probablement pas close. Au sein des 2 domaines d'homologie à la pleckstrine (PH1 et 2) de BBS5, des domaines DM16 ont été identifiés. Leur fonction reste inconnue mais l'alignement des séquences protéiques (prise en compte de dix espèces allant du Trypanosome à l'Homme) montre une grande conservation au cours de l'évolution traduisant ainsi leur importance (J. B. Li et al. 2004). Des faisceaux de 3 hélices  $\alpha$  sont présents à l'extrémité C-terminale. Cette structure globale n'a jamais été retrouvée dans d'autres protéines du BBSome. BBS5 est une protéine qui se lie au phosphatidylinositol-phosphate (PIP) (Nachury et al. 2007b) et permet l'interaction du BBSome avec la membrane plasmique comme BBS2, BBS7 et BBS9 (Chou et al. 2019b).

Il existe un modèle poisson-zèbre *Bbs5* créé en recourant aux morpholinos qui présente un phénotype rétinien, cardiaque et rénal. Les poissons *Bbs5* sont recourbés par rapport à leurs congénères WT : un phénotype typique de ciliopathie (Al-Hamed et al. 2014). Le modèle murin Knock-out (KO) *Bbs5*<sup>-/-</sup> (*Bbs5*<sup>tm1b(EUCOMM)Wtsi</sup>) a été généré par l'IMPC « International Mouse Phenotyping Consortium » : une organisation dont l'objectif est de générer des modèles murins invalidés pour chacun des ~20000 gènes répertoriés dans le génome humain afin de comprendre leurs fonctions (Dickinson et al. 2016). Les souris *Bbs5*<sup>-/-</sup> présentent des anomalies des membres (tibia plus court). Elles deviennent obèses, ont des problèmes au niveau du métabolisme glucidique et présentent une anomalie rétinienne : un phénotype récapitulant plusieurs traits cliniques typiques des patients BBS (Meehan et al. 2017). Les 2 modèles animaux récapitulent la plupart des signes cliniques rapportés dans le cas de ciliopathies.

➤ *IFT27/BBS19*

*IFT27* est un gène situé sur le chromosome 22 (22q12.3) composé de 7 exons codants pour un transcrit de 1073 pb et une protéine de 186 AA (20kDa) appartenant au complexe des IFT. Il s'agit d'une petite GTPase ayant une forte homologie de séquences avec Rab8 et Rab11. *IFT27* est en interaction avec *IFT25*. Ce complexe protéique n'est pas nécessaire pour la formation du cil car les fibroblastes *IFT27*<sup>-/-</sup> ne présentent aucune anomalie de nombre ou de longueur du cil. En revanche, le complexe *IFT27*-*IFT25* est nécessaire pour le transport de cargos protéiques hors du cil car les cellules *IFT27*<sup>-/-</sup> présentent une accumulation au cil du récepteur SMO (voie SHH) et de certaines sous-unités du BBSome (Figure 12). *IFT27* est une protéine du transport antérograde qui régule également le transport rétrograde des protéines hors du cil (Eguether et al. 2014a).

Aldahmesh et al. 2014 a rapporté, pour la première fois, *IFT27* comme le 19<sup>ème</sup> gène BBS en présentant le cas d'une famille avec un phénotype typique de BBS (RP, obésité, déficience intellectuelle, polydactylie et atteinte rénale). Une validation fonctionnelle au moyen du modèle poisson-zèbre (généralisé avec l'injection de morpholino anti-*Ift27*) a montré des caractéristiques typiquement retrouvées dans le cas de poisson-zèbres modèles de ciliopathie (courbature de l'axe antéro-postérieur, défaut dans le transport rétrograde de mélanosomes). Un modèle murin KO pour *Ift27* a également été étudié par Eguether et al. 2014. Les souris présentent des anomalies compatibles avec une dérégulation de la voie SHH : polydactylie, malformations squelettiques, malformations du septum atrial, défaut de positionnement des organes. Le phénotype des souris *Ift27*<sup>-/-</sup> est similaire à celui des souris *Ift25*<sup>-/-</sup> et proche de celui des patients. L'ensemble de ces données expérimentales confirme et précise les fonctions rapportées pour cette protéine.



**Figure 12 :** Le complexe IFT27-IFT25. Représentation schématique de l'impact du complexe IFT27-IFT25 sur le fonctionnement du cil primaire (inspiré de Nakayama et Katoh 2018).

### 3. Les autres ciliopathies et les gènes associés

#### A) *IQCE*

La polydactylie est l'anomalie des membres la plus fréquente dans la population générale (3%). Elle fait partie des signes cliniques typiquement rapportés pour des patients atteints de ciliopathies syndromiques. Par exemple, les syndromes de Ellis Van Creveld (EVC) et de Weyers sont 2 ciliopathies syndromiques causées par des mutations identifiées dans les gènes *EVC* et *EVC2* codants pour des protéines localisées au niveau de l'EVC zone. Elles sont causées par une perturbation de la voie SHH lors du développement des tissus cardiaques, squelettiques et orofaciaux (Pusapati et al. 2014).

La polydactylie peut être l'unique symptôme d'une maladie monogénique causée par des mutations dans 1 des 8 gènes de polydactylie rapportées actuellement. Si ce gène code pour une protéine impliquée dans le fonctionnement du cil alors la polydactylie est considérée comme une ciliopathie non syndromique, comme c'est le cas pour *IQCE*.

Ce gène est situé sur le chromosome 7 (7p22.3), se compose de 22 exons codants pour un transcrit de 6887 pb et une protéine de 695 AA (77 kDa) qui est localisée au niveau de l'EVC zone en interaction avec EVC et EVC2. Une mutation dans le gène *IQCE* a été rapportée par Umair et al. 2017 comme responsable de polydactylie post-axiale (PAP). Les patients issus d'une famille consanguine ont subi un examen global qui a révélé une PAP localisée uniquement au niveau des pieds. Ils sont porteurs à l'état homozygote de la mutation c.395-1G>A, p. Gly132Valfr\*22 dans le gène *IQCE*, à présent considéré comme un gène de ciliopathie non syndromique (PAP). Le modèle murin KO *Iqce*<sup>-/-</sup> présente des atteintes squelettiques multiples (polydactylie, anomalies des doigts, longueurs de tibias anormales) accompagnées d'autres atteintes (métabolique, visuelle, cardiovasculaire).

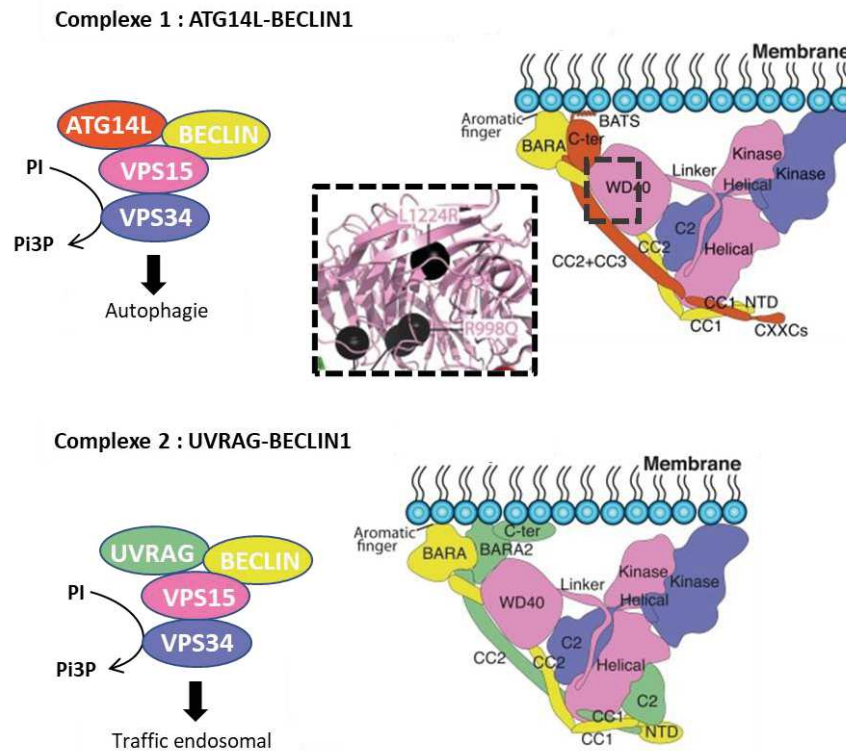
#### B) VPS15

En 2016, le Laboratoire de génétique médicale a publié un article rapportant pour la première fois *VPS15* comme un gène responsable de ciliopathie (Stoetzel et al. 2016). *VPS15* est une protéine de 1358 AA codée par le gène *VPS15* (3q22.1) qui comporte 20 exons, chez l'homme. Le gène *Vps15* est situé sur le chromosome 9 chez *Mus musculus* et code également une protéine de 150 kDa présentant 96% d'identité avec le gène humain. Il s'agit d'une protéine habituellement connue pour son implication dans le processus d'autophagie.

##### ➤ Le rôle de VPS15 dans le processus d'autophagie

La Vacuolar Protein Sorting associated protein 15 (*VPS15*) est également appelée PIK3R4 (sous unité régulatrice n°4 de la phosphatidyl-inositol 3 kinase). Il s'agit d'une protéine prédite comme étant une sérine-thréonine kinase à cause de son homologie de séquence avec certaines protéines telle que la protéine kinase A (Knighton et al. 1991). Sa fonction de kinase n'a jamais été confirmée. Cette protéine, découverte chez la levure est appelée p150 en référence à son poids moléculaire de 150 kDa (Panaretou et al. 1997). Il s'agit d'une sous unité régulatrice de la phosphatidylinositol 3 kinase (PI3K) de classe III qui permet la formation de Pi3P à partir de PiP. Elle subit une modification post traductionnelle (myristoylation) en N-terminale qui permet son attache aux membranes (notamment celles du RE ou de l'autophagosome) (Backer 2016). Le complexe *VPS15-VPS34* joue un rôle important dans l'autophagie (en interaction avec ATG14L/BECLIN-1, complexe 1) et dans le transport des vacuoles cytoplasmiques (en interaction avec ATG14L/UVRAG, complexe 2) (Pampliega et al. 2013) (Figure 13).

L'article Stoetzel et al. 2016 rapporte pour la première fois un rôle inédit de Vps15 dans le fonctionnement du cil primaire. La mutation R998Q ne perturbe pas l'interaction de VPS15 avec VPS34 et la formation de Pi3P. La ciliopathie développée est donc provoquée par une anomalie de fonctionnement de VPS15 indépendamment de VPS34 et indépendamment du processus d'autophagie.



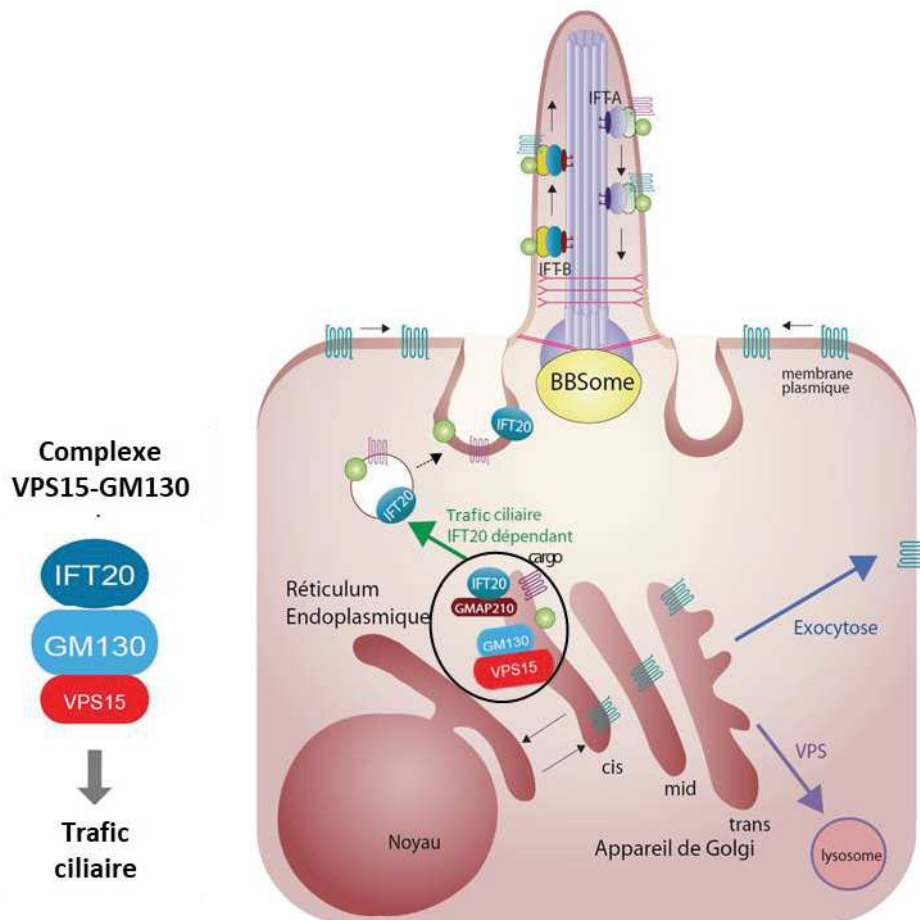
**Figure 13 :** Les complexes 1 et 2 de la PI3K de classe III. Zoom sur la mutation R998Q localisée au niveau du domaine WD40 de VPS15. (adaptée de Ohashi, Tremel, et Williams 2019) PI : phosphatidyl-inositol, Pi3P, phosphatidyl-inositol 3 phosphate.

➤ Le rôle de VPS15 dans le cil primaire

Les patients présentent tous le phénotype suivant : RP précoce, anomalie rénale (petits reins, kystes, insuffisance rénale), retard de croissance, anomalies des membres (clinodactylie et brachydactylie) et problèmes d'apprentissage. 1 des 3 patients est en surpoids. La mutation faux sens p.R998Q a été identifiée dans le gène *VPS15* à l'état homozygote et se situe au niveau du domaine fonctionnel WD40 de VPS15, connu pour être impliqué dans des interactions protéines-protéines (Figure 13).

Comme le résume la Figure 14, il ressort de ce travail que VPS15 est localisée au niveau du cis golgi et interagit avec GM130 (une protéine marqueur du cis golgi). La mutation R998Q n'impacte pas les interactions de VPS15 avec GM130 et d'IFT20 avec GMAP210. Par contre, IFT20 s'accumule au cis golgi, un nombre réduit de vésicules IFT20 positives cytoplasmiques est dénombré. IFT20 est retenue par GMAP210 au niveau du cis golgi grâce à un domaine coiled-coiled de 163 AA (Follit et al. 2008).

GMAP210 est une protéine qui intervient dans le trafic des protéines ciliaires par l'intermédiaire d'IFT20. GMAP210 et GM130 sont toutes les 2 localisées au niveau du cis Golgi et sont partiellement redondantes (Roboti, Sato, et Lowe 2015). Il n'existe pas encore de données rapportant une interaction entre le complexe VPS15-GM130 et IFT20, au niveau du cis golgi. Mais l'hypothèse formulée à l'issue du travail publié en 2016 est la suivante : le complexe VPS15-GM130 interagit avec IFT20 et sert de « plateforme de chargement d'IFT20 » dans les vésicules du cis golgi (Figure 14). Les vésicules IFT20 positives contenant des protéines ciliaires transitent depuis le cis golgi vers le cil, le long des MT. L'implication de VPS15 dans le fonctionnement du cil primaire serait donc associée à un mécanisme IFT20 dépendant (Stoetzel et al. 2016).



**Figure 14 :** Schéma bilan tiré de Stoetzel et al. 2016. Illustration des interactions et du rôle de VPS15 dans le fonctionnement du cil primaire. (Stoetzel et al. 2016)

#### 4. Les maladies neurodégénératives

L'atteinte cérébrale est l'un des signes cliniques développés par les patients atteints de ciliopathies. La membrane ciliaire est composée de récepteurs transmembranaires tels que des récepteurs à la somatostatine (SSTR3), à la dopamine (DR) qui déclenchent des cascades de signalisations sous-jacentes permettant l'établissement de fonctions neurologiques complexes (Goetz, Liem, et Anderson 2012). Des maladies neurodégénératives telles que l'ataxie spino-cérébelleuses ou la maladie de

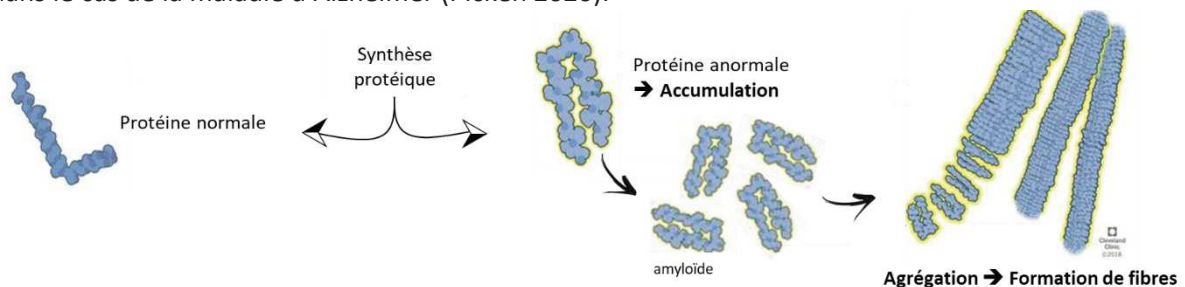


Huntington ont été rapportées comme pouvant être causées par des gènes codants des protéines impliquées dans la régulation de la ciliogénèse telles que *TTBK2* et *HTT* (Youn et Han 2018).

Les maladies neurodégénératives sont des maladies caractérisées par une mort progressive des neurones du système nerveux central menant à des troubles cognitifs tels que la perte de mémoire et des troubles moteurs. Les mécanismes moléculaires à l'origine de ces pathologies sont multiples et impactent diverses régions du cerveau.

#### A) Les amyloïdoses

Il s'agit d'un ensemble de maladies caractérisées par l'agrégation de protéines anormales sous la forme de fibres amyloïdes conduisant à un dysfonctionnement de l'organe touché. Les fibres forment des dépôts extracellulaires insolubles et résistants à la protéolyse, formés à partir de feuilletts  $\beta$  antiparallèles (Figure 15). Une multitude d'organes peut être atteints comme le cœur, le pancréas, le cerveau... Le diagnostic est établi par observation histologique du tissu coloré au rouge Congo sous un microscope à lumière polarisée : les fibres amyloïdes sont caractérisées par une réfringence jaune-verte. La classification des amyloïdoses se base sur la nature de la protéine impliquée dans les dépôts et les symptômes cliniques (Baker et Rice 2012). 36 protéines ont été identifiées comme étant amyloïdogéniques chez les humains. Par exemple, les peptides  $A\beta$  s'accumulent au niveau du cerveau dans le cas de la maladie d'Alzheimer (Picken 2020).



**Figure 15 : Les amyloïdoses.** Représentation schématique du principe général.

#### B) La maladie d'Alzheimer :

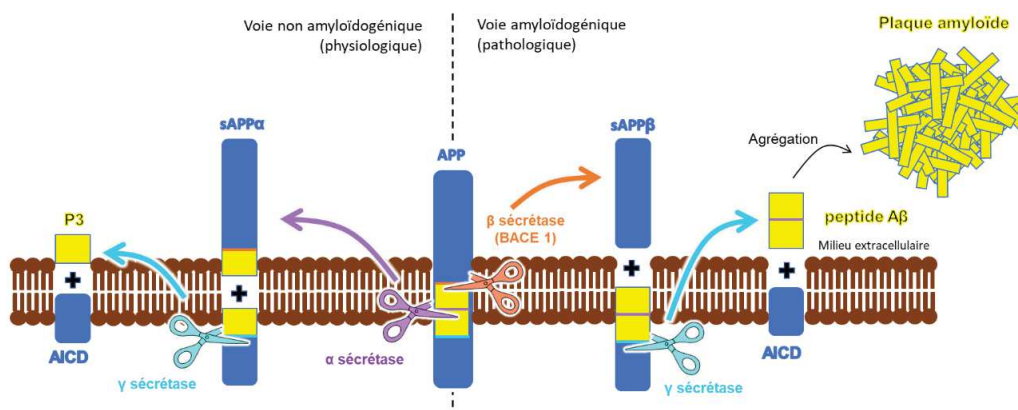
Elle fait partie des amyloïdoses emblématiques et s'avère être la première cause de démence chez les personnes âgées. Elle touche le plus souvent des personnes de plus de 65 ans et son incidence explose après 85 ans : une donnée majeure dans une société vieillissante comme la nôtre où l'espérance de vie est en hausse. D'après les données de la Fondation pour la Recherche Alzheimer, en France, 900 000 malades ont été recensés : un nombre qui devrait doubler d'ici 2050 sachant que 250 000 nouveaux cas s'ajoutent chaque année. Au niveau mondial, selon le *World Alzheimer report* édité par l'Alzheimer's disease international en 2015, 35,6 millions de personnes sont atteintes et le chiffre est estimé à plus de 92 millions d'ici 2050. A ce jour, les seuls traitements existants sont palliatifs face aux

troubles cognitifs des patients provoqués par la mort neuronale, la présence de plaques amyloïdes et de dégénérescences neurofibrillaires (DNF) composées de protéines TAU hyperphosphorylées.

Dans 25% des cas, les formes de maladies d'Alzheimer sont dites familiales c'est-à-dire quand au moins 3 patients de la même famille sont atteints. Dans 2 à 3% de ces formes familiales, la maladie d'Alzheimer est une maladie génétique à transmission autosomique dominante liée à des mutations identifiées dans les gènes *APP*, *PSEN1*, *PSEN2* et souvent associée à un âge de développement précoce. Dans 75% des cas, les formes de maladies d'Alzheimer ne sont pas familiales, il s'agit alors de cas dits sporadiques dont les causes sont multifactorielles : l'accumulation de variants dans des gènes de susceptibilités augmentant le risque de survenue de la maladie. Par exemple, il est estimé qu'une personne hétérozygote pour l'allèle *APOEε4* présente 10 à 20% de risque de développer une maladie d'Alzheimer vers 75 ans. Ce pourcentage atteint 25-35% pour les personnes porteuses homozygotes de cet allèle. Toutefois, même en l'absence de cet allèle chez un individu, le risque de développer la maladie n'est pas nul (Bird 1993).

➤ Les plaques amyloïdes

Le précurseur APP (Amyloïd Precursor Protein) subit des clivages successifs afin de générer des peptides relargués par la cellule. 2 voies co-existent dans les cellules pour le métabolisme d'APP. Comme le résume la Figure 16, la voie non amyloïdogénique se déroule de manière physiologique dans les cellules : APP est clivé par l' $\alpha$  sécrétase puis par la  $\gamma$  sécrétase pour finalement permettre le relargage des peptides sAPP $\alpha$ , P3 et AICD. La voie amyloïdogénique aboutit à la synthèse des peptides A $\beta$  1-40 et A $\beta$  1-42 après un premier clivage par la  $\beta$  sécrétase suivi d'un deuxième clivage par la  $\gamma$  sécrétase. En condition normale, les peptides A $\beta$  sont dégradés et la voie non amyloïdogénique est favorisée. Mais, en condition pathologique, la production d'APP est activée et la clairance des peptides A $\beta$  est inhibée : ce déséquilibre mène à l'accumulation des peptides qui s'agrègent et forment des plaques  $\beta$  amyloïdes.

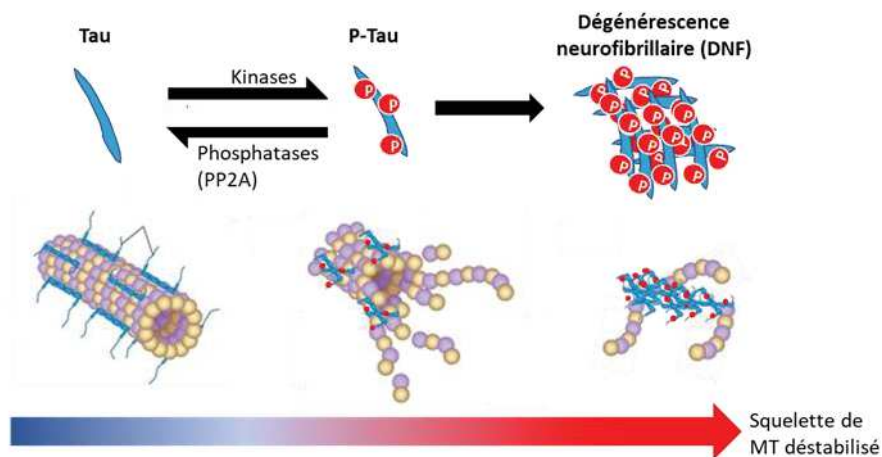


**Figure 16 : Les voies non amyloïdogéniques et amyloïdogéniques.** Représentation schématique du métabolisme du précurseur APP par les différentes sécrétases (inspirée de Bachurin, Bovina, et Ustyugov 2018).



➤ Les protéines TAU phosphorylées et les DNF

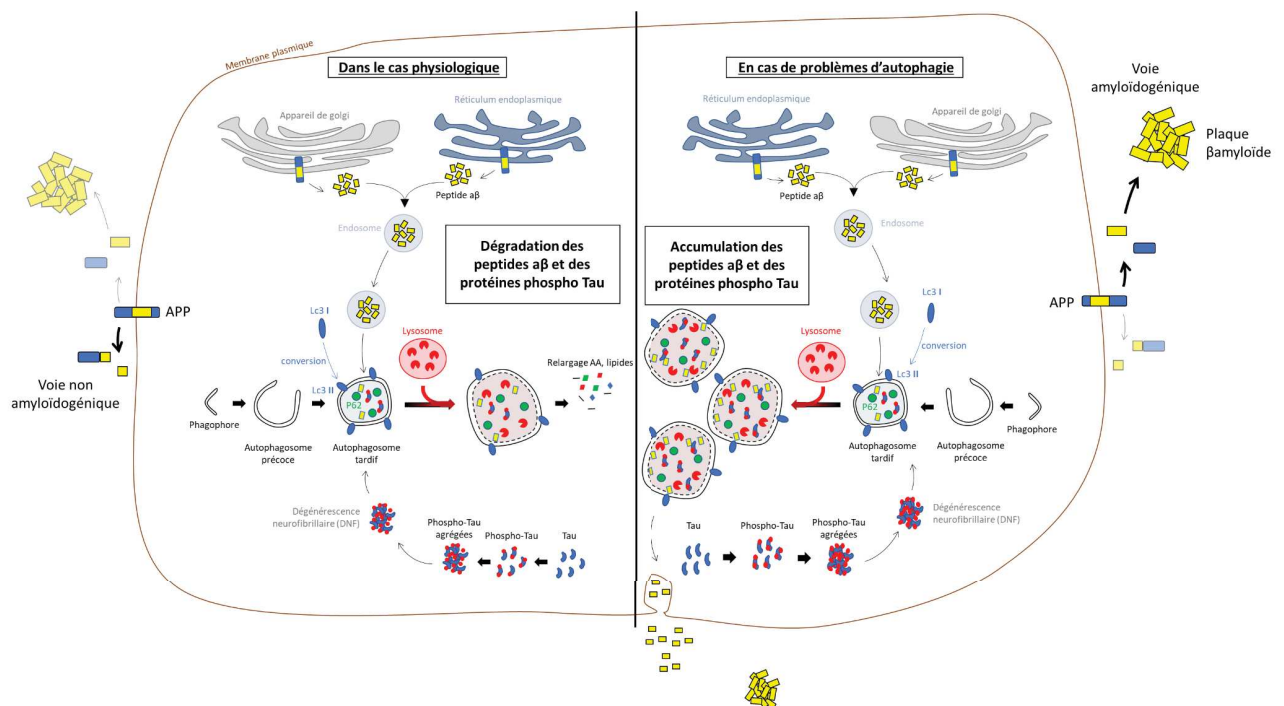
TAU est une protéine associée aux MT neuronaux qui permet leurs formations et leurs stabilisations. Dans des conditions normales, TAU est une protéine dont le degré de phosphorylation est hautement régulé. Dans des conditions pathologiques, les protéines TAU sont hyperphosphorylées et s'accumulent sous forme de filaments intraneuronaux conduisant à une dégénérescence neurofibrillaire appelée DNF (Alonso et al. 2018). L'augmentation de ce niveau de phosphorylation peut être due (1) à l'hyperactivation des kinases de TAU telle que GSK3 $\beta$  qui phosphoryle TAU au niveau des sérines 199, 202, 396, 404, (2) à un défaut de la phosphatase (PP2A), (3) à un défaut de dégradation des protéines TAU hyperphosphorylées (Miao et al. 2019). Ces formes séquestrent la protéine TAU normale et la remplace au niveau du squelette de MT qui se retrouve déstabilisé (Figure 17). Les tauopathies forment un groupe de maladies caractérisées par des anomalies de TAU et la maladie d'Alzheimer en fait partie (Alonso et al. 2018).



**Figure 17: Tau, phospho-Tau et DNF.** Représentation schématique de la formation des formes hyperphosphorylées de TAU, de leurs agrégations définies par le terme « dégénérescences neurofibrillaires » (DNF) et des conséquences sur la stabilité du squelette de microtubules (MT). (Utilisation d'images de © Alila Medical Media, Shutterstock)

➤ L'impact de l'autophagie

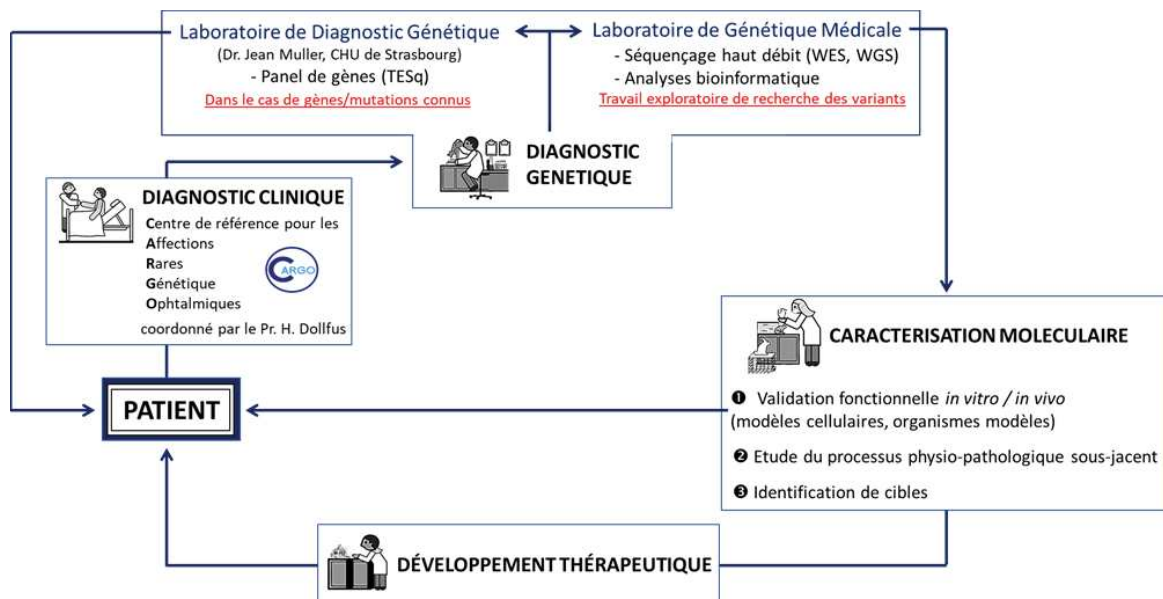
Plusieurs études ont déjà rapporté qu'un dysfonctionnement du processus d'autophagie a un impact majeur sur le développement de la maladie d'Alzheimer (Figure 18). En cas de perturbation du processus d'autophagie, les protéines TAU hyperphosphorylées et les peptides A $\beta$  qui sont synthétisés dans le cytoplasme ne sont pas dégradés à cause, par exemple, d'un défaut de fusion entre autophagosome et lysosome (Cai et al. 2015). L'observation des neurites de souris modèles pour la maladie d'Alzheimer a déjà révélé l'accumulation d'autophagosomes (Nilsson et Saido 2014). Par ailleurs des études *in vitro* ont montré que des neurones exposés à la rapamycine (un inhibiteur de la voie mTORC1 qui stimule le processus d'autophagie) présentent une sécrétion accrue de peptides A $\beta$ .



**Figure 18 : Autophagie et maladie d'Alzheimer.** Représentation schématique comparant la dégradation ou le maintien des peptides Aβ et des protéines Tau dans 2 cas : le processus d'autophagie fonctionne normalement (physiologique) ou le processus est perturbé (pathologique).

## IV. La recherche translationnelle

Les scientifiques qui s'intéressent aux ciliopathies doivent faire face à plusieurs difficultés liées à la grande hétérogénéité de ces pathologies. D'abord, ils sont confrontés à l'hétérogénéité clinique : les signes cliniques sont variables entre des individus, même au sein d'une famille et se chevauchent entre plusieurs ciliopathies différentes (Figure 20). Ensuite, ils font face à l'hétérogénéité génétique puisque d'une part, plusieurs gènes sont responsables d'une même pathologie et d'autre part, certains gènes peuvent donner lieu à des pathologies différentes (Figure 21). La sévérité des ciliopathies va dépendre du nombre d'organes atteints et varier selon le mécanisme ciliaire impacté (la ciliogénèse, le maintien d'un fonctionnement normal du cil, les voies de signalisation) (Wheway, Nazlamova, et Hancock 2018). Le travail du laboratoire de génétique médicale s'inscrit dans une recherche dite translationnelle qui fait converger le travail des cliniciens, des bioinformaticiens, des chercheurs vers un objectif commun : l'identification des causes moléculaires pour des familles concernées par « l'errance diagnostique » et la meilleure compréhension des mécanismes physiopathologiques (Figure 19). Dans ce cas, les données de la recherche fondamentale sont appliquées au cas d'un patient qui s'avère être à la fois le point de départ et le point d'arrivée de l'étude.



**Figure 19 : La recherche translationnelle.** Représentation schématisée des différentes étapes/acteurs de cette recherche appliquée au Laboratoire de génétique médicale (UMR\_S 1112).

### 1. Le diagnostic clinique

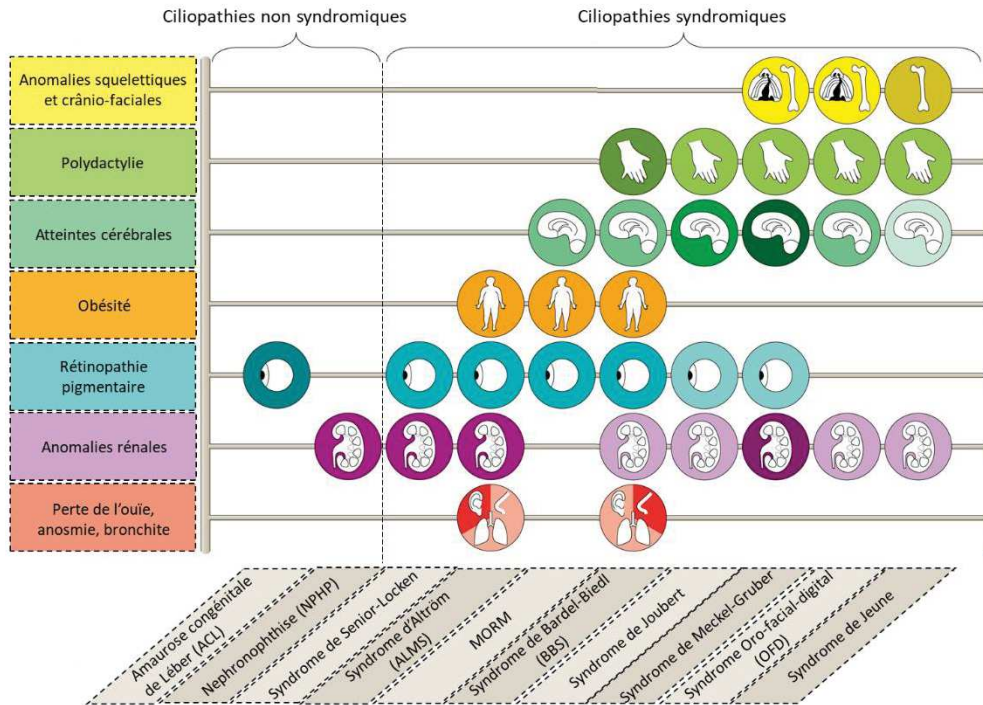
Strasbourg est un centre de référence pour de nombreuses maladies rares dont les ciliopathies. En effet, le CARGO (Centre de référence pour les maladies génétiques rares ophtalmiques) centralise et accueille tous les patients présentant des anomalies ophtalmologiques comme des RP. Les cliniciens vont émettre une hypothèse diagnostique plus ou moins forte en répertoriant les signes cliniques des patients. Elle sera confirmée ou non en fonction du diagnostic génétique/moléculaire. L'identification

des causes génétiques est réalisée en routine au Laboratoire de diagnostic génétique (NHC, Strasbourg) en recourant à des techniques de biologies moléculaires classiques parmi lesquelles on peut trouver le séquençage Sanger, la PCR quantitative, les puces à ADN et le séquençage à haut débit (panels de gènes ou target exome sequencing TES) selon un arbre décisionnel établi par les laboratoires de référence. Ces panels de gènes contiennent une liste établie plus ou moins larges de gènes connus pour être impliqués dans les ciliopathies. Dans le cas où le résultat est négatif (aucune mutation n'a été identifiée) et si le patient ou sa famille a donné son consentement, les études moléculaires peuvent se poursuivre au sein du Laboratoire de recherche (LGM, U1112). L'identification des mutations responsables est réalisée avec d'autres techniques telles que le séquençage complet de l'exome WES (analyse des séquences codantes) et/ou du génome complet (WGS) voire de l'ARN complet (RNASeq). Une analyse bioinformatique sur les données générées permet de filtrer les variants et de retenir les plus pertinents. Une caractérisation moléculaire par des expériences de biologies moléculaires et cellulaires et de l'utilisation de modèles *in vivo* adéquats vont permettre de confirmer et de valider la pathogénicité des mutations identifiées, d'étudier le mécanisme physiopathologique sous-jacent et d'établir une éventuelle corrélation génotype-phénotype. Le dernier pan de ce type d'étude est l'identification de protéines pouvant devenir de potentielles cibles pour le développement des molécules thérapeutiques. Le retour vers le patient permettra (1) de mettre un nom sur les causes de la maladie, (2) de permettre une prise en charge optimale avec une possible anticipation de la survenue des signes cliniques suivant le diagnostic, (3) la possibilité de proposer un conseil génétique au patient et à ces apparentés incluant un diagnostic préimplantatoire et/ou prénatal.

#### A) Les difficultés associées au diagnostic clinique

La présence d'une polydactylie, d'une obésité, d'anomalies au niveau de la vision, de l'ouïe, du fonctionnement rénal et des fonctions cognitives constituent des signes cliniques typiquement retrouvés dans le cas des ciliopathies. La présence ou l'absence d'un signe devrait théoriquement permettre de distinguer une ciliopathie d'une autre. Cependant, il existe des chevauchements cliniques entre 2 ciliopathies différentes et des variabilités entre des individus atteints de la même pathologie qui rendent d'autant plus compliqué le rendu d'un diagnostic clinique (Mitchison et Valente 2017a). Par exemple, en théorie, le syndrome d'Alström se distingue du BBS par l'absence de polydactylie et de déficience intellectuelle (Figure 20). Des différences qui ne sont pas si marquées en pratique. Il existe des cas de patients qui peuvent ne pas présenter les signes typiques du syndrome à cause de leurs habitudes de vie. Par exemple, il existe des patients BBS qui ne présentent pas d'obésité car ils suivent un régime alimentaire ou pratiquent beaucoup de sport. A l'inverse, il existe des familles pour lesquelles un seul des patients d'une fratrie présente un surpoids, une caractéristique clinique courante dans la population générale.

L'obtention d'un diagnostic clinique est privilégiée durant l'enfance afin de permettre une prise en charge optimale le plus tôt possible. Mais dans certains cas, les signes cliniques peuvent se déclarer plus tardivement car ils sont consécutifs à une dégénérescence progressive : un suivi régulier du patient permet d'affiner le diagnostic et la prise en charge avec l'âge grandissant.



**Figure 20 : Abaque des ciliopathies.** Représentation des signes cliniques en fonction du type de ciliopathies. L'intensité de couleur des signes cliniques est proportionnelle à la sévérité de l'atteinte. MORM : retard mental, obésité troncale, rétinopathie pigmentaire, micropénis.

## 2. Le diagnostic génétique

Le diagnostic génétique est une étape primordiale des études translationnelles car il permet d'identifier les causes génétiques du phénotype du patient. Les cliniciens l'utilisent pour éliminer les hypothèses de diagnostic qu'ils ont pu formuler face aux tableaux cliniques très hétérogènes de certains patients ou pour confirmer le diagnostic émis.

### A) Les difficultés associées au diagnostic génétique

Une exploration de l'ADN des patients et de leurs parents est généralement réalisée dans la mesure du possible par une étude dite « en trio ». Elle permet d'établir plus précisément la ségrégation familiale et le diagnostic génétique. Ce dernier peut confirmer soit le diagnostic clinique dans le cas de gènes connus ou soit étoffer les listes de gènes responsables dans le cas de gènes initialement candidats pour lesquels la caractérisation fonctionnelle a prouvé la pathogénicité. La difficulté d'émettre un diagnostic génétique réside à 2 niveaux (Figure 21). D'une part, un gène peut être responsable de 2 ciliopathies différentes : par exemple, des mutations dans *BBS1* peuvent provoquer

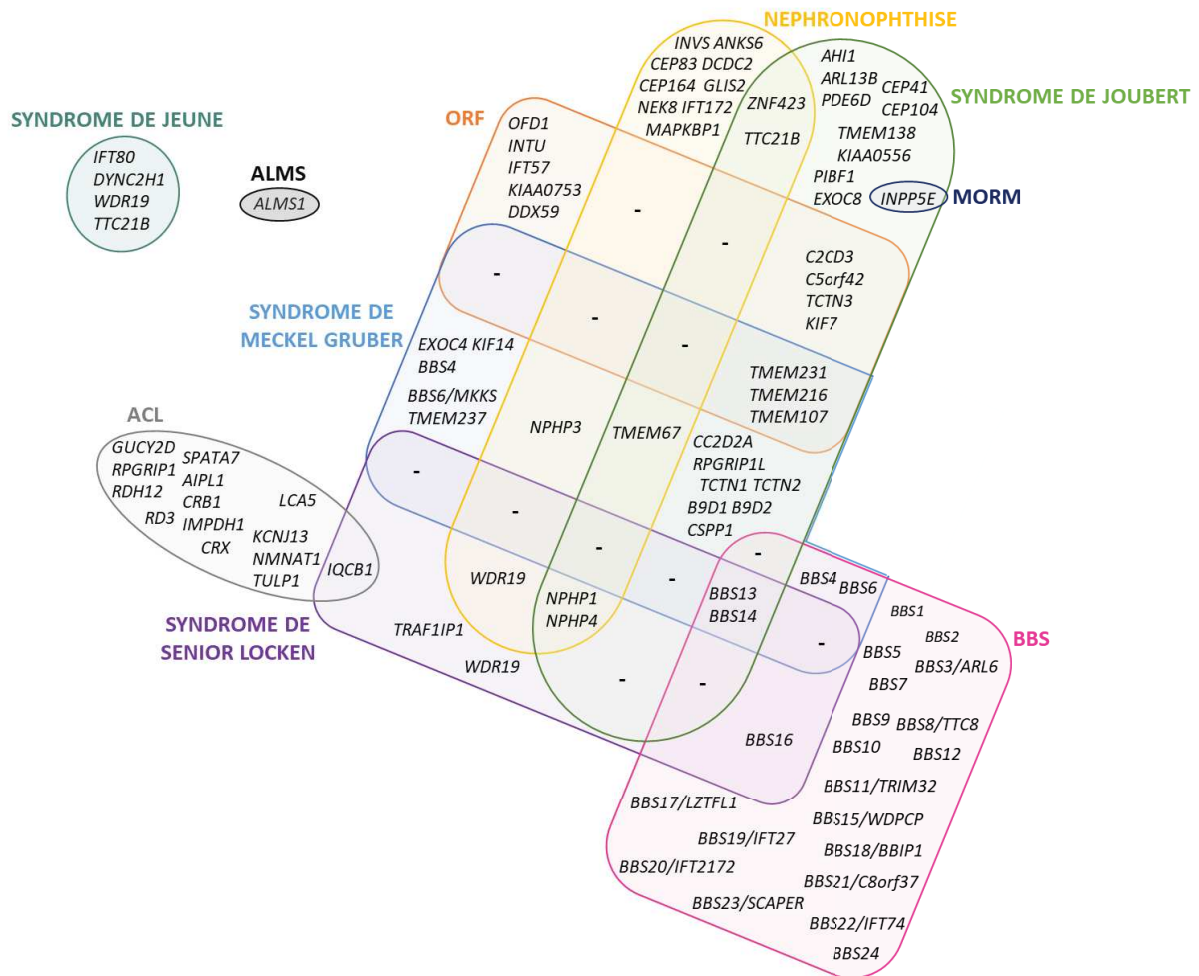
un BBS (ciliopathie syndromique) ou une RP isolée (ciliopathie non syndromique). D'autre part, plusieurs gènes peuvent causer les mêmes phénotypes : par exemple, 24 gènes BBS ont été identifiés comme responsables d'une même ciliopathie, le BBS. Dans certaines familles, les patients présentent les mêmes mutations mais la sévérité des symptômes cliniques varie d'un individu à l'autre. Un mode de transmission oligogénique est suggéré : c'est la présence d'une 3<sup>ème</sup> mutation à l'état hétérozygote dans d'autres gènes BBS en plus de 2 mutations bialléliques (homozygotes ou hétérozygotes composites) dans un premier gène BBS qui rend le phénotype pénétrant (Mitchison et Valente 2017a). Il s'agit d'une hypothèse déjà avancée pour expliquer les variabilités cliniques mises en évidence entre des apparentés BBS ayant les mêmes mutations génétiques. L'hypothèse du triallélisme a été émise uniquement dans le cas du BBS (Badano et Katsanis 2002). Il existe d'autres cas où l'association de 2 mutations hétérozygotes dans 2 gènes différents causent la maladie : on parle d'hérédité digénique, une hypothèse avancée dans le cas de patients atteints de néphronophtise (Bachmann-Gagescu 2014). Il existe également plusieurs publications qui rapportent l'existence de variants modificateurs génétiques : par exemple, le variant R930W présent dans le gène *AHI1* est associé au développement préférentiel d'une RP chez des patients atteints de néphronophtise (Bachmann-Gagescu 2014).

B) L'apport des techniques NGS

Avant l'essor des techniques de séquençage haut débit il y a une dizaine d'années, les chercheurs pratiquaient principalement en routine des analyses gène par gène tels que le séquençage Sanger, le cas échéant des MLPA (multiplex ligation-dependent probe amplification) permettant de mettre en évidence des variations du nombre de copie (CNV) connues dans des gènes connus, des SNP arrays (cartographie par autozygotie/homozygotie). L'objectif de ces techniques était de réaliser à moindre coût le test d'une liste de gènes connus assez réduite à l'époque. Ces stratégies diagnostics se sont avérées consommatrices de temps, d'argent et ne permettaient pas une étude à grande échelle. La cartographie par autozygotie est encore appliquée pour des maladies autosomiques récessives dans le cas de familles consanguines. L'objectif est de déterminer les régions chromosomiques identiques chez tous les individus atteints qui sont présentes à l'état homozygote. Les gènes des régions concernées sont ensuite séquencés par Sanger à la recherche de mutations (Bachmann-Gagescu 2014). Nous avons utilisé cette technique dans plusieurs projets dont celui portant sur le gène *BBS1*.

Face à la difficulté de relier le phénotype complexe des patients à un seul gène, les cliniciens étaient de plus en plus demandeurs d'analyser plusieurs gènes potentiellement responsables de la pathologie.



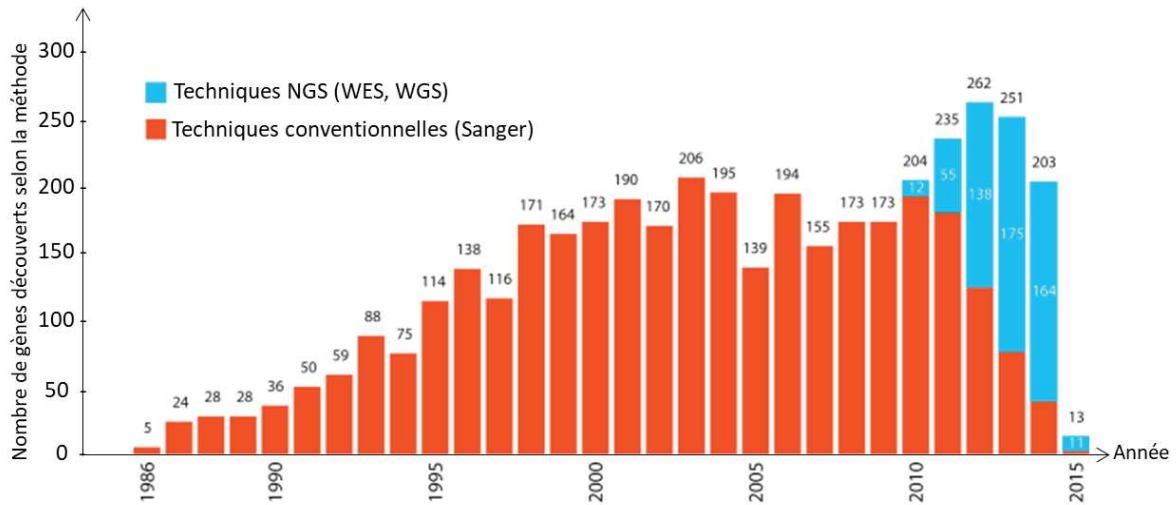


**Figure 21 : L'hétérogénéité génétique des ciliopathies.** Diagramme de Venn représentant les chevauchements des diagnostics moléculaires : certains gènes sont responsables de plusieurs ciliopathies. (inspiré de Mitchison et Valente 2017a) ACL : Amaurose congénitale de Leber, OFD : syndrome oro-facial-digital, ALMS : syndrome d'Alström, BBS : Syndrome de Bardet-Biedl. MORM : Retard Mental, Obésité troncale, Rétinopathie pigmentaire, Micropénis.

Le séquençage haut débit ou séquençage de nouvelle génération (NGS) regroupe un ensemble de méthodes qui vont interroger simultanément un grand nombre de régions de notre génome et détecter un grand nombre de variations génétiques à l'échelle : d'un groupe de gènes pour le TES, des régions codantes du génome pour le WES (2% du génome) et du génome entier pour le WGS. Il s'agit de techniques devenues de plus en plus abordables à partir de 2010, période à partir de laquelle leur utilisation s'est démocratisée dans les laboratoires. Par rapport aux méthodes utilisées avant, le NGS est plus rapide, moins coûteux et permet l'analyse simultanée d'énormes quantités d'ADN (le traitement de plusieurs patients en parallèle) (Mitchison et Valente 2017a).

Le recours aux techniques NGS a permis des avancées non négligeables dans l'identification et la résolution des cas de ciliopathies (Figure 22). Ces techniques ont facilité la mise en place de corrélations génotype-phénotype et ont permis de reclasser les phénotypes compliqués de certaines ciliopathies en fonction des causes génétiques. A l'heure actuelle, le BBS est la ciliopathie pour laquelle

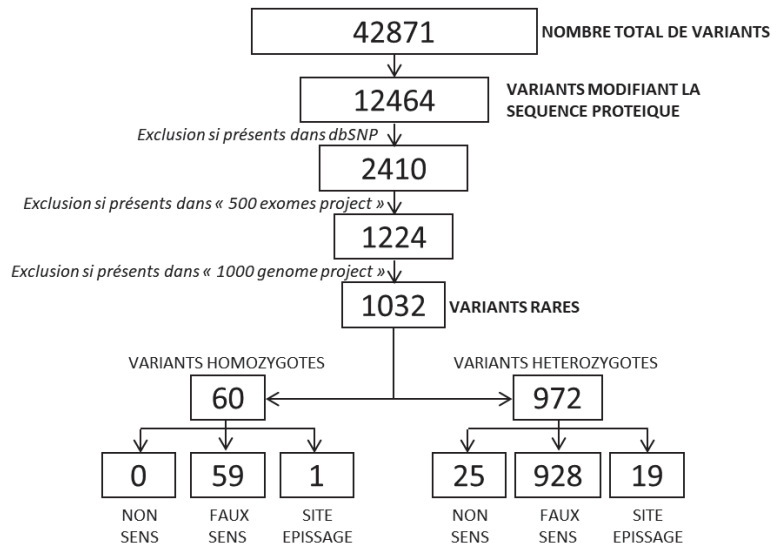
la proportion de cas résolus est la plus élevée en atteignant 80% des cas. Ce nombre s'élève à 70% dans le cas du syndrome de Meckel-Gruber, 60% pour le syndrome de Joubert et seulement 20% dans le cas des néphronophtises (Mitchison et Valente 2017a).



**Figure 22 : L'apport des techniques NGS dans le diagnostic génétique.** Histogramme représentant le nombre approximatif de gènes découverts en fonction de la technique utilisée. L'émergence des techniques NGS depuis 2010 est marquée en bleu (inspiré de Chog et al. 2015).

L'un des enjeux majeurs face à la quantité massive de données générées est la capacité de les stocker, de les annoter et surtout de les interpréter avec des logiciels dédiés. Comme le résume la Figure 23, une série de filtre est appliquée aux données obtenues pour réduire le nombre de variants et sélectionner ceux potentiellement responsables de la pathologie en fonction du mode de transmission, des mécanismes physiopathologiques connus, des données de pénétrance. Dans le cas des maladies autosomiques récessives, la recherche est d'abord orientée vers des variants présents à l'état homozygote, dans le cas de familles consanguines. Dans la majorité des cas, on considère le cas de patients hétérozygotes composites pour lesquels 2 mutations différentes dans le même gène sont identifiées et sont responsables du phénotype. Il est possible que certains variants aient été déjà rapportés dans les bases de données telles que ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>) par exemple. Les variants filtrés doivent être absents ou très peu fréquents dans les bases de données de la population générale telles que 1000 genomes (<https://www.internationalgenome.org/>) ou gnomAD (<https://gnomad.broadinstitute.org/>). Des logiciels de prédictions tels que SIFT (<https://sift.bii.a-star.edu.sg/>) et PolyPhen-2 (<http://genetics.bwh.harvard.edu/pph2/>) permettent de déterminer l'impact des mutations sur la structure et la fonction protéique. L'impact sur l'épissage peut être testé en recourant à MaxEntScan (Yeo et Burge 2004), NNSplice (Reese et al. 1997) et Splice Site Finder (Shapiro et Senapathy 1987).





**Figure 23 : La filtration de variants.** Exemple appliqué aux données de WES (inspiré de Raffan et Semple 2011).

Une classification des variants a été établie selon les recommandations de l’ACMG (American college of medical genetics) et de l’AMP (Association for molecular pathology) (Richards et al. 2015). Il existe 5 classes de variants dont la classe 1 correspond aux variants bénins et la classe 2 à ceux probablement bénins : dans ces 2 cas, le lien de causalité avec la pathologie est exclu.

A l’inverse, la classe 5 correspond aux variants pathogènes et la classe 4 à ceux probablement pathogènes. Dans ces 2 cas, leur effet sur la fonction de la protéine est avéré ou fortement suspecté. Le lien de causalité entre le variant et la pathologie est établi s’il est présent à l’état homozygote ou s’il est présent en trans d’une autre mutation classée 4 ou 5, dans le cas de maladies récessives. Dans la majorité des analyses, les variants de la classe 3 (de signification inconnue (VSI ou VUS)) posent problème car leur effet sur la fonction de la protéine ne peut pas être exclu. Des études complémentaires (études de co-ségrégation, études fonctionnelles *in vitro*, études d’épissage) doivent être mises en œuvre afin de re-classer les variants en bénins ou pathogènes et permettre le rendu d’un diagnostic univoque aux patients. Beaucoup de ces variations sont des mutations faux-sens qui conduisent à des substitutions d’AA et dont l’effet protéique doit être déterminé. Les mutations non-sens (conduisant à l’apparition d’un codon stop prématuré), les mutations d’épissage (touchant les sites donneurs ou accepteurs d’épissage), les délétions/insertions d’un nombre non multiple de 3 (entraînant un décalage du cadre de lecture et par conséquent l’apparition d’un codon stop prématuré) constituent un groupe de variants dont la preuve de pathogénicité est plus aisée (classe 4 et 5).

### ➤ Le séquençage ciblé ou panel de gènes (TES)

Il s'agit d'une technique NGS qui permet de faire le séquençage d'un groupe défini de gènes (panels de gènes) dont la pathogénicité dans la maladie d'intérêt est avérée. Par exemple, le laboratoire de diagnostic génétique a déterminé un panel de 58 gènes incluant les gènes BBS et ceux de diagnostics différentiels (MKS, NPHP, JBTS, ALMS...). L'utilisation de panels de gènes permet de prioriser l'analyse par rapport à une entité phénotypique spécifique. Cela permet une interprétation simplifiée en garantissant une qualité importante des données et une diminution du coût (augmentation du nombre de patients séquencés simultanément). Néanmoins, ces types d'analyses sont par définition restreintes aux gènes choisis. Il s'agit de la première étape d'analyse, rapide et peu coûteuse qui est mise en place dans le schéma aboutissant au diagnostic génétique.

### ➤ Le séquençage complet de l'exome (WES)

Il s'agit d'une technique NGS qui permet de réaliser le séquençage complet des séquences codantes du génome (soit 2%). L'annotation et l'interprétation des variants permettent de mettre en évidence des mutations dans des gènes déjà connus ou des mutations dans des gènes considérés comme candidats dont l'implication dans la pathologie n'est pas encore avérée. Il peut s'agir de gènes dont le lien avec le cil a été évoqué précédemment par des expériences de transcriptomiques, protéomiques ou toutes autres analyses globales. Par exemple, les protéines identifiées au cours d'études ayant permis de déterminer la composition du ciliome sont pour certaines codées par des gènes connus et d'autres par des gènes candidats. Le ciliome correspond à un ensemble de 2500 gènes/protéines à l'heure actuelle (représentant plus de 10% de nos gènes) impliqués dans le fonctionnement du cil primaire (Bachmann-Gagescu 2014). Ces informations sont présentes dans des bases de données telles que Syscilia (<http://syscilia.org/>), CilDB (<http://cildb.cgm.cnrs-gif.fr/>). Le WES permet également d'identifier de nouveaux gènes dans des cas où la pathogénicité des variants dans des gènes connus a été écartée et où les variants filtrés restants sont présents dans des gènes dont le rôle est totalement éloigné du cil primaire. C'était le cas pour *VPS15* à l'époque : la caractérisation moléculaire et le recours à des modèles d'étude de ce gène est nécessaire pour valider l'implication d'un gène dans les ciliopathies.

Même si le WES est une technique qui s'est démocratisée au sein des laboratoires de recherche, l'analyse par WES présente quelques limitations qu'il faut conserver à l'esprit : cette technique ne permet pas de détecter des mutations non codantes, des translocations équilibrées (échanges réciproques de matériels entre 2 chromosomes non-homologues). Les variations du nombre de copies (CNV) sont difficilement mises en évidence. Les régions difficiles du génome humain (pseudogènes, régions riches en GC, séquences répétées) sont également peu couvertes par cette analyse : le nombre

de lecture de ces séquences est faible et par conséquent la détection de mutations dans ces régions est limitée (Oud, Lamers, et Arts 2017).

➤ Le séquençage du génome complet (WGS)

Il s'agit de la technique NGS la plus complète car elle permet de réaliser le séquençage complet de l'ensemble du génome (régions codantes et non codantes). En plus des régions exoniques qui sont très bien annotées et interprétées, le WGS donne accès aux variants présent dans les régions promotrices proximales, les régions régulatrices distales appelées « enhancers » (séquences génomiques permettant l'activation transcriptionnelle d'un gène à distance). Des progrès sont nécessaires pour l'annotation de ces séquences non codantes dont l'exploration commence à se généraliser.

Contrairement au TES et au WES, le WGS ne nécessite pas d'étape de capture (étape permettant la sélection des régions séquencées). Comme je l'ai déjà mentionné, les régions plus complexes (riches en GC, hautement répétées) ne sont pas aussi bien capturées que le reste de l'exome et la couverture n'est pas homogène. Le WGS s'affranchit de cette étape et la couverture du génome est donc plus homogène. Cette approche se perfectionne encore avec la génération de séquences plus longues permettant de mettre en évidence plus facilement les variations structurales (Meienberg et al. 2016). Les fournisseurs de ces technologies (Illumina par exemple) proposent des prix de plus en plus attractifs ce qui explique la popularité grandissante pour le recours au WGS dans l'identification des mutations génétiques.

## V. Les modèles d'étude pour les ciliopathies

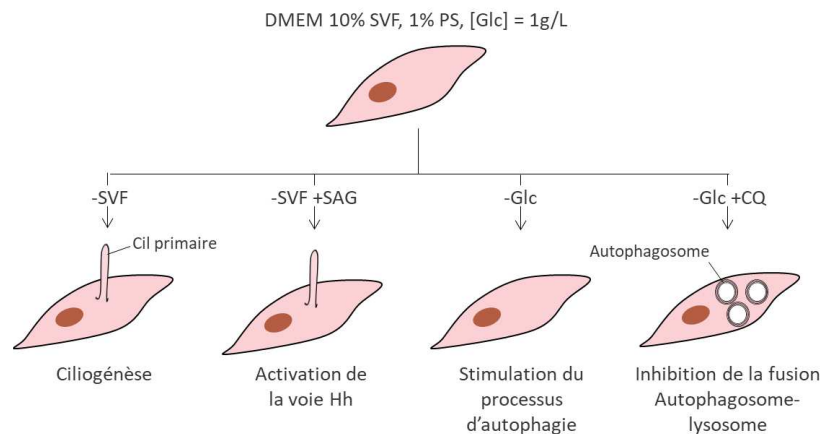
Dans le schéma général de la recherche translationnelle (Figure 19), le diagnostic génétique est suivi d'une validation fonctionnelle dans le cas de variants de signification inconnue (classe 3) pour lesquels la preuve de pathogénicité doit être montrée. Cette étape nécessite le recours à des modèles d'études : (1) cellulaires *in vitro*, à partir de fibroblastes maintenus en culture, (2) animaux *in vivo*, certaines lignées de souris et de poisson-zèbres modèles de ciliopathie ou de BBS ont d'ailleurs été mentionnées précédemment. Ce sont des outils ayant permis d'étendre les connaissances portant sur les gènes et les voies physiopathologiques impliquées. Le choix du modèle va dépendre de la conservation du gène entre l'humain et les espèces considérées (gènes orthologues), de sa fonction, de la nécessité ou pas d'étudier un tissu dans sa globalité, des outils moléculaires disponibles pour ce modèle, du temps et du budget pouvant être investis par les scientifiques (Molinari et Sayer 2020).

Dans la suite de cette partie, j'ai choisi de restreindre mon propos aux 3 modèles d'études utilisés aux cours de la thèse : les modèles cellulaires, les modèles poisson-zèbres et les modèles murins. Pour ce dernier modèle, je présenterais plus précisément les 3 organes explorés chez les souris *Vps15* : l'œil (notamment la rétine), le rein et le cerveau (notamment l'hippocampe).

### 1. Les modèles cellulaires

La culture de cellules est un moyen peu coûteux utilisant du matériel facilement manipulable afin de disséquer les mécanismes pathologiques responsables d'une maladie. Un phénotype peut s'avérer différent entre 2 patients ayant le même génotype, ce qui montre l'impact majeur du fond génétique. Les fibroblastes sont des cellules retrouvées de manière quasi ubiquitaire dans le corps humain. Ils sont facilement accessibles par biopsies de peau. Ce geste est peu invasif et indolore (la procédure est décrite dans Scheidecker et al. 2015) et permet de récupérer des fibroblastes pour la mise en place d'une culture primaire. Les fibroblastes forment un modèle humain simple, individu-spécifique. Quand cela est possible, l'accès aux fibroblastes de patient est un avantage puisque les cellules possèdent déjà la mutation dans le gène d'intérêt avec un fond génétique donné. Dans le cas contraire où les fibroblastes ne sont pas accessibles, des expériences permettant de reproduire la mutation dans des cellules ayant un fond génétique différent doivent être mises en place (Molinari et Sayer 2020). Il faut établir en parallèle des lignées de fibroblastes d'individus contrôles en privilégiant des apparentés sains qui ne présentent pas la mutation à l'état hétérozygote ou des personnes de même âge, même sexe, même origine ethnique. Une fois la culture primaire de fibroblastes établie, les cellules sont soumises à des conditions de culture qui varient en fonction de la question posée. Comme le résume la Figure 24, la culture des fibroblastes dans des conditions déprivées en glucose (-Glc) permet de stimuler le processus d'autophagie. L'ajout de chloroquine (CQ) dans le milieu de culture permet l'inhibition des phases tardives de l'autophagie. Une condition de culture sans sérum (-SVF) pendant

24h permet d'induire la ciliogénèse et d'étudier les cellules ciliées. C'est une condition très utilisée dans le cadre de l'étude des ciliopathies. En effet, il est possible de confirmer un phénotype de ciliopathie en mesurant la taille du cil ou en déterminant la fréquence de ciliation des cellules. L'ajout de SAG dans le milieu permet d'activer la voie SHH : une voie dont le fonctionnement est fréquemment altéré dans le cadre des ciliopathies.



**Figure 24 : Les conditions de culture des fibroblastes.** Représentation schématique récapitulant les conditions de culture des fibroblastes de peau utilisées au cours de cette thèse pour étudier différents mécanismes.

## 2. Les modèles poisson-zèbres

Pour générer ces modèles, le Laboratoire de génétique médicale a établi une collaboration de longue date avec l'équipe du Pr Uwe Strähle et notamment le Dr Christelle Etard basée au Karlsruhe Institute of Technology (KIT, Karlsruhe, Allemagne).

Avec ~25000 gènes (GRCz11, ensembl 100.11), le poisson-zèbre ou *Danio rerio* partage des orthologues fonctionnels avec 70% des gènes humains. Plusieurs avantages peuvent être avancés pour justifier l'étude de ces poissons. Ils sont transparents, leur organogénèse est rapide. Ils sont de petite taille, leur taux de fécondité est élevé et le coût de leur maintenance est réduit par rapport à d'autres modèles animaux. Pour mettre en place l'étude d'un gène et des voies de signalisations associées, il est primordial de vérifier que l'orthologue est présent dans cette espèce et qu'il n'est pas dupliqué. En effet, une duplication du génome entier des poissons téléostéens dont fait partie le poisson-zèbre a eu lieu très tôt dans l'évolution de leur branche et certaines duplications de gènes ont été maintenues au cours du temps (Molinari et Sayer 2020). Pour caractériser un gène, il est possible d'étudier les conséquences de sa surexpression par micro-injection d'ARNm. A l'inverse, l'injection de morpholinos ou oligonucléotides anti-sens permet d'atténuer l'expression de gènes dès le stade embryonnaire et d'étudier les conséquences. Plus récemment, la démocratisation du système CRISPR-Cas9 (une endonucléase d'ADN qui induit des coupures double-brin d'ADN associée à un ARN guide qui permet la création de mutations dans les gènes cibles) a conduit à l'établissement de lignées Knock-Out (KO)

de poisson-zèbres. Les morpholinos et le système CRIPR-Cas9 sont très utilisées. Ils sont rapidement applicables et peuvent être couplés avec l'injection d'ARNm du gène WT ou muté dans le cadre d'expériences de rescue/sauvetage qui permettent de confirmer la pathogénicité de nouvelles mutations à valider. Il existe des lignées transgéniques permettant après croisements de visualiser, par fluorescence, des structures précises comme le cil (Molinari et Sayer 2020).

### 3. Les modèles murins

La souris, *Mus musculus*, est un mammifère de petite taille qui partage 70% des gènes codants des protéines chez l'homme. C'est l'organisme modèle des vertébrés le plus utilisé pour les recherches portant sur la génétique médicale. La caractérisation des multiples modèles murins existants a contribué à la clarification des mécanismes pathologiques de nombreuses maladies ainsi qu'à l'identification de cibles thérapeutiques potentielles. Depuis la première lignée de souris KO rapportée en 1986, des milliers d'autres modèles ont été générés par différents consortiums dont la mission est d'assigner une fonction à chacun des gènes du génome murin (Barbaric, Miller, et Dear 2007). L'essor de ces modèles est dû à l'utilisation d'outils permettant la mise en place d'une mutagenèse dirigée par recombinaison homologue en utilisant les cellules souches embryonnaires. Depuis peu, les technologies d'éditions de gènes, telle que CRISPR-Cas9, ont simplifié et rendu accessible la génération de modèles KO et Knock-in (KI).

#### A) Comparaison des tissus humains et murins

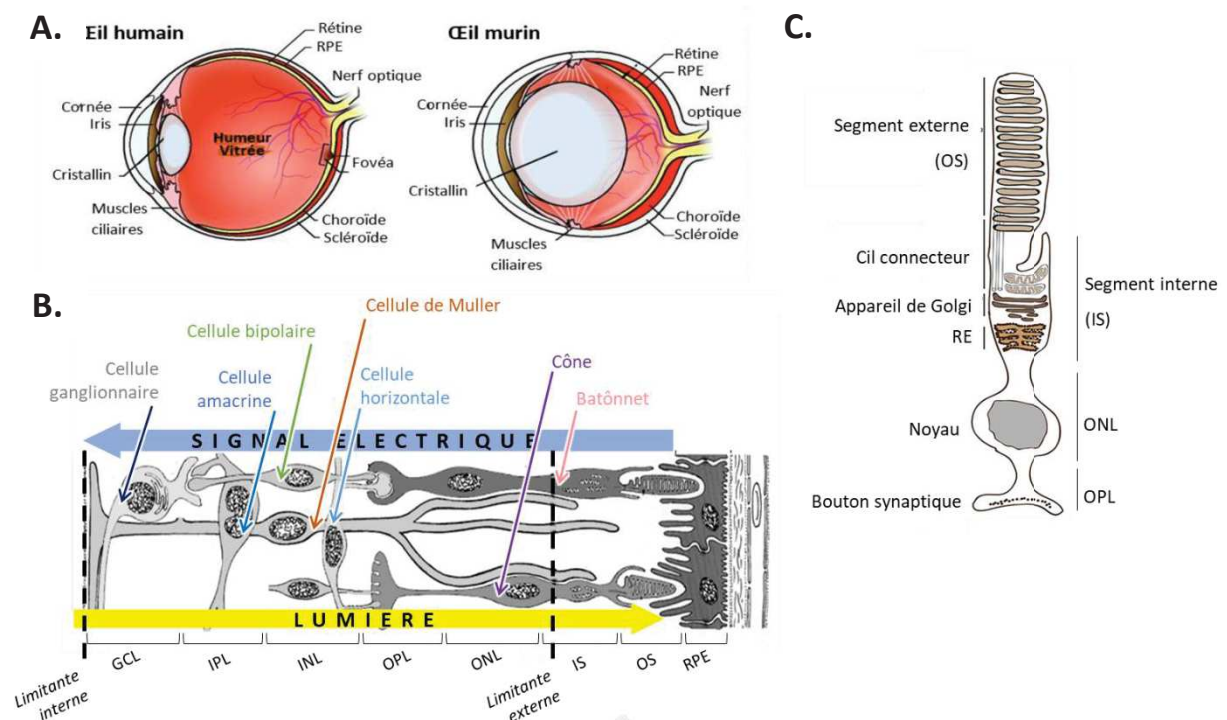
##### ➤ L'œil

Il s'agit d'un organe d'origine neurectodermique dérivé de la placode optique, elle-même issue du diencephale. Cet organe senseur est composé de différents tissus dont la rétine. Comme le résume la Figure 25, elle tapisse le fond de l'œil et se compose de 10 couches contenant 3 types cellulaires majeurs : les photorécepteurs, les cellules bipolaires et les cellules ganglionnaires dont le regroupement des axones forme le nerf optique. Les photorécepteurs réalisent la phototransduction : transformation du signal lumineux (flux photonique) en signal électrique (potentiel d'action) transmitt aux cellules bipolaires, elles-mêmes en contact avec les cellules ganglionnaires. Les photorécepteurs contiennent un cil spécialisé composé du « cil connecteur » qui fait le lien entre le segment externe (OS) et le segment interne (IS) de la cellule (Sjöstrand 1953). Il permet la prise en charge des protéines entre les 2 segments, le long des MT, grâce à des protéines telles que les IFT. L'OS est la zone d'empilement des pigments photosensibles (rhodopsine) qui détectent la lumière. L'IS contient les organites et les ressources énergétiques nécessaires à la production de nouvelles protéines pour réaliser la phototransduction (R. L. Collins et al. 2020). Il existe 2 types de photorécepteurs : les cônes et les bâtonnets. Les cônes sont peu sensibles à la lumière et permettent la vision en couleur (photopique). 3 types de cônes ont été recensés chez l'homme en fonction des longueurs d'ondes

auxquelles ils sont sensibles (S, M ou L). La souris en possède seulement 2 types (S et M). Les souris sont dichromates et ne sont pas sensibles à la couleur rouge. Il existe un seul type de bâtonnets pour les 2 espèces. Ils sont très sensibles à la lumière et permettent la vision en noir et blanc (scotopique) (Veleri et al. 2015).

La Figure 25 souligne les différences anatomiques existant entre les humains et les souris. Chez l'homme, les 6 millions de cônes sont tous concentrés au niveau de la fovéa : une zone hautement spécialisée pour la vision précise. Les 100 millions de bâtonnets sont répartis partout ailleurs dans la rétine. En comparaison, les souris ont un cristallin occupant un important volume, au détriment de l'humeur vitrée (=uvée). Comme la plupart des mammifères, leur rétine contient une plus grande proportion de bâtonnets par rapport aux cônes, mais ces derniers sont répartis dans tout le tissu à la différence de l'homme. Les souris n'ont pas de fovéa (Huberman et Niell 2011).

L'ONL (couche nucléaire externe) est la zone contenant les noyaux des photorécepteurs. En mesurant l'épaisseur et en comptant le nombre de noyaux, il est possible de mettre en évidence une diminution du nombre de photorécepteurs. Les synapses entre les photorécepteurs et les cellules bipolaires sont situées dans l'OPL (outer plexiform layer). Elles contiennent du glutamate : un neurotransmetteur associé aux hyperpolarisations (R. L. Collins et al. 2020).



**Figure 25 : Représentation schématique de l'œil et de la rétine.** A. Comparaison entre l'œil humain et l'œil de la souris. B. Annotation des 10 couches de la rétine et des différents types cellulaires. C. Annotation des 3 parties d'un photorécepteur (adapté de Athanasiou et al. 2013; Veleri et al. 2015) GCL : couches des cellules ganglionnaires, IPL : couche plexiforme interne, INL : couche nucléaire interne, OPL : couche plexiforme externe, ONL : couche nucléaire externe, IS : segment interne/OS : segment externe des photorécepteurs, RPE ; épithélium pigmentaire rétinien.



La rétinopathie pigmentaire (RP) est un des signes cliniques majeurs des ciliopathies. Il s'agit d'une dégénérescence de la rétine causée par la perte progressive des photorécepteurs (Athanasiou et al, 2013). Le fond d'œil d'un patient présentant une telle anomalie est très pigmenté en raison de l'amincissement de la couche des photorécepteurs laissant apparaître les pigments de l'épithélium pigmentaire. Cette dégénérescence progressive des photorécepteurs explique la perte progressive de la vue, jusqu'au stade ultime à savoir la cécité. Les RP sont caractérisées par l'atteinte des 2 types de photorécepteurs : d'abord des bâtonnets puis des cônes (Veleri et al. 2015). L'étude de la rétine et des cellules qui la composent constitue un enjeu majeur dans la compréhension des mécanismes physiopathologiques sous-jacents.

➤ Le rein

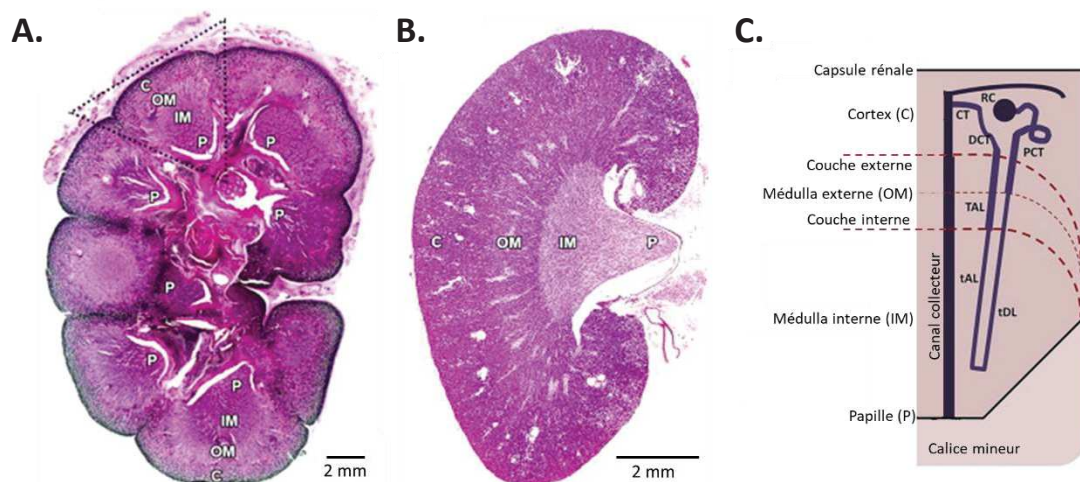
Le rein est un organe essentiel qui permet d'extraire une partie des déchets de l'organisme par la sécrétion d'urine. La définition d'un néphron correspond à une unité fonctionnelle du rein et se compose de portions ayant chacune une organisation cellulaire particulière, associée à des tâches précises (Figure 26C). Le glomérule, composant majeur du corpuscule rénal, permet la filtration de l'urine primitive à partir du sang. La fente glomérulaire est formée à partir des pédicelles des podocytes. L'urine primitive est récupérée dans la lumière de la capsule de Bowman qui est délimitée par un épithélium. Les tubules proximaux et distaux permettent la réabsorption depuis l'urine vers le sang grâce aux nombreux replis membranaires des cellules (riches en mitochondries et en transporteurs). La concentration-dilution de l'urine se réalise au niveau des anses de Henlé : zones riches en aquaporines. Les canaux collecteurs drainent l'urine définitive en l'acheminant jusqu'à la papille puis les calices et enfin vers l'uretère qui rejoint la vessie.

Si l'on compare le développement du rein entre l'homme et la souris, il existe des différences en termes de temps nécessaire à la formation de l'organe durant l'embryogénèse mais le résultat est globalement similaire sur le plan de l'organisation et du fonctionnement. Le rein humain est multilobé (Figure 26A) et possède environ 1 million de néphrons : un nombre établi dès la naissance (Murawski, Maina, et Gupta 2010). Il pèse en moyenne 140g et produit par jour environ 180L d'urine primitive. Le cheminement de ce fluide le long des néphrons aboutit à la formation de 1,5-2L d'urine définitive ayant une composition normalisée (Tanner, Brown, et Muntner 2012). Des anomalies lors d'analyses urinaires de patients (protéinurie, glycosurie, polyurie...) aiguillent le clinicien sur la fonction et donc la portion du néphron qui dysfonctionne. La version réduite du rein murin pèse 18,6g, compte environ 16000 néphrons, possède un seul lobe (Figure 26B). Il permet l'excrétion de 1mL d'urine définitive par jour (Lindström et al. 2018; Mazzaccara et al. 2008). Les souris présentent de manière physiologique une protéinurie caractérisée par une quantité d'albumine élevée (Mazzaccara et al. 2008) et l'organe



présente un dimorphisme sexuel au niveau des cellules de la capsule de Bowman. En effet, les mâles présente un épithélium simple cubique, contrairement aux femelles pour lesquelles l'épithélium est aplati (Yabuki et al. 1999). Par contre, les femelles des 2 espèces présentent un nombre plus restreint de néphrons par rapport aux mâles.

Dans le rein, les cellules épithéliales sont pourvues à leur pôle apical d'un cil primaire ayant un rôle mécano-senseur nécessaire au maintien de l'homéostasie. Les ciliopathies sont caractérisées par une atteinte rénale dont la sévérité est telle qu'elle constitue la cause majeure de décès chez les patients. Par exemple, certains patients présentant une dérégulation des voies de signalisations telles que JAK-STAT, WNT, la voie médiée par l'AMPC ou par le  $Ca^{2+}$ . Une activation anormale de la voie WNT non canonique cause une prolifération cellulaire accrue qui conduit à la formation de kystes (Jonassen et al. 2008).



**Figure 26 : Coupes transversales de reins. A.** humain. **B.** murin. **C.** Représentation schématique annotée des différentes couches rénales et des différentes portions d'un néphron. (adapté de Cullen-McEwen, Sutherland, et Black 2016) RC : corpuscule rénal, PCT : tubule contourné proximal, tDL : anse de Henlé/branche fine descendante, TAL : anse de Henlé/branche fine ascendante, TAL : anse de Henlé/ branche large ascendante, DCT : tubule contourné distal, CT : tubule collecteur.

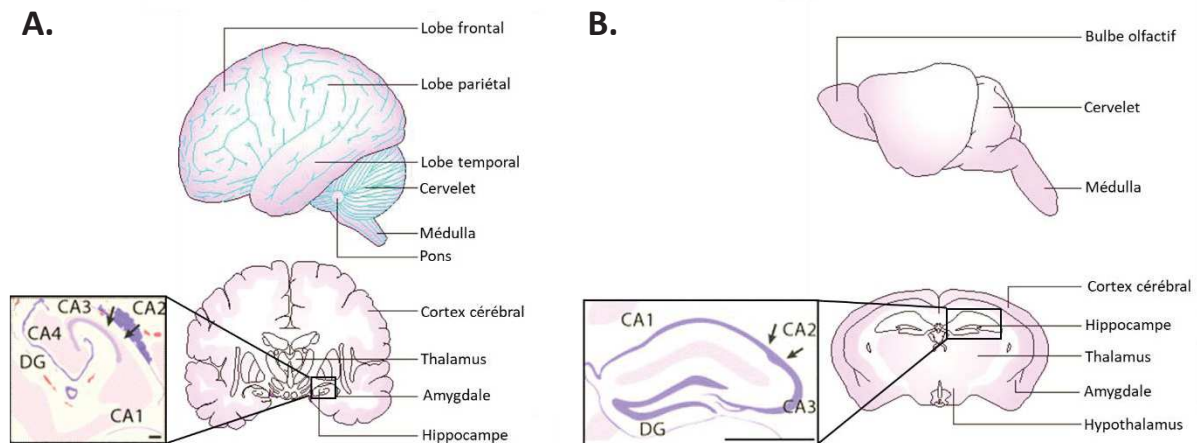
➤ Le cerveau

Il s'agit d'une structure formée à partir du tube neural, lui-même issu d'un repliement du feuillet embryonnaire ectodermique (Semple et al. 2013). Les différentes parties du cerveau ont évolué à partir de 3 puis de 5 bourgeons encéphaliques. Le télencéphale est à l'origine des hémisphères cérébraux (lobes frontaux, pariétaux, temporaux, bulbe olfactif). Le diencephale a donné le thalamus et l'hypothalamus. Le mésencéphale a donné naissance aux pédoncules du tronc cérébral. Le métencéphale est à l'origine du cervelet et du pons. Le myélocéphale a donné la médulla. L'organogénèse cérébrale commence au cours de la gestation : elle dure 34 semaines chez l'humain et 11 jours chez la souris. La neurogénèse continue après la naissance : la maturation et l'établissement de certaines connexions neuronales se poursuivent jusqu'à l'adolescence (Zhao et Bhattacharyya 2018). Les modèles de rongeurs et plus particulièrement les souris se sont imposées depuis ces

dernières années comme un modèle d'étude majeur pour les maladies neurodégénératives. Ce modèle a effectivement permis de modéliser certains aspects de la physiologie et du fonctionnement cérébrale face aux similitudes mises en évidence. Par exemple, les étapes de bases aboutissant à la formation des neurones inhibiteurs (interneurones) et excitateurs sont conservées. Par contre, chez l'homme le nombre de cellules, de types cellulaires et de connections établies sont beaucoup plus importants et diversifiés. Il s'agit d'une espèce dont la neuroanatomie est plus complexe et sous-tend des capacités plus spécialisées comme le langage et la conscience par exemple. Ces fonctions spécifiques sont également mises en avant sur les profils d'expressions géniques établis entre humains et souris qui montrent des différences pour certains groupes de gènes (Zhao et Bhattacharyya 2018). Pour les gènes orthologues, les séquences nucléotidiques et protéiques sont conservées et les patrons d'expression concordent (Strand et al. 2007). Globalement, les stades du développement cortical ont été conservés au cours de l'évolution des espèces mammifères. La mise en place de cette structure en 6 couches forme un empilement construit grâce à la migration des neurones depuis les couches les plus internes vers la superficie. Le cortex ainsi formé est impliqué dans les fonctions cognitives des mammifères tels que la perception sensorielle, les mouvements moteurs (Zhao et Bhattacharyya 2018). Le cortex humain, beaucoup plus gros que celui des souris, est caractérisé par la présence de circonvolutions qui sont absentes chez cette 2<sup>ème</sup> espèce (Semple et al. 2013). La forme et les différentes portions du cerveau adulte pour les 2 espèces sont récapitulées dans la Figure 27.

Afin d'analyser les différentes fonctions cérébrales chez les souris, plusieurs analyses sont mises en place. Par exemple, les capacités d'interaction sociale sont étudiées en réalisant des tests de préférence et des tests de reconnaissance d'individus. La communication est appréhendée par l'enregistrement des vocalisations. Les capacités cognitives sont analysées grâce à des tests de mémoire et d'apprentissage tels que la reconnaissance d'objet, la piscine de Morris, le Y-Maze (Zhao et Bhattacharyya 2018).

L'hippocampe est une structure cérébrale, d'origine télencéphalique, considéré comme le siège de la mémoire (V. S. Chen et al. 2017). Il s'agit d'une région sous corticale localisée au niveau du lobe temporal formé à partir de 2 structures en U inversées qui sont le gyrus denté (DG) et la corne d'Amon (CA), elle-même divisée en 3 parties (CA1, 2 et 3) (Figure 27). Le flux d'information dans l'hippocampe est unidirectionnel et forme une boucle. Les cellules granulaires du DG reçoivent l'information du cortex entorhinal et la transmettent aux cellules pyramidales de CA3. Ces dernières sont en contact avec les cellules pyramidales de CA1 qui envoient les informations au subiculum qui est relié au cortex entorhinal (Molinari et Sayer 2020). En cas d'atteinte de l'hippocampe, les patients présentent des troubles cognitifs avec perte de mémoire, associés à d'autres troubles cérébraux mettant en évidence les connexions entre les diverses régions cérébrales.



**Figure 27 :** Représentation schématique de cerveau A. humain B. murin et d'une coupe coronale où l'hippocampe est visualisé et agrandi. CA : Corne d'Amon, DG : Girus denté (inspiré de Cryan et Holmes 2005; Brunner 2016).

De nombreux modèles murins ont été générés et phénotypés afin de mettre en place des corrélations génotype-phénotype dans le cadre de l'étude des ciliopathies. J'ai récapitulé de manière non exhaustive, certains modèles murins BBS dans le Tableau 1. Je rapporte également les résultats portant sur les modèles murins *Vps15* sur lesquels je me suis basée pour définir les âges, les effectifs et les expériences à mettre en place pour débiter la caractérisation du nouveau modèle murin KI *Vps15*.

#### B) Les modèles murins *Vps15*

##### ➤ Le modèle KO *Vps15* de Nemazanyy et al. 2013

Afin de déterminer l'effet de l'inactivation de *Vps15* sur un organisme entier, cette équipe a généré et étudié le premier modèle KO pour *Vps15*, par recombinaison homologue et technique Cre-Lox en ciblant l'exon 2 du gène. Il faut noter que le modèle KO total pour *Vps15* de cette étude (Nemazanyy et al. 2013) présente une létalité embryonnaire donc il n'a pas été étudié. C'est finalement un modèle KO conditionnel et spécifique du muscle qui a été étudié. L'absence de protéines *Vps15* dans le muscle a été confirmé par WB. Une diminution de la quantité protéique de *Vps34* et *Beclin-1* a été mise en évidence, tout comme une accumulation des protéines *P62* et *LC3* associée à une quantité accrue de protéines ubiquitinylées et une accumulation d'autophagosomes et de lysosomes dans le cytoplasme des cellules musculaires. La souris KO *Vps15* dans le muscle présente une AVM (autophagy vacuolar myopathy) : un phénotype cohérent avec la perte de fonction totale de l'activité du complexe PI3K de classe III, dans cet organe.

➤ Le modèle *Vps15* de Gstrein et al. 2018.

Ce deuxième modèle murin pour *Vps15* a été induit par la technique d'ENU (N-éthyl-N-nitrosourea). Par cette technique, les scientifiques créent des mutations de manière aléatoire dans le génome murin. Il s'agit d'une technique de créations de lignées murines inversée par rapport aux techniques habituelles : les phénotypes des lignées sont d'abord caractérisés puis une analyse génétique par WES est mise en œuvre afin d'identifier la mutation et le gène responsable du phénotype observé.

Dans le cas de cette étude (Gstrein et al. 2018), une substitution de nucléotides (T>A) dans l'intron 4 de *Vps15* a été mise en évidence dans le cas de la première lignée étudiée dite *marble*, sachant que 2 autres lignées murines ont également été analysées. Selon les logiciels de prédictions, cette mutation crée un nouveau site accepteur d'épissage et aboutie à l'apparition d'un codon stop prématuré. En comparaison de la protéine de 1358 AA attendue normalement, une protéine tronquée de seulement 289 AA peut être produite (qui n'est pas détectée par WB). Il s'agit d'une mutation hypomorphe (entraînant une forte diminution de l'expression de la protéine normale) qui provoque une désorganisation de la couche des cellules pyramidales notamment au niveau de la région CA1 de l'hippocampe. Elle est associée à une altération de la mémoire de travail. Les souris homozygotes de la lignée *marble* présentent des quantités fortement réduites de protéines *Vps15*, *Vps34* et *Beclin-1* au niveau des régions hippocampiques.

Le deuxième modèle étudié correspond à des embryons pour lesquels la déplétion de *Vps15* est totale. L'étude a été faite avant leurs naissances et a révélé un défaut de migration neuronale sévère.

Enfin, le dernier modèle murin rapporté est un KO conditionnel *Vps15* au niveau du télencéphale (partie la plus frontale du cerveau en cours de développement) qui montre une dégénérescence progressive sévère de l'hippocampe et du cortex, associée à une accumulation des protéines LC3 et P62 et une stimulation de l'apoptose.

Cet article rapporte la 2<sup>ème</sup> famille mutée dans *VPS15* après l'article Stoetzel et al. 2016. Par contre, le phénotype des patients est très éloigné de celui d'une ciliopathie. En effet, l'individu est porteur homozygote de la mutation p.L1224R dans *VPS15* et présente une sévère atrophie du cortex et du nerf optique, une déficience intellectuelle, une ataxie, un retard psychomoteur, une épilepsie. La caractérisation moléculaire menée sur ces cellules a révélé une baisse des quantités protéiques de *VPS15*, *VPS34* et *Beclin-1* et une augmentation de la quantité de P62.

Génotype	<i>Bbs1</i> <sup>M390N/M390R</sup>	<i>Bbs2</i> <sup>-/-</sup>	<i>Bbs4</i> <sup>-/-</sup>	<i>Bbs10</i> <sup>-/-</sup>	<i>Bbs10</i> <sup>-/-</sup>	<i>Bbs12</i> <sup>-/-</sup>	<i>Bbs17</i> <sup>-/-</sup>	<i>Cep290</i> <sup>-/-</sup>
Référence	Davis & al. 2007	Nishimura & al. 2004	Mykytyn & al. 2004	Cognard & al. 2015	Yu & al. 2016	Marion & al. 2013	Datta & al. 2015	Yu & al. 2016
Anomalies rétinienne	<input checked="" type="checkbox"/> - MET dès 5M - H&E dès 6W - ERG dès 11W	<input checked="" type="checkbox"/> - ERG dès 5W - H&E dès 7W - IHC $\alpha$ -rhodopsine dès 5W - TUNEL dès 5M - MET dès 5M	<input checked="" type="checkbox"/> - H&E dès 7M - IHC $\alpha$ -opsine dès 6W - ERG dès 6W - TUNEL dès 6W	<input checked="" type="checkbox"/> - OCT dès 3M - Bleu de toluidine dès 3M - MET dès 3M - ERG dès 8W	<input checked="" type="checkbox"/> - ERG dès 3M - Bleu de toluidine dès 3M - IHC $\alpha$ -rhodopsine dès 2W - MET	<input checked="" type="checkbox"/> - IHC $\alpha$ -rhodopsine dès 3M	<input checked="" type="checkbox"/> - H&E dès 2M - ERG dès 7W - MET dès P45	<input checked="" type="checkbox"/> - ERG dès 1M - Bleu de toluidine dès 2W - IHC $\alpha$ -rhodopsine dès 2W - MET
Anomalies rénales	<input checked="" type="checkbox"/> Pas de kyste	<input checked="" type="checkbox"/> Kystes - H&E dès 5M.	-	Pas de kyste - H&E à 3M <input checked="" type="checkbox"/> Hyperalbuminurie et polyurie	-	<input checked="" type="checkbox"/> Pas de kyste	-	-
Obésité	<input checked="" type="checkbox"/> Dès 3M	<input checked="" type="checkbox"/> Dès 4M (femelles) Dès 5M (mâles)	<input checked="" type="checkbox"/> Dès 3M	<input checked="" type="checkbox"/> Dès 3M	-	<input checked="" type="checkbox"/> Dès 10W	<input checked="" type="checkbox"/> Dès 4M	-
Hyper Leptinémie	<input checked="" type="checkbox"/>	-	-	<input checked="" type="checkbox"/> Dès 3M	-	-	-	-
Dépôts de graisse	<input checked="" type="checkbox"/> - IRM dès 3,5M	-	<input checked="" type="checkbox"/> Dès 10M	<input checked="" type="checkbox"/> Dès 3M	-	-	<input checked="" type="checkbox"/> - IRM dès 6M	-
Hyperphagie	<input checked="" type="checkbox"/> Dès 10W.	<input checked="" type="checkbox"/> Dès 10W.	<input checked="" type="checkbox"/> Dès 9W.	<input checked="" type="checkbox"/>	-	<input checked="" type="checkbox"/> Dès 8W.	-	-
Perte d'activité	<input checked="" type="checkbox"/>	-	-	-	-	-	-	-
Ventriculo mégalie	<input checked="" type="checkbox"/> - IRM coronaires dès 3,5M	-	-	-	-	-	<input checked="" type="checkbox"/> - IRM dès 6M	-
Rétrécissement Striatum et Hippocampe	<input checked="" type="checkbox"/> - coupes coronaires	-	-	-	-	-	-	-
Affinement du cortex cérébral	<input checked="" type="checkbox"/> - coupes coronaires	-	-	-	-	-	-	-
Infertilité	<input checked="" type="checkbox"/> - H&E dès 6M	<input checked="" type="checkbox"/> - H&E	<input checked="" type="checkbox"/> - H&E et Bleu de toluidine dès P45	-	-	-	-	-
Anomalies des cils motiles 9+2	<input checked="" type="checkbox"/> Pas d'anomalie des cellules épendymaires - MET à 7M	<input checked="" type="checkbox"/> Pas d'anomalies des cellules trachéales - MET	<input checked="" type="checkbox"/> Anomalies des cellules rénales - MET	-	-	-	-	-
Polydactylie	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	-	-	-	-	-

**Tableau 1 : Présentation de certains modèles murins BBS.** Annotation des âges d'apparition et des techniques réalisées pour déterminer le phénotype de 8 modèles murins Bbs.  : présence,  absence, - : non étudié, MET : microscopie électronique à transmission, H&E : coloration hématoxyline-éosine, ERG : électrorétinogramme, IHC : immunohistochimie, IRM : imagerie à résonance magnétique, M : mois, W : week (semaine), P45 : 45 jours après la naissance

## VI. Les problématiques

Mon projet de thèse s'inscrit dans la recherche translationnelle appliquée aux ciliopathies et se décompose en 2 grandes parties qui abordent toutes les étapes de cette recherche, par l'intermédiaire de plusieurs projets. Les études que je vous présente décrivent toujours la démarche scientifique mise en place entre le diagnostic clinique, l'identification des variations dans les données de séquençage et l'étude fonctionnelle de leur impact sur les voies de signalisation attenantes au cil primaire en utilisant les 3 modèles d'études présentés.

La première partie de ma thèse aborde le cas de familles présentant des phénotypes typiques de ciliopathies pour lesquelles le diagnostic génétique n'avait initialement pas encore été résolu : soit parce que les variants identifiés étaient considérés comme des VUS (classe 3), soit parce que le 2<sup>ème</sup> allèle n'avait pas été identifié avec les techniques NGS mises en place en 1<sup>ère</sup> intention et devait encore être déterminé. Il s'agissait donc de mettre en place les études moléculaires et les validations fonctionnelles afin de déterminer la pathogénicité et l'impact fonctionnel de mutations identifiées dans certains gènes de ciliopathies déjà connus. Cette première partie se décomposera en 4 projets dont 3 portent sur des gènes responsables de BBS (*IFT27*, *BBS1* et *BBS5*) pour lesquels les validations fonctionnelles ont été conduites à partir des fibroblastes des patients mis en culture. Le dernier projet porte sur le cas particulier du gène *IQCE* : un gène déjà rapporté dans une seule famille comme responsable de polydactylie post-axiale (PAP), une ciliopathie non syndromique pour laquelle les patients présentent un phénotype ressemblant à celui d'une ciliopathie syndromique. Pour cette partie, nos recherches ont été réalisées sur les cellules des patients et un modèle d'étude animal : le poisson-zèbre. A ce stade, 2 des 4 projets sont publiés :

- Identification and Characterization of Known Biallelic Mutations in the *IFT27* (*BBS19*) Gene in a Novel Family With Bardet-Biedl Syndrome. Schaefer E, **Delvallée C**, Mary L, Stoetzel C, Geoffroy V, Marks-Delesalle C, Holder-Espinasse M, Ghoumid J, Dollfus H and Muller J. *Frontiers in Genetics* (2019).
- Novel *IQCE* variations confirm its role in postaxial polydactyly and cause ciliary defect phenotype in zebrafish. **Delvallée C\***, Estrada-Cuzcano A\*, Etard C\*, Stoetzel C, Schaefer E, Scheidecker S, Geoffroy V, Schneider A, Studer F, Mattioli F, Chennen K, Sigaudy S, Plassard D, Poch O, Piton A, Strahle U, Muller J and Dollfus H. *Human Mutation* (2020).

Le projet portant sur *BBS1* est en révision dans *Human genetics* :

- An SVA F retrotransposon insertion is a frequent cause of *BBS1* mutation. **Delvallée C**, Nicaise S, Antin M, Leuvrey AS, Nourisson E, Leitch C.C, Kellaris G, Stoetzel C, Geoffroy V, Scheidecker S, Keren B, Depienne C, Klar J, Dahl N, Deleuze JF, Génin E, Redon R, Demurger F, Devriendt K, Mathieu-Dramard M, Poitou-Bernert C, Odent S, Katsanis N, Mandel JL, Davis E. E, Dollfus H, Muller J.

La seconde partie de ce manuscrit porte sur l'étude fonctionnelle d'un autre modèle d'étude animal, le modèle murin *Vps15*<sup>R998Q</sup> généré au laboratoire. Le projet s'inscrit dans la continuité des travaux du laboratoire portant sur la ciliopathie humaine causée par la même mutation dans le même gène. Je vous présenterais les expériences menées sur les souris afin de déterminer (1) s'il s'agit du bon modèle d'étude pour cette ciliopathie et (2) l'impact de cette mutation chez la souris. Les 2 modèles murins déjà publiés pour *Vps15* ont été utilisés dans des études qui se sont concentrées sur une altération en lien avec le rôle canonique de *Vps15* dans le processus d'autophagie. Les organes atteints en cas de ciliopathie n'ont jamais été analysés. Le modèle KI porteur de cette mutation apparaîtrait comme un modèle très prometteur pour définir les mécanismes physiopathologiques à l'origine du phénotype de ciliopathie s'il existe. Cela permettrait de confirmer la corrélation génotype-phénotype, dans le cadre des ciliopathies, avec un modèle d'étude supplémentaire.

# II.

## Matériel et méthodes



# I. Le matériel

## 1. Tableau 2 : Listes des amorces

Gene	Nom	Exon	Amorces sens (5'-3')	Nom	Exon	Amorces anti-sens (5'-3')	Taille (pb)
<b>Amorces utilisées lors de la partie 1</b>							
<b>PROJET 1 portant sur <i>IFT27</i></b>							
Amorces utilisées pour l'amplification et le séquençage sanger de l'ADNg							
<i>IFT27</i>	Ex2F		CATAGGGCGTGTTCAGTGGAAGA	Ex2R		CCAAGTTGCTTGTGGCCATTCTG	600
	Ex5F		TCTACCGAGTTGGGAGTTTCACA	Ex5R		TCCCTGGGGTTAGCTTTCTGGAAT	710
Amorces utilisées pour l'amplification et le séquençage sanger de l'ADNc							
<i>IFT27</i>	Ex2F		CAGATCTTCCGAGTGATGG	Ex3-4R		GTTCCACACTGTCTCCCGT	112
	Ex3F		TGAAGACAGTGCCAGTTCCT	Ex6-7R		AGTTTCCATCTCTTTCACGGA	342
	ex6-7F		TCCGTGAAAGAGATGGAAACT	Ex7R		TCTGTCTTCTCCGGTTGTGC	146
	Ex2F		CAGATCTTCCGAGTGATGG	Ex7R		TCTGTCTTCTCCGGTTGTGC	537
	Ex1F		CCATGGTGAAGCTGGCAG	Ex3-4R		GTTCCACACTGTCTCCCGT	177
<b>PROJET 2 portant sur <i>BBS1</i></b>							
Amorces utilisées pour l'amplification et le séquençage Sanger de l'ADNg							
<i>BBS1</i>	1F	Int12	CACAGTACTCCACAAATAACTGCT	1R	Alu	CAGCCTGGGCACCATTGA	270
				2R	Int13	ATTTCCCCAGCTTTGCTGT	440
<i>BBS1 &amp; SVA-F</i>	A	Int11	CCATCCCTGTCTTGCTTTC	B	Int12	CATGCTTCATTTCACCTCC	189
	F0	TSD	CTACAAGATCCTGAAGCGTACTC	R3	SINE-R	TCTTGGGTGTTTCTCGAGA	~2300
	F4	VNT R	CTCTGCGAGAAACACCCAAG	R5	Ex13	GTCTTTCGGGGCACATTGAG	~250
	F9	SINE- R	AGTACCCAGGGACAACACT				~350
	F6	SINE- R	CAGCAGCCACCCCATCTG	R2	SINE-R	GCACCGCCCTTAATCCATTT	
Amorces utilisées pour l'amplification et le séquençage Sanger de l'ADNc							
<i>BBS1</i>	EX10-11		CCCACAAGGGGAAGAAGC	13R		TCAGGATCTTGATGATCAGG	264
				14R		AGCACGTAGGCGCAGCA	446
Amorces utilisées pour la qPCR sur l'ADNg							
<i>BBS1</i>	Ex13F		AAACTCAATGTGCCCGA	Ex13R		CTACCTAGTTGGCCGAGATCTCCT	189
	Ex14F		ACACCAACTGAGCAGGAGT	EX14R		ACGTAGGCGCAGCAGGAT	135
Amorces utilisées pour la qRT-PCR sur l'ADNc							
<i>BBS1</i>	Ex4F		TGGAACCAGGCCAAAGAG	6R		AAGGCTCCTCTGCCGTCT	85
	Ex10-11F		CCCACAAGGGGAAGAAGC	1R	Alu	CAGCCTGGGCACCATTGA	227
	Ex15F		CCCACCTTTAAGCTCACAC	17F		CTTCTCGAAGCACCAGCAC	296
<b>PROJET 3 portant sur <i>IQCE</i></b>							
Amorces utilisées pour l'amplification et le séquençage Sanger de l'ADNg							
<i>IQCE</i>	Ex12F		CTCTGAGCTCACACCAACAG	Ex12R		GTTGCCAAGGGTTCACA	399
	Ex16F		TGGTTTCTGGTGTGTCAGG	Ex13R		TTTTCAGTGGCTGGTCAGAA	382
<i>ATP6V1B 1</i>	Ex3F		AGCCGTGGGAAGTAGGTGT	Ex3R		ATGTCCATTGACCCACT	385
<i>TULP1</i>	Ex12F		GGGATGTAGGATCCCCTCA	Ex13R		CATGCCAGGAATGATGACG	481
Amorces utilisées pour la qRT-PCR sur l'ADNc							
<i>IQCE</i>	Ex3-4F		ACATCGCAAAGTCACTTATC	Ex5R		AGGCCTGGGTCAGACTT	158
<i>GSTT1</i>	Ex2-3F		ACCTTGACGGAGAGTGTTG	Ex3R		GCAGCTTCTCCGAGAGT	144
<i>SLC30A8</i>	Ex9-10F		CCATCTTACTCATGGAAGGTGTG	Ex10R		GCTGTAGCAACATGAGCTGAGA	154
<i>RSPO1</i>	Ex4F		TGGAGAGGAACGACATCCG	Ex5R		CCTTACACTTGGTGCAGAAGTTA	156

Gene	Nom	Exon	Amorces sens (5'-3')	Nom	Exon	Amorces anti-sens (5'-3')	Taille (pb)
<i>FGF9</i>		Ex2F	GGCCTGGTCAGCATTGAG		Ex3R	GTATCGCCTCCAGTGTCCAC	180
<i>SFRP1</i>		Ex1F	GCTTCTACTGGCCGAGATG		Ex2R	TGGCCTCAGATTTCAACTCGT	180
<i>MECOM</i>		Ex4F	AAACTCGAAAGCGAGAATGATCT		Ex5R	TGGTGGCGAATTAATTTGGACTT	185
<i>ALPL</i>		Ex6-7F	AACATCAGGGACATTGACGTG		Ex7R	GTATCTCGGTTTGAAGCTCTTCC	159
<i>IGF2</i>		Ex2-3F	TTCTACTTCAGCAGGCCCG		Ex3-4R	GAAGTTGTCGGGAAGCACG	168
<i>GAPDH</i>		Ex4-5F	GGAGCGAGATCCCTCCAAAAT		Ex6R	GGCTGTTGCATACTTCTCATGG	197
<i>HPRT1</i>		Ex1-2F	CCTGGCGTCGTGATTAGTGAT		Ex2-3R	AGACGTTCAGTCCTGTCCATAA	131
<i>EFCAB7</i>		Ex6F	GTCAGCAACCAGGAAGTTCA		Ex6-7R	CATGTGTTGCCAGTCCTTTATT	115
<i>EVC2</i>		Ex6-7F	CTCGTCACGGAACAGAACAC		Ex8R	CAGGTCAGCACAAGGGAGAG	135
<i>EVC</i>		Ex2F	TCGAAGGACAAGGAAGCTGT		Ex4R	TTTCATGCAGAGACGGGTTG	148
<i>GLI2</i>		Ex2F	AAGCAAGAAGCCAAAAGTGG		Ex3R	TGGTACCTTCTCTCGGTG	188
<i>GLI3</i>		Ex2F	GGCATTITTTGGTCGAAGAGA		Ex3R	GGACATTCTGTGGCTGCATA	238
<i>SMO</i>		Ex2F	TGCCAAGTGTGAGAATGAC		Ex3R	TACCAGCTTTGGGGTTGTC	229
<b>PROJET 4 portant sur BBS5</b>							
Amorces utilisées pour l'amplification et le séquençage Sanger de l'ADNg							
<i>BBS5</i>	F2		GGCACCAGGATTCTCATACC	R2		TGGGTCACCTGAAGTCAGGA	389
Amorces utilisées pour l'amplification et le séquençage Sanger de l'ADNc + duplex							
<i>BBS5</i>		Ex6-7F	TCCAGTGATCAGGGCAATTT		EX8R	TGAGAGCTTTCTATGACAAGAGC	167
Amorces utilisées pour la qRT-PCR sur l'ADNc							
<i>BBS5</i>		ex1F	ATGCCAAAGAGTGCCAGAGA		ex3R	TCCGTTTCGACCTGTCCG	153pb
		Ex5-6F	TGGCAGTACACGAGCTTATGA		ex6-7R	TGATCAGGGCAATTTAGGAACC	164pb
		ex9F	GTGGTGGATATGTTCTTGGCT		ex9-10R	CTGGGGCTTTTCTTCCATCTC	140pb
<i>IQCE</i>		Ex3-4F	ACATCGCCAAAGTCACCTTATC		Ex5R	AGGGCCTGGGTCAGACTT	158
<i>EVC</i>		Ex2F	TCGAAGGACAAGGAAGCTGT		Ex4R	TTTCATGCAGAGACGGGTTG	148
<i>EVC2</i>		Ex6-7F	CTCGTCACGGAACAGAACAC		Ex8R	CAGGTCAGCACAAGGGAGAG	135
<i>EFCAB7</i>		Ex6F	GTCAGCAACCAGGAAGTTCA		Ex6-7R	CATGTGTTGCCAGTCCTTTATT	115
<i>GLI1</i>		Ex10-11F	GGTGCCATGAAGCCACAG		Ex11R	AATGCAAGGTCCTCGTCC	162 pb
<i>GLI2</i>		Ex2F	AAGCAAGAAGCCAAAAGTGG		Ex3R	TGGTACCTTCTCTCGGTG	188
<i>GLI3</i>		Ex2F	GGCATTITTTGGTCGAAGAGA		Ex3R	GGACATTCTGTGGCTGCATA	238
<i>SUFU</i>		Ex2-3F	ACAGAGTCCATGAGTTACAGG		Ex3R	CGTATCGTGCCAAAGCCCT	138
<i>PTCH1</i>		Ex2-3F	GTGGAAGTTGGAGGACGAGT		Ex3R	AGTCCAGGTGTTGTAGGAGC	154
<i>SMO</i>		Ex2F	TGCCAAGTGTGAGAATGAC		Ex3R	TACCAGCTTTGGGGTTGTC	229
<i>GAPDH</i>		Ex4-5F	GGAGCGAGATCCCTCCAAAAT		Ex6R	GGCTGTTGCATACTTCTCATGG	197
<i>HPRT</i>		Ex1-2F	CCTGGCGTCGTGATTAGTGAT		Ex2-3R	AGACGTTCAGTCCTGTCCATAA	131
<b>Amorces utilisées lors dans la partie 2</b>							
Amorces utilisées pour le génotypage et séquençage Sanger de l'ADNg							
<i>Vps15</i>	Ef <sup>2</sup>	In13	CTGGAGTGTCTATAAGACAC	Er <sup>2</sup>	In13	CCAATCCCTGCAATCACAGTGACG	KI=462 WT=383
	Mf	In12	GTCATGGTCCAGTAGACCCCTCTG G	Mr	In13	GCACTGCTGACGTAGCTGGATTGG	466
<i>Crb1</i>	mF1	Ex9	GTGAAGACAGCTACAGTTCTGATC	mR	Ex9	GCCCCATTTGCACACTGATGAC	220
	mF2	Ex9	GCCCCTGTTTGCATGGAGGAAACT TGGAAAGACAGCTACAGTTCTTCTG				244
Amorces utilisées pour le séquençage Sanger de l'ADNc							
<i>Vps15</i>		Ex11-12F	TCCACAGGTCTCAGAAGCAG		Ex14-15R	AGAATCCTGCTTTGTAGAGGGT	555

## 2. Tableau 3 : Liste des anticorps (AC)

Type d'AC	Hôte	Nom protéine ciblée	Dilution	Référence	Fournisseur
<b>PARTIE 1</b>					
PROJET 2 portant sur <i>BBS1</i>					
Application : WB					
Primaire	Souris	Anti-βtubuline	1/2500 TBST 0,1% LAIT 5%	TUB-2A2	Euromedex
	Souris	Anti-BBS1	1/500 (LAIT)	H00000582-A01	Euromedex
Secondaire	Cheval	Anti-souris	1/2000 (LAIT)	7076	Cell signaling
PROJET 3 portant sur <i>IQCE</i>					
Application : immunocytofluorescence					
Primaire	Souris	Alpha tubuline acétylée	1/200 PBS1X SVF 2%	AB2461	Invitrogen
	Lapin	Anti-EVC2	1/250	ab198930;	Abcam
secondaire	Chèvre	Anti-souris (alexa 594)	1/1000	A11032	Invitrogen
	Ane	Anti-lapin (alexa 488)	1/1000	#A21206	Thermo-scientific
Application : WB					
Primaire	Chèvre	Anti-gli2	1/2000 TBST 0,1% LAIT 5%	AF3635	R&D systems
	Souris	Anti-lamine a/c	1/500 (LAIT)	Sc-376248	Santa cruz
	Souris	Anti-βtubuline	1/2500 (LAIT)	TUB-2A2	Euromedex
secondaire	Souris	Anti-chèvre	1/5000 (LAIT)	Sc 2354	Cell signaling
	Cheval	Anti-souris	1/2000 (LAIT)	7076	Cell signaling
	Ane	Anti- lapin	1/2000 (LAIT)	7074	Cell signaling
PROJET 4					
Application : WB, Co-IP (*)					
Primaire	Chèvre	Anti-GLI2	1/2000 TBST 0,1% LAIT 5%	AF3635	R&D systems
	Souris	Anti-βtubulin	1/2500 (LAIT)	TUB-2A2	Euromedex
	Souris	Anti-LAMIN A/C	1/500 (LAIT)	Sc-376248	Santa cruz
	Lapin	Anti-BBS1*	1/500 (LAIT)	21118-1-AP	Proteintech
	Lapin	Anti-BBS5*	1/500 (LAIT)	14569-1-AP	Proteintech
	Lapin	Anti-bbs9	1/500 (LAIT)	14460-1-AP	Proteintech
Secondaire	Souris	Anti-chèvre (HRP linked)	1/5000 (LAIT)	Sc 2354	Cell signaling
	Cheval	Anti-souris (HRP linked)	1/2000 (LAIT)	7076	Cell signaling
	Chèvre	Anti-lapin (HRP linked)	1/2000 (LAIT)	7074	Cell signaling
<b>PARTIE 2</b>					
Application : WB, Co-IP (*)					
Primaire	Souris	Anti-βtubuline	1/2500 TBST 0,1% LAIT 5%	TUB-2A2	Euromedex
	Lapin	Anti-vps15*	1/1000 (LAIT)	NBP1-30463	Novus biological
	Lapin	Anti-Beclin1*	1/1000 TBST 0,1% BSA 5%	#3495	Cell signaling
	Lapin	Anti-Kap0*	1/2000 (LAIT)	20358-1-AP	Proteintech
	Souris	Anti-bin1*	1/1000 (LAIT)	C99D	Merck
	Lapin	Anti-Dapk1*	1/1000 (LAIT)	LS-C413195	Cliniscience
	Lapin	Anti-tau	1/1000 (LAIT)	46687	Cell signaling
	Lapin	Anti-phospho tau S404	1/1000 (BSA)	20194	Cell signaling
	Lapin	Anti-phospho tau S416	1/1000 (BSA)	15013	Cell signaling
	Lapin	Anti-phospho tau S202	1/1000 (BSA)	39357	Cell signaling
	Souris	Anti-phospho tau S396	1/1000 (LAIT)	9632	Cell signaling
	Lapin	Anti-phospho tau S404	1/1000 (LAIT)	35834	Cell signaling

Type d'AC	Hôte	Nom protéine ciblée	Dilution	Référence	Fournisseur
Secondaire (couplé à HRP)	Cheval	Anti-souris	1/2000 TBST 0,1% LAIT 5%	7076	Cell signaling
	Ane	Anti-lapin	1/2000 (LAIT)	7074	Cell signaling
	Souris	Anti-chèvre	1/5000 (LAIT)	Sc 2354	Cell signaling
	Chèvre	Anti-souris	1/1000 (LAIT)	A-11032	Thermofisher
Application : Immunohistochimie-paraffine (IHC-P)					
Primaire	Lapin	Anti-vps15	1/100	HPA036032	Sigma
	Lapin	Anti-vps15	1/100	Ab124817	Abcam
	Lapin	Anti-vps15	1/500	GTX132466	Genetex
	Lapin	Anti-Uromoduline	1/500	AB223540	Abcam
	Souris	Anti-mégaline	1/200	AB184676	Abcam
	Lapin	Anti-Aqp2	1/200	SC-28629	Santa Cruz
Secondaire	Chèvre	Anti-lapin	1/200	BA-1000	Vector Lab
	Cheval	Anti-souris	1/200	BA-2000	Vector Lab
Application : immunocyto/histofluorescence					
Primaire	Souris	Anti-alpha tubuline acétylée	1/200 PBX1X-SVF 2%	AB2461	Invitrogen
	Lapin	Anti-P62	1/200	#7695	Cell signaling
	Lapin	Anti-Lc3A	1/1000	#12741	Cell signaling
	Souris	Anti-synucléine	1/1000 PBS1X-BSA5%	#2647	Cell signaling
	Lapin	Anti-β amyloïde peptide	1/1000	#8243	Cell signaling
Secondaire	Chèvre	Anti-souris (alexa 594)	1/1000	A11032	Invitrogen
	Ane	Anti-lapin (alexa 488)	1/1000	#A21206	Thermo-scientific

### 3. Tableau 4 : Liste des buffers

Nom	Composition (concentration finale mentionnée)
Tampon PK	6,25mL Tris 2M (pH=7,8) + 50mL EDTA 0,5M + 6,25 SDS 20% qsp 250mL H2O
Acétate sodium (3M)	204,12g d'acétate (MM=136,08 g/mol) qsp 500mL H2O
MIX PCR (5X)	Buffer PCR 10X (100 mM tris-HCl, pH 8,3, 500 mM KCl), d'une solution de MgCl <sub>2</sub> (25 mM), des 4 dNTP (200 μM final)
RIPA	RIPA base, Na <sub>3</sub> VO <sub>4</sub> 1M, NaF 1M, cocktail inhibiteur de protéase 25X (Complete EDTA free, 05056489001, Roche)
Tampon de lyse (LB)	Triton 1%, NaCl 50 mM, tris 50 mM (pH8), complete roche
Tampon de lavage (WB)	Triton 0,1%, NaCl 50 mM, tris 50 mM (pH8), complete roche
Tampon au citrate de sodium	10mM Sodium Citrate, 0.05% Tween 20, pH 6.0

### 4. Tableau 5 : Liste des nourritures pour souris

Nom	Composition	Référence
High tout « I2 »	EF Western mod. 25kJ% Fat, 50% Fructose, 5% Glc, 1,63% Xanthan gum, High cholesterol - POWDER	Ssniff Spezialdiäten GmbH D-59494 Soest
« Chow diet », Safe A03	52% glucides, 21,4% protéines, 5,1% lipides	

## 5. Tableau 6 : Liste des logiciels

Nom	Application	Source
Microsoft Excel	Analyse et calcul	Microsoft office
Microsoft Power point	Présentation et figure	
Microsoft Word	Rédaction	
Graphpad Prism6/8	Test statistique	<a href="https://www.graphpad.com/scientific-software/prism/">https://www.graphpad.com/scientific-software/prism/</a>
Image Lab	Analyse et quantification protéique	<a href="https://www.bio-rad.com/fr-fr/product/image-lab-software?ID=KRE6P5E8Z">https://www.bio-rad.com/fr-fr/product/image-lab-software?ID=KRE6P5E8Z</a>
CFX maestro software	Analyse et quantification ARN	<a href="https://www.bio-rad.com/fr-fr/sku/12004110-cfx-maestro-software?ID=12004110">https://www.bio-rad.com/fr-fr/sku/12004110-cfx-maestro-software?ID=12004110</a>
Bioedit	Visualisation des séquences sangers	<a href="https://bioedit.software.informer.com/">https://bioedit.software.informer.com/</a>
Funrich	Analyse d'enrichissement	<a href="http://funrich.org/download">http://funrich.org/download</a>
Primer 3	Design d'amorces	<a href="http://primer3.ut.ee/">http://primer3.ut.ee/</a>
Zen pro	Traitement d'image	<a href="https://www.zeiss.fr/microscopie/produits/microscope-software/zen.html">https://www.zeiss.fr/microscopie/produits/microscope-software/zen.html</a>

## II. Les méthodes

### 1. La culture cellulaire de fibroblastes humains

Les cultures primaires de fibroblastes humains issues de biopsie de peau sont cultivées au laboratoire dans des incubateurs dédiés, à 5% de CO<sub>2</sub>, sous 98% d'humidité, à 37°C. Le milieu de culture DMEM-Glutamax à 1g/L de glucose (21885-025, Gibco) utilisé est supplémenté en Sérum de veau foetal SVF (10%) et en antibiotiques Pénicilline-Streptomycine PS (1%).

#### A) Induction de la ciliogénèse

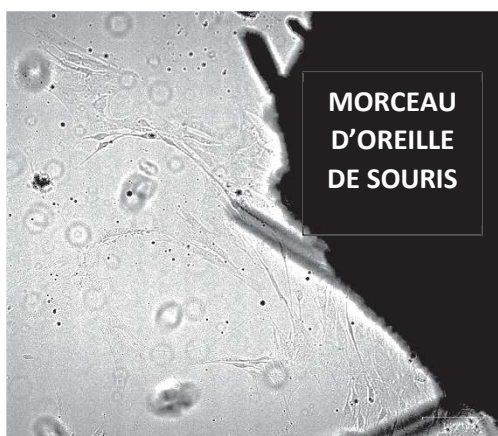
Pour induire la formation du cil (ciliogénèse), les cellules sont cultivées dans un milieu DMEM +1% de PS, déprivé en sérum (-SVF) pendant 24h.

#### B) Stimulation de la voie Sonic Hedgehog (SHH)

Pour permettre l'activation de la voie SHH dans nos cellules, nous les cultivons pendant 24h dans un milieu DMEM +1% PS -SVF + SAG (Smo AGonist) à cf = 100nM. Il s'agit d'une poudre resuspendue dans du DMSO et stockée à ci = 4mM à -20°C.

### 2. Les explants : culture primaire de fibroblastes murins

Les souris sont anesthésiées et des bouts d'oreilles sont prélevés avec un ciseau stérile et immergés dans un milieu F10-nutrient mix (ref 11550043, Thermofisher) supplémenté à 3% PS. Les morceaux sont découpés en petits bouts avec un scalpel stérile, disposés sous une lamelle de verre dans une mini-pétri et immergés dans du milieu DMEM + 10% SVF + 3% PS. Les premiers fibroblastes sont visibles 10-15 jours après la mise sous lamelle (Figure 28). Les fibroblastes sont ensuite cultivés dans un milieu normal DMEM + 10% SVF + 1% PS dans des incubateurs dédiés à 5% de CO<sub>2</sub>, sous 98% d'humidité, à 37°C.



**Figure 28 :** Photo d'un explant. On peut observer que les fibroblastes sont en train d'émerger du morceau d'oreille mis en culture.

#### A) Induction de la ciliogénèse

Pour induire la formation du cil (ciliogénèse), les cellules sont cultivées dans un milieu DMEM +1% de PS, déprivé en sérum (-SVF) pendant 24h.

#### B) Stimulation/blocage du processus d'autophagie

Pour stimuler le processus d'autophagie dans les cellules, elles sont cultivées dans un milieu DMEM +10% SVF +1% PS -Glc, pendant 24h. Nous inhibons les phases tardives du processus en incubant de la chloroquine à 50 $\mu$ M pendant 6h. La solution est diluée à partir d'une solution stock à [100mM], préparée à partir de 520mg de poudre (MM=515,9 g/mol) (Chloroquine diphosphate salt C6628, Sigma Aldrich) resuspendue dans 10mL d'H<sub>2</sub>O.

### **3. Les techniques de biologie moléculaire**

#### A) Design des amorces

Nous générons les séquences des amorces nous-mêmes en utilisant comme base de données « Ensembl-Browser », les logiciels Primers 3, UCSC browser « in silico PCR » et NCBI primers. Pour les couples d'amorces dédiés à la qRT-PCR, nous choisissons des séquences de 25 pb (paire de base) maximum qui sont soit à cheval entre 2 exons consécutifs, soit, séparées par des séquences introniques de plus de 2000 pb (pour ne pas risquer d'amplifier d'ADNg). Les couples sont choisis pour obtenir des fragments d'environ 150 pb, avec un T<sub>M</sub> aux alentours de 60°C, une proportion en GC entre 40 et 60%. Les couples doivent avoir 100% d'identité avec le gène d'intérêt lors du BlastN et présenter un score de complémentarité (entre les 2 séquence d'amorces ou auto-repliement d'une des séquences) inférieur à 5.

Pour les couples d'amorces dédiés à l'amplification d'ADNg, nous choisissons des séquences situées dans des exons ou des introns en fonction de la zone à cibler. Chaque amorce contient au maximum 25 pb, ayant un T<sub>M</sub>~60°C, une proportion en GC entre 40 et 60%. Les couples doivent avoir 100% d'identité avec le gène d'intérêt lors du BlastN et présenter un score de complémentarité (entre elles ou auto-repliement) inférieur à 5.

Les amorces sont systématiquement testées sur l'ADN d'un contrôle (ADNg ou ADNc) avec un protocole de PCR gradient (à des températures d'hybridation 59, 60, 61, 62 °C) pour vérifier l'amplification d'un unique fragment à la taille attendue et déterminer la température d'hybridation optimale à utiliser. Lors d'amplification d'ADNg, le protocole et la stringence peuvent être optimisés.

#### B) PCR et électrophorèse, élution du fragment depuis le gel

Pour la partie 1, l'extraction d'ADNg est réalisée à partir du kit DNeasy Blood & Tissue Kit (ref 69504, Qiagen). La concentration d'ADNg est mesurée par Spectrophométrie (Eppendorf Bio Photometer

Plus). Les PCR de géotypage utilisent le mix PCR 5X. La Taq DNA polymérase est stockée dans une solution composée de Tris-HCl (20 mM), d'EDTA (0,1 mM), de DTT (1 mM), de stabilisants et de glycérol (à 50%). Le mix PCR et la Taq polymérase sont dilués 5 fois dans les tubes. 50 ng d'ADN est ajouté dans chaque tube PCR. Des variations de protocole PCR ont été appliquées pour optimiser l'amplification des fragments d'intérêt. Pour des séquences riches en GC et/ou en répétitions, nous appliquons un protocole de PCR « touch down » pour lequel, les 14 premiers cycles sont réalisés à une température d'hybridation plus basse pour favoriser l'amplification de fragments sous représentés. 35 autres cycles sont ensuite effectués à une température d'hybridation normale. Pour rendre les conditions plus stringentes, il est possible d'ajouter du DMSO en quantité variable (3, 5, 10%) et de bétaïne (1M). La visualisation des fragments se fait au moyen d'une électrophorèse sur gel d'agarose 2%. Les produits PCR sont alourdis avec un bleu de charge et déposés sur gel avec un marqueur de taille. Dans le cas où plusieurs fragments sont générés lors de la PCR (plusieurs bandes visualisées sur le gel d'électrophorèse), il n'est pas possible de les envoyer au séquençage car nous obtiendrions une séquence illisible avec des superpositions de pics. Ainsi, comme ce fut le cas pour le projet 1 portant sur le gène *IFT27*, nous avons élué des fragments migrés dans le gel afin de les réamplifier séparément et d'obtenir un fragment unique. Dans ce cas, nous découpons le gel sur une plaque UV pour obtenir un mini cube de gel contenant le fragment d'intérêt. Nous extrayons le fragment par immersion dans du Tris-EDTA buffer 10.1, toute la nuit à 4°C et nous réamplifions en utilisant 5µL de la solution contenant le produit d'éluion et appliquons exactement le même protocole. A l'issue de la PCR, nous déposons les produits PCR sur le gel et vérifions que le fragment réamplifié est unique et correspond à la taille attendue.

#### C) Séquençage Sanger

Lorsque nous voulons obtenir la séquence exacte des fragments amplifiés, nous purifions le produit PCR pour éliminer les résidus de Taq, de dNTP, d'amorces en le déposant sur une membrane placée sous vide pendant 10 min. Seuls les fragments PCR persistent sur la membrane. Ils sont resuspendus dans l'eau pendant 15min sous une légère agitation. La solution récupérée est ensuite envoyée avec l'amorce à utiliser pour le séquençage (diluée au ¼). Le laboratoire sous-traite l'étape de séquençage par GATC Biotech (Allemagne).

#### D) Extraction d'ARN et synthèse d'ADNc

Pour la partie 1, l'extraction d'ARN s'effectue en suivant les instructions des 2 kits, QUIAshredder (Qiagen, ref 79654) et RNeasy Minikit (Qiagen, ref 74104).

Pour la partie 2, les reins et les rétines ont préalablement été congelés à sec par immersion dans l'azote liquide après la dissection. Pour les extractions d'ARN de ses tissus, la lyse est réalisée par immersion du tissu congelé dans le Trizol (ref. 15596-018, Invitrogen, Life technologies™, Carlsbad, California,



USA). Après broyage au TissueRuptor (9001272, Qiagen, Nederland), le chloroforme permet d'isoler les ARN totaux. La précipitation de l'ARN est réalisée dans l'isopropanol. Le culot est lavé à l'éthanol 75%.

La quantification de l'ARN resuspendu dans l'eau RNase free est toujours réalisée par spectrophotométrie (Eppendorf Bio PhotometerPlus). L'ARN est déposé sur le gel d'agarose pour vérifier son intégrité : les 2 bandes correspondantes aux ARNr 28s et 18S doivent être visibles. L'étape de Reverse Transcription (RT) est réalisée en utilisant le kit « iScript™ cDNA Synthesis Kit » (170-889, Biorad, USA) selon les instructions du fournisseur. Un maximum de 1 µg d'ARN est utilisé.

#### E) qRTPCR

Nous réalisons une qPCR différentielle avec le C1000™ Thermocycleur (CFX96, Real-time System, Bio-Rad, USA). La liste de toutes les amorces utilisées est disponible dans le Tableau 2. Dans le cas de la partie 1, j'utilise systématiquement 3 ADNc différents dits « contrôles » (triplicats biologiques). Je réalise des triplicats techniques avec l'ADNc des patients sauf s'il est possible d'avoir de l'ADNc issus d'extractions distinctes (duplicats/triplicats biologiques). Pour la partie 2, j'ai toujours utilisé 3 ADNc différents de souris WT et de souris KI (triplicats biologiques pour les 2 génotypes).

L'ADNc est dilué au 1/5<sup>ème</sup> ou 1/10<sup>ème</sup> dans chaque puit. L'iQTM SYBR Green Supermix Biorad (170-8886, Biorad) est mélangé avec les amorces (ci=100 µM), diluées de manière à avoir 10 µmol/puit. Les références d'expression sont les gènes *GAPDH* (Hu)/*Gapdh* (Mu) et *HPRT* (Hu)/*Hprt* (Mu). La qPCR est réalisée en suivant le protocole suivant : première dénaturation à 95°C (5 min), suivie de 39 cycles à 95°C (10s), 60°C (30s) suivie d'une augmentation de 0,5°C toutes les 5s entre 65°C et 95°C pour générer les « melting curves ». Le niveau d'expression des gènes a été calculé avec la méthode de quantification relative « threshold  $\Delta\Delta Ct$  » grâce au logiciel CFX Manager Software V.1.5 (Biorad) et grâce au traitement Excel, en normalisant par rapport aux gènes de référence. Les niveaux d'expression des contrôles et des souris WT sont rapportés à 1. Une analyse statistique a été menée sur les données obtenues afin de montrer leur niveau de significativité en utilisant le logiciel « GraphPad prism ». Un test de Fisher est systématiquement mis en œuvre suivi d'un test non paramétrique de type Mann-Whitney ou un test paramétrique comme le Test de Student en fonction de l'effectif analysé.

### 4. Les techniques de biochimie

#### A) Extractions protéiques

Pour l'ensemble des projets, l'extraction des protéines totales utilise le RIPA. Lorsqu'il s'agit de tissus de souris (partie 2), nous utilisons le TissueRuptor (9001272, Qiagen, Nederland) pour la lyse. Lorsqu'il s'agit de culots cellulaires (partie 1), nous lysons par aspiration-reflux à la pipette. Dans tous les cas,

une centrifugation à 15000 rpm (rotation par minute) pendant 20 min à 4°C permet la séparation des protéines (surnageant) et des membranes (culot). Le surnageant est récupéré et soniqué pendant 30s avec un Sonic ruptor 250 (Omni international) en mode pulsé. Une seconde centrifugation de 10 min à 4°C à 15000 rpm est réalisée. Le surnageant est récupéré et la quantification des protéines totales est effectuée par un test de Bradford.

Dans le cadre des projets 3 et 4 portant respectivement sur *IQCE* et *BBS5*, nous avons effectué une extraction protéique différentielle sur des culots cellulaires fraîchement récupérés à partir de cellules cultivées dans des conditions -SVF + SAG. Le protocole décrit dans Estrada-Cuzcano et al. 2020 a été appliqué. La lyse est réalisée par aspiration-reflux dans 200 µL PBS1X-NP40 (Nonidet P40, 19628, USB corporation). 100 µL est récupéré et constitue la fraction « W » correspondant aux protéines totales. Une première centrifugation rapide du volume restant permet de générer un culot et le surnageant de 100 µL qui correspond à la fraction « C » (cytoplasmique). Enfin, après 3 lavages avec le buffer d'extraction suivis de 3 centrifugations rapides, le culot est resuspendu dans 90 µL de buffer d'extraction : nous obtenons la fraction « N » (nucléaire). Une sonication pendant 10s est réalisée sur les fractions W et N pour dégrader les membranes résiduelles.

#### B) Quantification protéique (test de Bradford)

200 µL du réactif de Bradford (Bio-rad protein assay, cat.#500-0006, Bio-rad laboratories GmbH, München) est mélangé à 800 µL d'eau distillée. 5 µL du tampon de lyse (pour le blanc) ou de la solution d'intérêt est utilisé dans le mélange. En parallèle, une gamme est réalisée à partir de BSA (Bovine serum albumin) à des concentrations connues pour générer une droite de référence à laquelle se référer pour déterminer la concentration protéique du surnageant à tester. Un changement de couleur du jaune au bleu prouve la présence de protéines dans le mélange. Plus la couleur bleue est soutenue, plus la concentration en protéines est importante. Nous faisons la lecture de l'absorbance à 595 nm avec une spectrophotomètre (Eppendorf BioPhotometer Plus). La concentration déterminée est divisée par 5 pour obtenir la concentration protéique initiale.

#### C) Poudre d'organes et test ELISA (partie 2 uniquement)

Nous avons réduit les organes des souris en poudre pour pouvoir réaliser plusieurs expériences différentes à partir du même organe. Ils ont été congelés à sec et maintenus au -80°C. Ils sont déposés, dès leur sortie, dans un mortier lui-même immergé dans l'azote liquide. Le pilon, également refroidi, est utilisé pour écraser l'organe et générer la poudre qui est ensuite pesée. Une quantité équivalente pour chaque souris est resuspendue dans du PBS1X (avec une proportion de 500µL pour 30 mg de poudre). Une centrifugation permet de séparer le culot et le surnageant. Ce dernier est récupéré et une quantification protéique est réalisée par un test de Bradford. L'extraction au PBS1X est demandée

dans les protocoles des tests ELISA car d'autres solutions (comme NAF, Na<sub>3</sub>VO<sub>4</sub>, inhibiteurs anti-protéases...) pourraient interférer avec les réactifs du test et potentiellement fausser les résultats. Plusieurs kits ont été appliqués en utilisant les protéines extraites de poudres de cerveaux selon les instructions des fournisseurs. Ils ont été référencés dans le tableau suivant.

**Tableau 7 : Liste des tests Elisa**

Protéine d'intérêt	Référence et fournisseur	Dilution appliquée
PP2A	Abcam Ab209878	1/600
Bace 1	Cloud corp SEA718Mu	1/100
Bace 2	My bio source MBS2882850	1/100
γsecrétase	Cusabio CSBE12142m	1/100
Psen1	Cloud corp SEC200Mu	1/100

#### D) Co-immunoprécipitation (Co-IP)

Dans le projet 4 portant sur le gène *BBS5*, nous avons effectué une co-immunoprécipitation contre la protéine BBS5 à partir de culots cellulaires, suivie d'un western blot. Les protéines ont été extraites à partir des culots de fibroblastes de la patiente et de contrôles, cultivés en condition –SVF. L'extraction a été réalisée en utilisant le tampon de lyse (LB). Les culots sont resuspendus dans le LB (150µL/culot), puis broyés au mortier.

Dans la partie 2, nous avons également effectué une Co-IP contre la protéine Vps15 à partir d'une tranche d'hippocampe contrôle, suivie d'un WB. Le tissu est resuspendu dans le tampon de lyse LB (4mL) puis broyé avec TissueRuptor (9001272, Qiagen, Nederland) pendant 3 x 10s.

Ensuite, c'est le même protocole qui est suivi : les volumes broyés sont récupérés et centrifugés 20 min, à 15000 rpm, à 4°C. Après avoir récupérer tous les surnageants, les solutions obtenues ont été quantifiées par le test de Bradford. Les tubes Co-IP + et – ont été préparés pour les deux conditions (patiente et contrôle). Les tubes Co-IP + contiennent 300 µg de protéines et 0,6 µg d'AC qsp (quantité suffisante pour) 500µL avec le LB. Les tubes Co-IP - ne contiennent pas d'AC. Les AC utilisés sont référencés dans Tableau 3.

Des « inputs » de 30 µL à [prot] = 1µg/µL sont aliquotés et stockés à -20°C, ils correspondent aux protéines totales extraites. Les tubes Co-IP + et - sont gardés à 4°C, sous agitation douce, toute la nuit, pour permettre la reconnaissance entre l'AC et la protéine appât. Le lendemain, 50 µL de billes de sépharose sont centrifugés rapidement. Le surnageant est jeté et le culot est resuspendu avec du LB. Ce cycle est répété 4 fois afin de remplacer l'éthanol (solvant initial des billes) par le LB (rééquilibrage). Par la suite, les 500 µL des tubes Co-IP + et – sont ajoutés aux culots de billes et conservés à 4°C, sous agitation douce, pendant 1h30 pour favoriser l'interaction entre les billes et le complexe AC-protéines. Des lavages avec le tampon dédié (WB) sont réalisés pour éliminer les

protéines n'ayant pas interagi avec les billes. Finalement, seul le culot de bille reste à l'issue des lavages. Enfin, 50 µL de Laemmli 2X (cat.#161-0767, Biorad) sont ajoutés aux culots (protéines + AC + billes), un chauffage de 7 min à 95°C à 1200 rpm permet de séparer les billes, les AC et les protéines. Le surnageant est récupéré après centrifugation et correspond aux complexes protéiques seuls.

#### E) Western Blot

30 à 40 µg de protéines sont mélangés avec le Laemmli 4X (dilué au ¼, cat.#161-0747, Bio-rad laboratories), chauffés à 95°C pendant 15 min, à 450 rpm. Le mélange est déposé dans les puits de gel polyacrylamide (Mini-PROTEAN<sup>R</sup> TGX<sup>TM</sup> Gels cat.#456-1034, Biorad). Dans le cas des extractions protéiques différentielles, le dépôt est réalisé sur un gel Criterion permettant un marquage stain free (Criterion<sup>TM</sup> TGX Stain free<sup>TM</sup> Precast gels, cat.#567-8024, Biorad). La migration est lancée dans un tampon TGS1X (préparé à partir de 10X, 25mM Tris, 192 mM Glycine, 0.1% (w/v) SDS, 161-0772, Biorad). Un marqueur de taille est systématiquement ajouté en parallèle (Precision plus protein<sup>TM</sup> all blue, cat.#161-0373, Biorad). Le transfert sur membrane (nitrocellulose ou PVDF) est réalisé au moyen du Transblot<sup>R</sup> Turbo<sup>TM</sup> Transfer System (Biorad).

Les inputs, les Co-IP + et - sont déposés sur un gel de polyacrylamide SDS-PAGE 10% Mini PROTEAN TGX, Bio-RAD (456-1034). L'électrotransfert du gel est effectué sur une membrane de PVDF (0,2µM, Trans-Blot Turbo, Transfer Pack, Bio-RaD (1704156).

Dans le cas des gels Criterion, la photo du « stain free » est faite sur la membrane après transfert. Dans les autres cas, une coloration au rouge ponceau est systématiquement faite sur la membrane pour vérifier la présence de protéines. Il faut saturer les sites aspécifiques en bloquant la membrane avec la solution de blocage à 5% lait dilué dans du TBS1X-T 0.1% pendant la nuit, à 4°C. L'incubation de l'AC primaire dilué dans la solution de blocage s'effectue toute la nuit à 4°C. L'AC secondaire est dilué dans la même solution et incubé 1h, à température ambiante. La révélation est réalisée avec un Chemidoc<sup>TM</sup> MP Imaging System (Biorad) à l'aide des solutions Pierce<sup>TM</sup> ECL Western (ref 32209, ThermoScientific) et SuperSignal R West Femto Maximum Sensitivity Substrate (prod #34095, ThermoScientific). La normalisation est effectuée en utilisant le logiciel Image Lab software 6.0.1 (Biorad), par rapport au stain free des membranes ou en prenant comme protéines de référence la  $\beta$ TUBULINE/ $\beta$ tubuline ou la GAPDH/Gapdh.

#### F) Immunocytofluorescence et Immunohistofluorescence

Dans le projet 3 portant sur le gène *IQCE* et la partie 2 portant sur les cellules des souris, nous avons effectué des Immunocytofluorescences à partir de cellules adhérentes sur labteks à 8 puits (Lab-tek chamber slide w/cover Permanox slide sterile 177445, ThermoScientific). La fixation est effectuée au paraformaldéhyde (PFA) 4% pendant 30 min à température ambiante ou au méthanol 100% pendant 20 min à -20°C. La perméabilisation est réalisée par incubation pendant 10 min de triton 0,01%. Du

PBS1X-SVF 20% est utilisé pour le blocage (1h). Les AC primaires ont été dilués dans du PBS1X-SVF 2% et incubés toute la nuit, à 4°C. Les AC secondaires également dilués dans la même solution, ont été incubés 1h, à l'abri de la lumière. Le marquage nucléaire est réalisé au DAPI (Hoechst, 33258, Invitrogen H3569), dilué au 1/2000<sup>ème</sup>, incubé 10 min, à température ambiante, à l'abri de la lumière. Les lames sont finalement montées avec le milieu Vectashield Mounting Medium. L'observation a été réalisée à l'aide du microscope Zeiss imager Z2. Les coupes ont été analysées grâce au logiciel ZEN 2012 (Carl Zeiss Inc., Oberkochen, Allemagne).

Dans le cadre de la partie 2, les cerveaux des mâles de 1 an ont été prélevés et découpés en 2 moitiés sagittales. Une moitié a été directement immergée dans l'OCT (Optimal Cutting Temperature, ref. 4583, Tissue-Tek, OCT<sup>TM</sup>, Satura Finetek, Torrance, Californie, USA) : un gel transparent qui se solidifie et blanchi dans l'azote liquide. Les cryocoupes de 10 µm d'épaisseur ont été obtenues avec le Cryostat Leica CM1950 (ref. 1404778001, Leica Biosystems, Wetzlar, Allemagne).

Pour l'immunohistofluorescence, les cryocoupes des cerveaux ont été fixées 3 min à la formaline 4% et perméabilisées pendant 5 min avec du PBS1X-Triton 0,01% (ET330, Euromedex, France). Du PBS1X-BSA 5% (A7030-100G, Sigma Aldrich, USA) est utilisé pour le blocage (30 min). Les AC primaires ont été dilués dans la solution de blocage et incubés toute la nuit à 4°C. Les AC secondaires également dilués dans la solution de blocage, ont été incubés 1h, à l'abri de la lumière. Le marquage nucléaire est réalisé au DAPI, incubé 5 min, à température ambiante, à l'abri de la lumière. Les lames ont été montées avec le milieu Vectashield Mounting Medium. L'observation de l'ensemble des coupes a été réalisée à l'aide du microscope Zeiss imager Z2. Les coupes ont été analysées grâce au logiciel ZEN 2012 (Carl Zeiss Inc., Oberkochen, Allemagne). Les tissus de souris WT du même âge ont subi le même traitement et servent de contrôles. Les analyses ont été réalisées en mode « z-stack » afin de déterminer la présence et la taille des agrégats protéiques à la surface des tissus.

## **5. Les techniques dédiées aux souris**

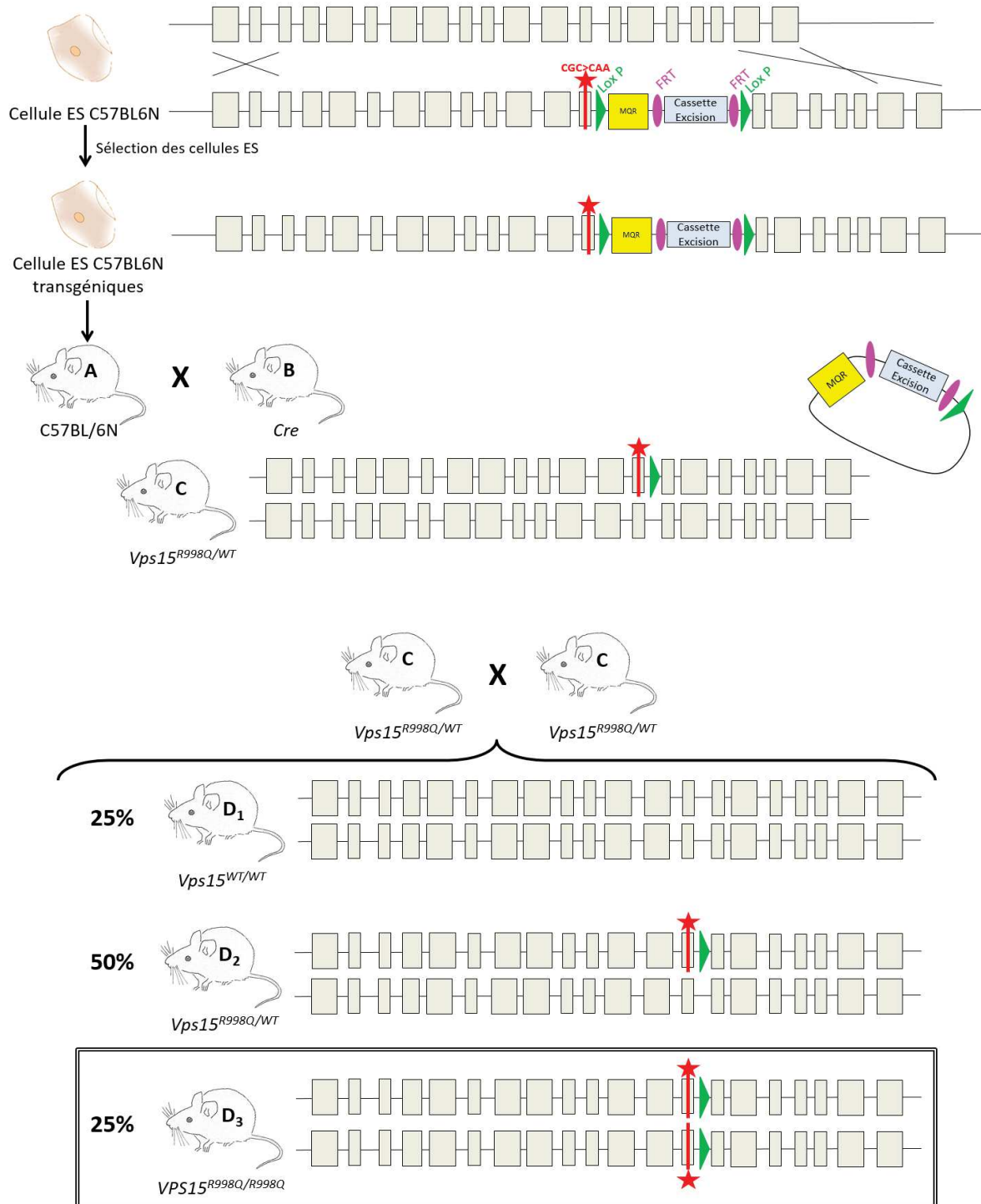
### **A) La génération de la lignée par l'ICS**

Les souris *Vps15* porteuse de la mutation faux-sens R998Q ont été générées à l'ICS, sur un fond génétique C57BL6N, grâce au financement de la Fondation maladie rare (FMR) par application d'une technique « point de mutation » utilisant le système Cre-Lox, basée sur la publication O'Gorman et al. 1997.

Une souris transgénique KI technique « point de mutation » est un modèle pour lequel des nucléotides ont été substitués de manière à obtenir la mutation d'intérêt. Initialement, 2 types de souris ont été commandés : un type avec le fond génétique C57BL/6J et un autre avec le fond C57BL/6N. Le seul type de souris généré porte un fond génétique C57BL/6N. Chez l'humain, la mutation est une substitution

c.2993G>A qui conduit à un faux-sens p.R998Q. Le gène *Vps15* murin, situé sur le chromosome 9, mesure 44,50 kb, compte 20 exons et code pour une protéine de 1358 AA. Chez la souris, pour obtenir cette même mutation, il a fallu substituer 2 nucléotides : CGC est devenu CAA aux positions 3429-3431. Comme le résume la Figure 29, La substitution des nucléotides est d'abord réalisée dans des cellules ES par recombinaison homologe entre le fragment génomique endogène et le fragment génomique exogène, porteur de la mutation R998Q. Après sélection des cellules, celles-ci sont réinjectées dans la masse interne d'une souris mère-porteuse. Les souris **A** « porteuses hétérozygotes pour l'allèle muté non excisé » sont croisées avec des souris **B** exprimant la Cre recombinase de manière ubiquitaire : il y a excision de la séquence encadrée par les sites LoxP. A l'issue du croisement, les souris **C** (*Vps15*<sup>R998Q/WT</sup>) sont obtenues. Ces dernières sont ensuite croisées ensemble, de manière à générer des souris **D**<sub>1</sub> (*Vps15*<sup>WT/WT</sup>), **D**<sub>2</sub> (*Vps15*<sup>WT/R998Q</sup>) et **D**<sub>3</sub> (*Vps15*<sup>R998Q/R998Q</sup>).

La publication Mattapallil et al. 2012 a rapporté la présence de la mutation Rd8 (c.3481delC) au niveau du gène *Crb1* seulement chez les souris de fond génétique C57BL6N. Cette mutation cause un phénotype rétinien. Au départ, toutes les souris *Vps15* étaient porteuses de la mutation Rd8 à l'état homozygote. Pour ne pas biaiser notre étude car nous nous attendons à un potentiel phénotype rétinien, nous avons décidé de faire des backcross (8 jusqu'à présent) de nos souris hétérozygotes avec des congénères WT de fond génétique C56BL6J. L'absence de la mutation Rd8 a été vérifiée par simple génotypage sur l'ensemble des souris.



**Figure 29 :** Génération des souris KI *Vps15*. Représentation schématique du processus mis en place pour la génération des souris transgéniques *Vps15*<sup>R998Q</sup> à partir de cellules ES. Mutation c.3430\_3431delinsAA ; p.R998Q, sites LoxP, sites FRT, MQR : gène marqueur.

#### B) Maintien de la lignée

Les souris sont stabilisées à l'animalerie centrale de la Faculté de médecine de Strasbourg. Les cohortes sont générées par croisement *in situ* entre souris hétérozygotes *Vps15*<sup>R998Q/WT</sup>. Les cohortes sont constituées de souris KI *Vps15*<sup>R998Q/R998Q</sup> et de contrôles WT *Vps15*<sup>WT/WT</sup> issues des mêmes portées



(« littermate ») dans la mesure du possible. Les animaux sont sevrés entre 3 et 4 semaines et considérés comme matures sexuellement à 2 mois. Leur environnement est maintenu à ~22°C et un cycle jour 12h - nuit 12h est mis en place dans la structure.

Certaines cohortes de souris ont été mises sous une nourriture spéciale que j'ai appelée « I2 » dans la suite du manuscrit et dont la composition est rapportée dans le Tableau 5. Elle se présente sous la forme d'une poudre à reconstituer en pâte avec l'ajout d'eau. Je la prépare et je fais le change de nourriture une fois par semaine. Je regarde la quantité de nourriture consommée par semaine selon les génotypes. Je pèse la quantité de nourriture à jeter et j'en déduis la quantité consommée par les souris. Je pèse la nourriture ajoutée, correspondant à celle qui sera consommée. Enfin, je pèse les souris. Finalement, je calcule la différence entre la nourriture ajoutée la semaine précédente et la nourriture consommée et je la rapporte au poids de chaque souris. Je divise enfin ce chiffre par le nombre de jours qui s'est écoulé entre 2 changes. Ces contrôles me permettent de confirmer que les souris mangent et d'avoir un différentiel entre les génotypes. Le reste des souris est nourri avec une nourriture normale de type « chow diet », changée une fois par semaine par les soigneurs, en même temps que le change des cages. Ces souris sont également pesées chaque semaine pour générer la courbe de poids en fonction du génotype.

Les manipulations des souris ont été réalisées dans le respect des lois d'expérimentation animale (article R214-87 à R214-137 du code rural, mis à jour en 2013) sous couvert d'une autorisation de projet référencée sous le numéro APAFIS#16840-2018092413511703 obtenue en Mars 2019.

### C) Génotypage de l'ADNg et de l'ADNc

L'extraction d'ADNg a été réalisée à partir de biopsies de doigts de souris. L'étape de digestion à la protéinase K est réalisée, sous agitation, à 55°C, toute la nuit. L'ajout du mélange acétate de sodium-éthanol 100% permet de faire précipiter l'ADN. Le culot, une fois lavé des sels résiduels grâce à l'éthanol 70% est resuspendu dans du TE 10.1, durant 1h à 55°C. La concentration d'ADN de la solution est mesurée par spectrophométrie (Eppendorf Bio PhotometerPlus). Les PCR de génotypage utilisent un mix PCR 5X. La Taq DNA polymérase est stockée dans une solution composée de Tris-HCl (20 mM), d'EDTA (0,1 mM), de DTT (1 mM), de stabilisants et de glycérol (à 50%). Le mix PCR et la Taq polymérase sont dilués 5 fois dans les tubes. 50 ng d'ADN sont ajoutés dans chaque tube PCR.

Pour le génotypage de *Vps15*, les amorces ER<sup>2</sup> et EF<sup>2</sup> (100 mM) sont diluées pour atteindre 6 µmol/tube. Chaque tube contient 5% de DMSO et ~50 ng d'ADN. Pour le génotypage de la mutation Rd8 de *Crb1*, les 3 amorces mF1, mF2, mR sont utilisées. Le protocole de Mattapallil & al. 2012 a été adapté : 2 PCR distinctes sont réalisées avec d'un côté l'association mF1 (0,5 µM)-mR (1 µM) et de l'autre côté, l'association mF2-mR (1 µM).

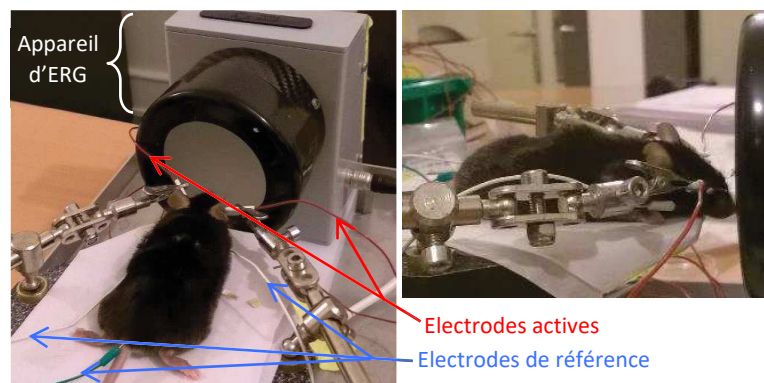
Pour déterminer la séquence exacte du fragment *Vps15* et visualiser la mutation c.3430\_3431delinsAA,



le couple Mf et Mr est utilisé. Les fragments d'ADNg amplifiés sont séquencés avec l'amorce Mf. Pour le séquençage de la mutation R998Q à partir de l'ADNc (obtenu par rétrotranscription de l'ARN extrait) de cerveaux, de reins et de rétines de souris, nous avons utilisé un couple d'amorces dédié. L'ensemble des séquences des amorces utilisées sont regroupées dans le Tableau 2. Les fragments d'ADN amplifiés sont séquencés avec l'amorce sens. Le laboratoire sous-traite l'ensemble des séquençages Sanger par GATC Biotech (Allemagne).

D) Les analyses fonctionnelles

➤ Electrorétinogramme (ERG)



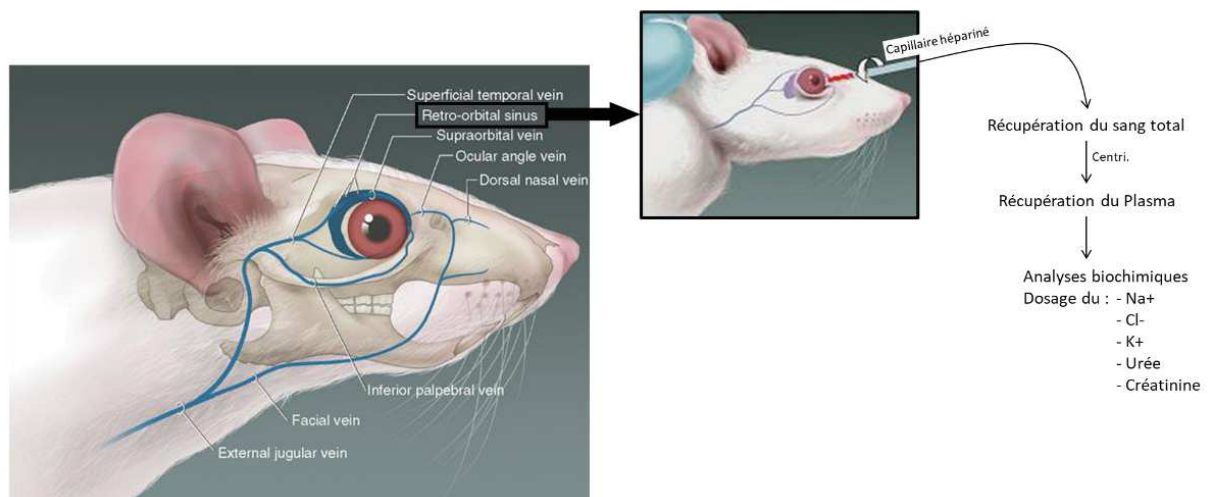
**Figure 30 : Electrorétinogramme.** Photo d'une souris soumise à un ERG.

Des ERG ont été réalisés sur la même cohorte de souris entre 1 et 9 mois avec le système HMsERG (Ocuscience, Kansas City, Missouri, USA) (Figure 30). Les animaux sont préalablement adaptés au noir pendant toute la nuit (condition scotopique). Les expériences se déroulent le lendemain dans le noir, sous lumière rouge : une couleur qui n'est pas perçue par les souris. Leurs pupilles ont été dilatées avec des gouttes d'Atropine (0,3%, Alcon, France) 20 min avant l'anesthésie. Les souris sont anesthésiées par injection intrapéritonéale d'un mélange Domitor (0,76  $\mu\text{g/g}$  de masse corporelle) (ref. 134739-3, Orion Pharma) et Kétamine (76  $\mu\text{g/g}$  de masse corporelle de souris) (ref. 81918704, Virbac France). Les électrodes actives (ERGACC8, Ocuscience) sont positionnées au contact de la cornée et recouvertes de lentilles (ERGACC3, Ocuscience) enduites de Goniovisc (ref. 17238-610-15, Hub pharmaceuticals). Parmi les 3 électrodes de référence, la première est piquée à la base de la queue et les 2 autres sont insérées en surface des joues. L'enregistrement ERG est effectué sur les 2 yeux, en simultané (Figure 30). Les valeurs d'impédance de toutes les électrodes ont toujours été inférieures à 5k $\Omega$ , avant et après l'acquisition. Le protocole permet d'évaluer la réponse visuelle scotopique (réponse des bâtonnets et des cônes) et se compose d'une série de flashes lumineux d'intensités croissantes de 0,1 à 25  $\text{cd.s.m}^{-2}$ . A la fin de l'expérience, les souris ont été réveillées par l'injection unique intra-musculaire d'Antisedan (0,25  $\mu\text{g/g}$  de masse de corporelle) (ref. 134785-1, Orion Pharma).

Les données ont été analysées grâce au logiciel ERG View system 4.3 (Xenotec, Ocuscience, Kansas City, Missouri, USA).

➤ Prélèvements sanguins au niveau du sinus rétro-orbital

Cette procédure a été appliquée à des cohortes de souris (les 2 sexes) d'âges croissants (2, 4, 6, 8 mois), préalablement anesthésiées par voie respiratoire à l'isoflurane dans une chambre d'induction (2L d'O<sub>2</sub>, 4% d'isoflurane). La tête de l'animal endormi est maintenue entre le pouce et l'index. Une compression de la veine jugulaire est mise en place pour générer un meilleur retour veineux. Une brèche dans le sinus orbital est réalisée grâce à un capillaire hépariné (ref 9100260, Hirshmann laborgerate) inséré au niveau du coin de l'œil exorbité de la souris (Figure 31) afin de recueillir le sang. Nous récupérons le plasma à partir du sang centrifugé afin de réaliser des analyses biochimiques.



**Figure 31 :** Illustration du prélèvement rétro-orbital. Annotations de la localisation du sinus et du positionnement du capillaire (inspirée de Yardeni et al. 2011).

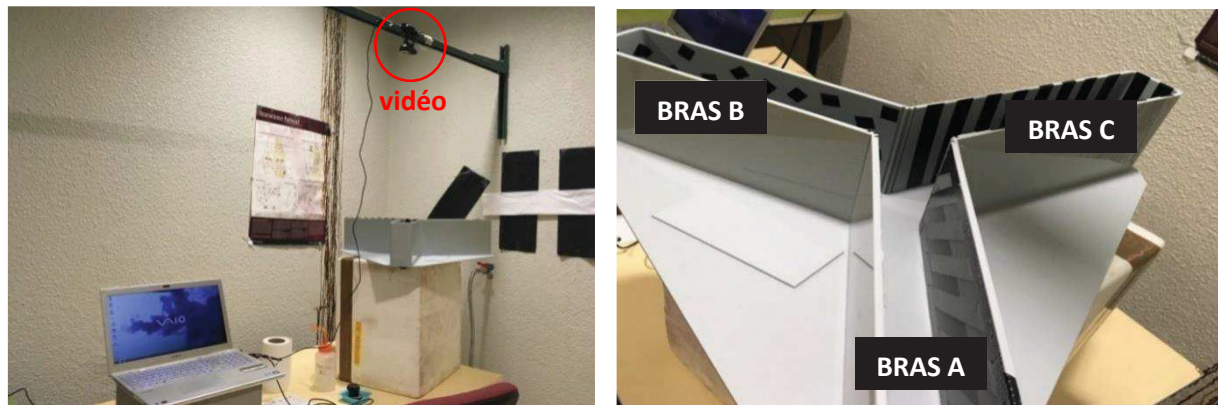
➤ Y-maze et frottis vaginaux

Les souris sont stabulées dans la pièce d'expérimentation 30 min avant le test sans qu'elles ne puissent voir le dispositif (phase d'habituation). Nous utilisons un Y-maze Pan lab (76-0079, Harvard Apparatus). Chacun des 3 bras a été distingué par la présence de motifs différents sur les parois (Figure 32). La luminosité est fixée à 150 lux dans chaque bras. Au début du test, la souris est toujours déposée dans le bras A et l'enregistrement dure 8 min. Le bras est considéré comme « visité » lorsque les 4 pattes de la souris ont passé l'entrée.

Le pourcentage d'alternance est obtenu en appliquant la formule suivante :

$$\text{pourcentage d'alternance} = \frac{\text{nombre de triplets}}{\text{nombre total de bras visités} - 2}$$

Pour les femelles, la phase du cycle menstruel est vérifiée le jour du test par lavage vaginal. Nous incorporons l'extrémité d'une pipette plastique contenant 20  $\mu$ L de PBS1X dans le vagin de la souris. Par aspiration-reflux, nous récupérons le liquide qui est déposé sur une lame de verre recouverte d'une lamelle pour permettre une observation sous microscope du prélèvement. C'est la présence et la quantité de certaines cellules qui nous permettent de différencier les 4 stades du cycle (Calgioni 2009).



**Figure 32 : Le dispositif du Y-maze** : Photos de l'environnement autour du Y-maze et des 3 bras, distingués les uns des autres par des motifs différents.

➤ Prélèvements et dissections

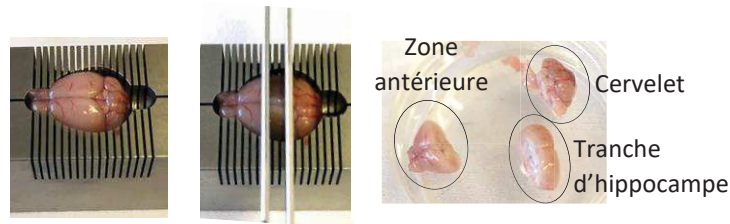
(a) Microdissection des yeux

Après l'euthanasie, les souris sont énucléées et les yeux sont immergés dans un milieu Ames (15230-097, Sigma) suppléé en glucose (0,325 g/L) dans lequel les yeux seront disséqués sous loupe. La cornée est découpée puis le cristallin est extrait. Ensuite, l'épithélium pigmentaire et la rétine sont séparés. Les rétines sont congelées à sec dans l'azote liquide.

(b) Découpe des cerveaux

Les cerveaux de femelles de 8M ont été disséqués et découpés transversalement à 2 niveaux pour générer 3 morceaux en recourant à un moule « brain slicer » (BSMAS001-1, Zivic instruments). Cet outil a permis de reproduire à l'identique la découpe pour chaque prélèvement réalisé à des dates différentes. Nous avons ainsi généré des tranches de 3mm englobant l'ensemble de l'hippocampe pour toutes les souris. Pour déterminer la zone de découpe, j'ai utilisé les informations issues de l'Atlas Brain Allen (<https://portal.brain-map.org/>) qui fournit les coupes sériées coronales des cerveaux de souris adultes espacées de 100 $\mu$ m chacune. L'hippocampe débute à la coupe n°70 soit à 7mm du bulbe olfactif et disparaît à la coupe n°91 soit 2,7mm d'épaisseur après. J'ai utilisé un brain slicer permettant des coupes coronales avec des rainures espacées de 1mm. J'ai positionné mes rasoirs aux rainures n°7 et 10 pour obtenir une tranche de 3mm de cerveaux contenant l'hippocampe (Figure 33). Le recours à ce type de moule m'a permis d'obtenir de manière reproductible des coupes de cerveaux de plusieurs

souris différentes. Ces tranches ont été utilisées pour réaliser les expériences de Co-IP, ce qui explique la nécessité d'avoir des échantillons les plus similaires possibles entre les souris.



**Figure 33 : Découpe d'un cerveau de souris.** Photos du moule « brain slicer » et des 3 morceaux obtenus.

(c) Perfusion au PFA 4%

Le thorax de la souris endormie, par injection intra-péritonéale d'un mélange d'anesthésiques Kétamine/Xylazine à dose létale (respectivement 100 et 10 mg/kg), est ouvert pour permettre l'accès au cœur. L'aiguille reliée à la pompe péristaltique réglée sur un débit de 0,80 mL/min est insérée dans le ventricule gauche. L'oreillette droite est dilacérée afin d'ouvrir le système circulatoire. Un lavage au NaCl 0,9% est d'abord réalisé. La décoloration hépatique est le premier indicateur confirmant que la perfusion est correcte. Ensuite, la perfusion au PFA 4% se poursuit pendant 10 min. L'ondulation de queue dans les minutes qui suivent est le 2<sup>ème</sup> indicateur d'une perfusion efficace atteignant les extrémités. Finalement, les organes sont prélevés et immergés dans les différents fixateurs selon les expériences envisagées. Pour la réalisation de l'étude histologique avec inclusion en paraffine, les organes sont immergés dans le PFA 4%. Pour les analyses par microscopie électronique à transmission (MET), les cerveaux et le pancréas des souris sont immergés dans le fixateur de MET.

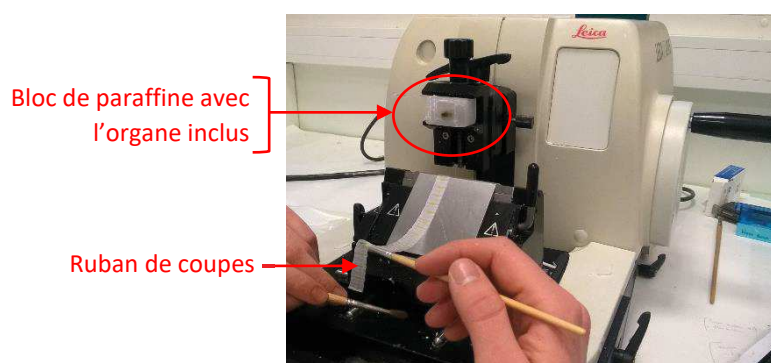
### III. Les méthodes réalisées en collaboration

#### 1. Analyse cérébrale : Yalcin Binnaz (IGBMC, Illkirch)

En 2018, une étude portant sur la morphométrie cérébrale de souris de 5M (*Vps15<sup>R998Q</sup>*, *Crb1<sup>Rd8/Rd8</sup>*) a été menée dans le cadre d'une collaboration avec l'équipe de Yalcin Binnaz (IGBMC, Strasbourg). Le protocole suivi et standardisé est rapporté dans la publication S. C. Collins et al. 2018.

#### 2. Etude histologique : Cristina Antal (Faculté de médecine, Strasbourg)

Nous avons débuté une étude histologique sur des souris de 5M (n=6 par sexes) dans le cadre d'une collaboration avec Cristina Antal (Institut d'histologie, Faculté de médecine, Strasbourg). Après perfusion de l'animal et immersion des organes dans le PFA 4%, les organes subissent une étape de déshydratation par immersion dans des bains croissants d'éthanol et immersion finale dans le xylène (solvant de la paraffine). Les organes ainsi préparés sont inclus dans des blocs de paraffine (liquide à 60°C, solide à température ambiante) conservés à 4°C. Ils sont ensuite découpés avec un microtome à une épaisseur de 5µm (Figure 34).



**Figure 34 : Découpe au microtome.** Photo d'un microtome permettant la découpe des blocs de paraffine dans lesquels des tissus sont inclus.

Les coupes de paraffine sont montées sur des lames Super Frost Plus (Thermoscientific) grâce à un bain-marie maintenu à 46-48°C. Avant la mise en œuvre des différents marquages par IHC-P ou des colorations classiques, les lames subissent une étape de déparaffinage-réhydratation : elles sont incubées dans un premier bain de xylène suivi de bains décroissants d'éthanol.

Pour la coloration hématoxyline-éosine (H&E), les lames déparaffinées sont immergées 1min30 dans le bain d'hématoxyline de Groat (1g sulfate d'ammonium ferrique ; 0,8mL d'acide sulfurique ; 1% solution alcoolique d'hématoxyline qsp 50mL d'eau distillée), différenciées à l'eau distillée pendant 2min, immergées 1min dans le bain d'érythrosine (à 1% dans l'eau distillée) puis rincées à l'eau distillée. Elles sont ensuite déshydratées et montées à l'Eukitt (Labonord).

Pour la coloration au trichrome de Gomori, les lames sont immergées 1min30 dans le bain d'hématoxyline de Groat, différenciées à l'eau distillée pendant 2min puis immergées pendant 5 min dans un bain de colorants acides (composé de 0,6g de 2R chromotrope ; de 0,3g de colorant vert ; de

0,7g d'acide phosphotungstique ; de 1mL d'acide acétique et 100mL d'eau distillée). Elles sont ensuite rincées abondamment à l'eau distillée, à l'acide acétique et encore à l'eau distillée. Elles sont déshydratées et montées à l'eukitt.

Concernant l'immunohistochimie (IHC-P), les coupes de tissus déparaffinés subissent une étape de démasquage de l'antigène par chauffage au microonde 800W 2x5min avec du buffer citrate 0.1M (pH6). Les signaux aspécifiques sont bloqués par incubation dans la solution de blocage (5% sérum normal de chèvre dilué dans le PBS1X-Tween 0,1%). Une étape d'inactivation des peroxydases endogènes est mise en place en incubant les coupes avec de l'eau oxygénée (3% H<sub>2</sub>O<sub>2</sub>, 10min). Les coupes ont ensuite été incubées toute la nuit avec l'AC primaire. Le lendemain, les coupes sont incubées avec les AC secondaires biotinylés. Cette étape est suivie par une incubation de 2h avec le complexe Avidin-Biotin du kit VECTASTAIN® Elite ABC-HRP Kit Peroxidase (Vector Laboratories). La révélation est réalisée en recourant au chromogène VIP (Vector Laboratories). Les coupes sont montées avec l'Eukitt pour les observations et les prises de photos au microscope optique. Des contrôles négatifs sont systématiquement réalisés en omettant l'AC primaire lors de l'incubation nocturne.

### **3. Microscopie électronique à transmission (MET) : Nadia Messadecq (IGBMC, Illkirch)**

Les organes ont été prélevés sur des souris perfusées au PFA 4% et ils ont été immergés dans le fixateur de MET à base d'un tampon Cacodylate (0.1M, pH 7.4) composé à 2.5% de Glutaraldehyde et à 2.5% de Paraformaldehyde. La suite du protocole a été réalisée à l'IGBMC par la plateforme de microscopie électronique.

La post-fixation se déroule à 4°C pendant 1h dans un tampon Cacodylate (0.1M, pH 7.4) contenant 1% de tétraoxyde d'osmium. La déshydratation des organes est ensuite réalisée avec des bains croissants d'alcool (50, 70, 90, 100%) et enfin l'immersion dans l'oxyde de propylène pendant 30 min. Les organes sont inclus dans une résine Epon™ 812 (Sigma-Aldrich, Saint-Louis, Missouri, USA). Les coupes semi-fines sont réalisées à 2 µm d'épaisseur avec un ultra-microtome (Leica Ultracut UCT, Leica Biosystems, Wetzlar, Germany) et colorées au bleu de toluidine. Le contrôle de la coloration se fait sous microscope optique. Le double contraste des coupes ultrafines de 70 nm d'épaisseur utilise l'acétate d'uracyle et le citrate de plomb. L'observation se déroule à 70 kV avec un microscope électronique Morgagni 268D. Les photos sont prises par une camera Mega View III camera (Soft Imaging System).

### **4. Spectrométrie de masse : Philippe Hamann (IBMC, Strasbourg)**

La protéine Vps15 et ces partenaires ont été immunoprécipités à partir des extraits protéiques issus des rétines, des reins, des cerveaux des souris. Nous avons appliqué la même méthode d'extraction décrite précédemment pour les expériences de Co-IP suivies de WB. La suite du protocole est réalisée



à l'IBMC et reprend celui déjà décrit dans la publication issue de notre laboratoire (Stoetzel et al. 2016). Les complexes protéiques contenant Vps15 ont été capturés avec un AC anti-Vps15 (NBP1-30463, Novus biological). La cible et ses interactants ont été purifiés avec des micro-billes de protéine G (Miltenyi Biotech) puis élués avec du Laemmli buffer 1X. Pour optimiser la reproductibilité des expériences, elles ont été réalisées en triplicats biologiques avec exactement la même quantité protéique pour chaque échantillon (2,5mg pour l'analyse portant sur les cerveaux par exemple). Pour les contrôles négatifs, les protéines ont subi l'étape d'immunoprécipitation sans ajout de l'AC anti-Vps15.

Pour les analyses LC-MS/MS, les protéines éluées ont été précipitées toute la nuit avec 5 volumes glacés d'acétate d'ammonium à 0,1M dans du méthanol 100%. Après centrifugation à 4°C, à 12000g, pendant 15min, les culots sont lavés 2 fois avec de l'acétate d'ammonium dans 80% de méthanol et séchés avec une pompe (Genevac centrifugation concentrator miVac Duo, Fisher Scientific). Les culots sont resuspendus dans 100µL de bicarbonate d'ammonium (50mM) et soumis à une réduction (5 mM dithiothreitol, 95°C, 10 min) et une alkylation (10 mM iodoacetamide à température ambiante pendant 20 min). Les protéines sont enfin digérées avec 200ng de trypsine (Promega). Les échantillons sont analysés soit par un dispositif nanoLC-MS/MS composé du spectromètre TripleTOF 5600 (SCIEX) couplé au système NanoLC-Ultra-2D-Plus (Eksigent) comme décrit dans Chicher et al. 2015, soit par un dispositif composé du spectromètre Q-Exactive Plus mass (Thermo Fisher Scientific) couplé au chromatographe EASY-nanoLC-1000 (Thermo Fisher Scientific) comme décrit dans Waltz et al. 2019. Chaque échantillon a été séparé sur une colonne Acclaim PepMap 100 C18 LC (Thermo Fisher Scientific) pendant 2h40 avec un gradient 300 nl/min d'acétonitrile. Les données MS sont comparées à la base de donnée *Mus musculus* Swissprot database (target database) et à la banque de séquences leurres (decoy database) (Swissprot release 2019\_07; 20432 protein sequences). Les peptides et les protéines ont été identifiés avec l'algorithme Mascot (Matrix Science). Les données ont été importées dans Proline (<http://proline.profi-proteomics.fr/>). Les protéines ont été validées pour un score Mascot =1 et un taux de fausse découverte (FDR) = 1% sur le score PSM et le score Protein Set. Proline a également été utilisé pour aligner le nombre de spectre identifié pour tous les échantillons en utilisant l'outil « Compare with SC ». Le nombre total de spectres issus de la fragmentation MS/MS a été utilisé pour quantifier chaque protéine. Pour l'analyse statistique, les données obtenues à l'issue des triplicats ont été comparées aux IP- en utilisant un logiciel « maison » comme décrit dans Chicois et al. 2018. Ce logiciel calcule le niveau d'enrichissement protéique et la p-value ajustée corrigée par la méthode *Benjamini-Hochberg* pour chacune des protéines identifiées. Ces données sont ensuite visualisables sous la forme d'un volcano plot qui représente le  $-\log_{10}$  de la p-value ajustée en fonction du  $\log_2$  de l'enrichissement. Les annotations d'enrichissement ont été analysées en recourant au logiciel Funrich.

## 5. Etudes portant sur les poissons zèbres : Uwe Strahle et Christelle Etard (KIT, Karlsruhe)

Nous avons établi une collaboration avec l'équipe de Uwe Strahle pour l'étude des poisson-zèbre *Iqce*, générés en recourant à des morpholinos. Le protocole est décrit dans (Westerfield 2000) . Brièvement, les poissons de la lignée WT AB utilisés ont été maintenus à 28°C avec un cycle alternant 14h de lumière et 10h de noir. Ils ont été nourris avec de la nourriture commerciale. Le maintien des animaux et les procédures expérimentales ont été réalisées en accord avec la réglementation de protection animale allemande (AZ35-9185.81/G-137/10). Un jour avant d'atteindre la maturité sexuelle, les animaux sont transférés dans la boîte d'accouplement où le mâle et la femelle sont maintenus séparés. Le matin de l'injection, la cloison est ôtée. La fécondation des œufs est réalisée dans les minutes qui suivent dans une solution 1× Instant Ocean salt solution (Aquarium Systems, Inc.). Ils sont collectés et transférés dans des boîtes de pétri. Les aiguilles de microinjection sont remplies avec 3 à 6 µL de solution contenant les morpholinos et l'injection est réalisé au stade « une cellule » ou « 2 cellules ». Pour vérifier que la même quantité de solution est injectée pour tous les œufs, 0,1% de rouge phénol est ajouté pour chaque injection. Les embryons injectés sont ensuite gardés à 28,5°C (ceux non fécondés ou morts sont retirés). Les embryons de poisson-zèbres sont injectés le matin entre 9h et 11h. La concentration des morpholinos nécessaire à l'injection d'un embryon au stade « une cellule » diffère d'un morpholino à l'autre. Dans le cadre du projet 3, les morpholinos ont été générés pour être dirigé contre le codon d'initiation de la traduction de *Iqce*. Le morpholino contrôle présente 5pb mal-appariées. Avant d'être capable d'observer un phénotype spécifique, plusieurs concentrations ont été testées, les 2 types de morpholinos ont finalement été injectés à une concentration de 0,6µM. Pour les expériences de « rescue », l'ARNm d'*Iqce* a été clonée dans un plasmide résistant aux morpholinos synthétisé avec SP6kit (Promega). Pour visualiser le *situs inversus*, une lignée de poisson-zèbre transgénique exprimant la GFP a été utilisée.



# III.

## Partie 1 :

La validation de l'impact  
fonctionnel de mutations  
identifiées dans des  
gènes de ciliopathies

Le projet I

portant sur *IFT27*

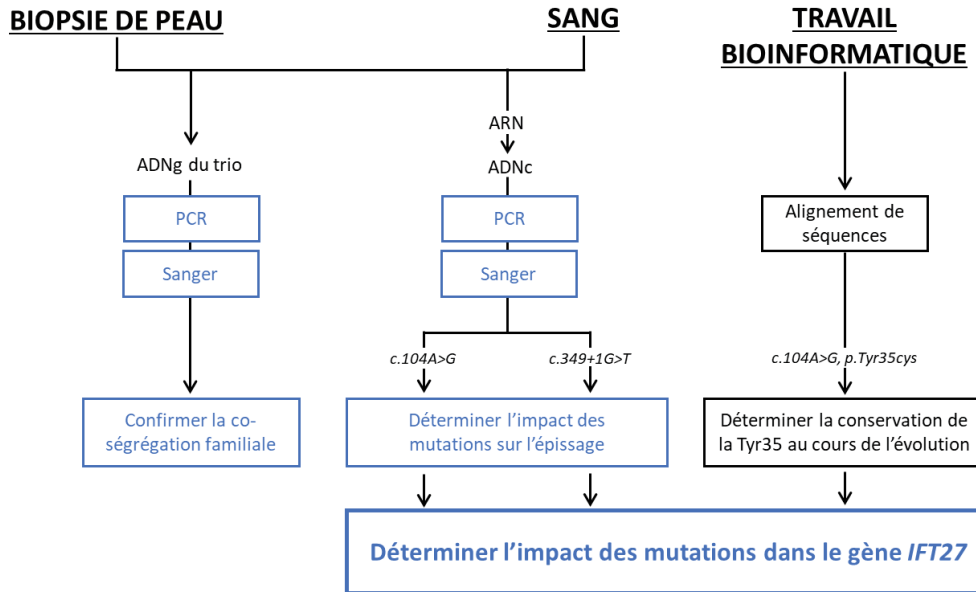
## I. Le synopsis

Ce projet publié en 2019 dans *Frontiers in genetics*, rapporte la résolution d'une famille BBS pour laquelle le patient est hétérozygote composite pour 2 mutations ayant déjà été rapportées indépendamment l'une de l'autre dans la littérature. Le patient présente un phénotype typique de BBS avec déficience intellectuelle, rétinopathie pigmentaire RP, anomalies du septum atrioventriculaire, polydactylie et obésité. La recherche de mutations réalisée par WES a révélé des mutations (c.104A>G ; p.Tyr35Cys et c. 349+1G>T) dans le gène *IFT27*. Considéré comme le 19<sup>ème</sup> gène BBS depuis la première publication d'Aldahmesh et al. 2014, nous avons rapporté la 4<sup>ème</sup> famille BBS porteuse de mutations dans ce gène.

Sanchez-Navarro et al. 2018 a publié le cas d'un patient BBS présentant une déficience intellectuelle, une RP, une polydactylie, une obésité et une atteinte rénale. Le patient rapporté est hétérozygote composite pour *IFT27*, l'un des variants correspond au faux-sens c.104A>G ; p.Tyr35Cys mais la co-ségrégation familiale n'a pas été validée. Cette mutation est localisée au niveau de l'exon 2 d'*IFT27*, dans une région hautement conservée. La Tyrosine 35 se trouve au niveau du domaine « switch 1 » de la protéine impliqué dans les interactions protéine-protéine et la fonction GTPase. Le changement d'AA est donc prédit comme pathogène et rend la protéine IFT27 non fonctionnelle. Le second variant c.349+1G>T a été précédemment rapporté par Quélin et al. 2018 dans le cas d'un fœtus présentant de nombreuses malformations dont une polydactylie et une atteinte rénale. Dans ce cas, la validation fonctionnelle de la mutation n'a pas pu être réalisée.

La substitution c.349+1G>T est localisée au niveau du site donneur d'épissage de l'exon 5. Même si la pathogénicité de ce variant semble évidente, nous avons déterminé son impact et confirmé sa pathogénicité sur le processus d'épissage en analysant l'ARN. Ce matériel a été extrait à partir d'une biopsie de peau et d'un prélèvement sanguin du patient. Cette analyse a révélé la présence de transcrits inexistant chez un contrôle. La séquence sanger de ces transcrits nous a confirmé soit la perte des exons 5 et 6 soit la perte des exons 4 à 6.

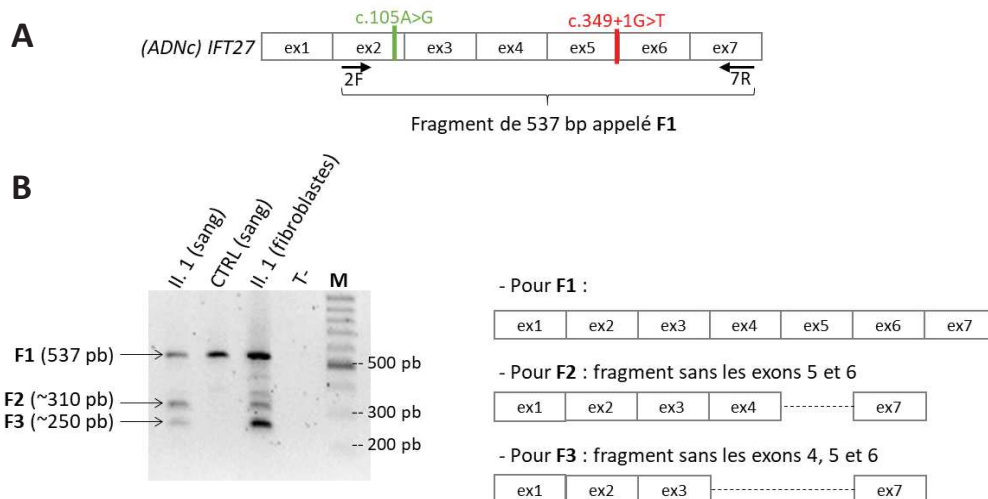
Finalement, notre étude rapporte le cas d'une famille pour laquelle nous avons mis en place une série d'expériences (récapitulée dans la Figure 35) afin de déterminer l'impact des 2 mutations identifiées. Nous avons validé la co-ségrégation familiale des 2 mutations. L'une rend la protéine produite non fonctionnelle. La seconde impacte l'épissage et crée des transcrits alternatifs. S'ils sont traduits, ils donneront des protéines non fonctionnelles. S'ils sont dégradés, la quantité de protéine IFT27 issue du transcrit normal sera fortement réduite. L'absence de la protéine IFT27 ou la présence d'une protéine non fonctionnelle ont déjà été rapportées comme responsables d'un phénotype de BBS (Eguether et al. 2014a) comme celui observé chez les patients.



**Figure 35 :** Le logigramme. Résumé des différentes expériences réalisées pour répondre aux objectifs initialement formulés. Mon travail est identifié en **bleu**.

## II. Les données supplémentaires.

Les PCR réalisées dans le cadre de ce projet ont utilisé l'ADNc issu de la rétrotranscription de l'ARN extrait des fibroblastes du patient ou d'un contrôle. J'ai également réalisé la même PCR (Figure 36A) avec le matériel extrait du sang des individus afin de déterminer si l'impact de la mutation était le même dans des tissus différents. Dans les 2 cas, nous avons mis en évidence chez le patient des bandes surnuméraires de tailles inférieures au fragment attendu F1 (Figure 36B) correspondant à des transcrits issus d'un épissage alternatif pour lesquels certains exons ne sont pas présents (voir Figure 1D de l'article).



**Figure 36 :** Comparaison des résultats de PCR sur l'ADNc selon le tissu. **A.** Représentation schématique de l'ADNc IFT27 avec le positionnement des amorces et des mutations. **B.** Résultats de la PCR avec le couple d'amorces 2F-7R, sur l'ADNc issu de l'extraction d'ARN du sang ou des fibroblastes. Représentation schématique expliquant les fragments obtenus chez le patient. T- : témoin négatif, ex : exon, CTRL : contrôle.

### III. La conclusion et les perspectives.

Avec ce projet, nous avons résolu le cas d'une famille BBS en montrant la pathogénicité de 2 mutations identifiées dans le gène *IFT27/BBS19*. Si nous avons voulu poursuivre la caractérisation de la mutation d'épissage, nous aurions pu déterminer la quantité de chaque transcrit (F2 et F3 vs F1) en créant des amorces allèles spécifiques et recourant à une qRT-PCR dépendante de sondes Taqman. Nous aurions pu faire des WB sur des extraits protéiques totaux afin de déterminer s'il existe plusieurs protéines IFT27 produites chez le patient en utilisant un AC dirigé contre la partie N-terminale de la protéine. Si nous avons voulu poursuivre la caractérisation de la mutation faux-sens, nous aurions pu envisager de faire un co-marquage de la protéine avec l'alpha-tubuline-acétylée par immunofluorescence sur des cellules en conditions ciliées afin de déterminer si la localisation d'IFT27 est maintenue dans le cil malgré la présence de la mutation. La Tyrosine 35 est un AA situé dans un domaine impliqué dans les interactions protéine-protéine, nous aurions pu envisager une Co-IP anti-IFT25 suivie d'un WB dirigé contre IFT27 afin de déterminer si l'interaction connue entre ces 2 protéines est maintenue malgré la mutation.

### IV. L'article



# Identification and Characterization of Known Biallelic Mutations in the *IFT27 (BBS19)* Gene in a Novel Family With Bardet-Biedl Syndrome

Elise Schaefer<sup>1,2</sup>, Clarisse Delvallée<sup>1</sup>, Laura Mary<sup>3</sup>, Corinne Stoetzel<sup>1</sup>,  
Véronique Geoffroy<sup>1</sup>, Caroline Marks-Delesalle<sup>4</sup>, Muriel Holder-Espinasse<sup>5</sup>,  
Jamal Ghoumid<sup>6</sup>, Hélène Dollfus<sup>1,2</sup> and Jean Muller<sup>1,3\*</sup>

<sup>1</sup> Laboratoire de Génétique Médicale, Institut de Génétique Médicale d'Alsace, INSERM U1112, Fédération de Médecine Translationnelle de Strasbourg, Université de Strasbourg, Strasbourg, France, <sup>2</sup> Service de Génétique Médicale, Hôpitaux Universitaires de Strasbourg, Strasbourg, France, <sup>3</sup> Laboratoire de Diagnostic Génétique, Hôpitaux Universitaires de Strasbourg, Strasbourg, France, <sup>4</sup> Service d'Exploration de la Vision et Neuro-Ophthalmologie, Hôpital Roger Salengro, CHRU Lille, Lille, France, <sup>5</sup> Clinical Genetics Department, Guy's Hospital, London, United Kingdom, <sup>6</sup> Service de Génétique Clinique, Hôpital Jeanne de Flandre, CHRU Lille, Lille, France

## OPEN ACCESS

### Edited by:

Musharraf Jelani,  
King Abdulaziz University,  
Saudi Arabia

### Reviewed by:

Muhammad Tariq,  
University of Tabuk, Saudi Arabia  
Sulman Basit,  
Quaid-i-Azam University, Pakistan

### \*Correspondence:

Jean Muller  
jeanmuller@unistra.fr

### Specialty section:

This article was submitted to  
Genetic Disorders,  
a section of the journal  
Frontiers in Genetics

**Received:** 29 October 2018

**Accepted:** 15 January 2019

**Published:** 30 January 2019

### Citation:

Schaefer E, Delvallée C, Mary L,  
Stoetzel C, Geoffroy V,  
Marks-Delesalle C,  
Holder-Espinasse M, Ghoumid J,  
Dollfus H and Muller J (2019)  
Identification and Characterization  
of Known Biallelic Mutations  
in the *IFT27 (BBS19)* Gene in a Novel  
Family With Bardet-Biedl Syndrome.  
*Front. Genet.* 10:21.  
doi: 10.3389/fgene.2019.00021

Bardet-Biedl syndrome (BBS; MIM 209900) is a rare ciliopathy characterized by retinitis pigmentosa, postaxial polydactyly, obesity, hypogonadism, cognitive impairment and kidney dysfunction. Mutations in 22 BBS genes have been identified to cause the disease. We report a family with typical BBS features (retinitis pigmentosa, postaxial polydactyly, obesity, cognitive impairment, and atrioventricular septal defect) mutated in *IFT27/BBS19*. *IFT27* is part of the Intraflagellar transport (IFT), a bidirectional mechanism allowing the protein motility within the cilia. Using whole exome sequencing, two compound heterozygous mutations were found in the proband (NM\_006860.4:c.[104A > G];[349+1G > T], p.[Tyr35Cys];[?]) consistent with the expected autosomal recessive inheritance mode. These two mutations have already been reported but independently in other families and lacking either familial segregation or functional validation. This is the third report of *IFT27* mutations in BBS patients confirming *IFT27* as a BBS gene (*BBS19*). Mutations in IFT genes (*IFT27*, *IFT172* and *IFT74*) confirm the IFT-pathway as a pathomechanism for BBS.

**Keywords:** Bardet-Biedl syndrome, *IFT27* gene, *BBS19*, ciliopathy, whole exome sequencing

## INTRODUCTION

Bardet-Biedl syndrome (BBS; OMIM 209900) is a recessive and genetically heterogeneous ciliopathy defined by the association of retinitis pigmentosa (RP), postaxial polydactyly, obesity, hypogonadism, kidney dysfunction and cognitive impairment. Mutations in 22 genes have been identified of which *WDPCP/BBS15*, *LZTFL1/BBS17*, *BBIP1/BBS18*, *IFT27/BBS19*, *IFT172/BBS20*, *C8Orf37/BBS21*, and *IFT74/BBS22* only account for one or a few families (Khan et al., 2016). Among the last 4 genes identified, 3 (*IFT27/BBS19*, *IFT172/BBS20*, and the candidate locus *IFT74/BBS22*) are members of the intraflagellar transport machinery (IFT) delineating a possible novel mechanisms for the BBS (Aldahmesh et al., 2014; Bujakowska et al., 2015; Lindstrand et al., 2016). The IFT machinery is composed of 2 large protein complexes named IFT-A and IFT-B

required, respectively, for retrograde and anterograde transport in the cilia. This machinery is part of an essential process for the assembly and the maintenance of the cilia (Taschner and Lorentzen, 2016). Most of the reported mutations in these genes, and in particular for *IFT172*, have been identified in patients with skeletal forms of ciliopathies including the Jeune asphyxiating thoracic dystrophy (OMIM 208500), Mainzer–Saldino syndrome (OMIM 266920), the Sensenbrenner syndrome (OMIM 218330).

In this study, we report and validate 2 compound heterozygous mutations in *IFT27* confirming the 19th BBS locus in a family with a clear BBS phenotype.

## MATERIALS AND METHODS

### Subjects

Bardet-Biedl syndrome studies have been approved by our Institutional Review Board “Comité de Protection des Personnes” (EST IV, N°DC-20142222) and written informed consent for research and publication was obtained from each participant. We declare that the present research complies with the declaration of Helsinki.

### Whole-Exome Sequencing

Whole exome sequencing (WES) was performed by the IGBMC Microarray and Sequencing platform. Genomic DNA (2 µg) was sheared to obtain a mean fragment size of 150 bp using the Covaris E210 (KBioscience, Herts, United Kingdom) followed by library preparation using the Agilent SureSelect XT Human all exon V6 (Agilent Technologies, Santa Clara, CA, United States PN G7530-90000, protocol B4). Sequencing was performed on an Illumina HiSeq 4000 (Illumina, San Diego, CA, United States) to generate 100-bp paired-end reads following the manufacturer’s protocols.

### Bioinformatics Analysis

Image analysis and base calling were performed using CASAVA v1.8.2 (Illumina) and the 109,658,782 reads were mapped to the reference human genome (GRCh37/hg19) using BWA v0.7.5a (Li and Durbin, 2009) leading to 95.89% of the bases covered at least by 20x. GATK UG v3.2-2 was used to call SNV and indel variations (DePristo et al., 2011). Annotation and ranking of SNV and indel were performed by VaRank (Geoffroy et al., 2015) in combination with the Alamut Batch software (Interactive Biosoftware, Rouen, France). Very stringent filtering criteria were applied to filter out non-pathogenic variants (**Supplementary Table S1**): (i) variants represented with an allele frequency of more than 1% in public variation databases including the 1000Genomes (Genomes Project et al., 2015), the gnomAD database (Lek et al., 2016), the DGV database (MacDonald et al., 2014) or our internal exome database, (ii) variants in 5′ and 3′ UTR, downstream, upstream or intronic locations and synonymous without pathogenic prediction of local splice effect, (iii) variants not in the ciliary genes (Nevers et al., 2017). Structural variants were predicted using CANOES (Backenroth et al., 2014) and annotated by AnnotSV (Geoffroy et al., 2018). Our analysis was focused on compound heterozygous

and homozygous variants consistent with a recessive mode of transmission. The *IFT27* nomenclature is based on the RefSeq accession number NM\_006860.4 (O’Leary et al., 2016).

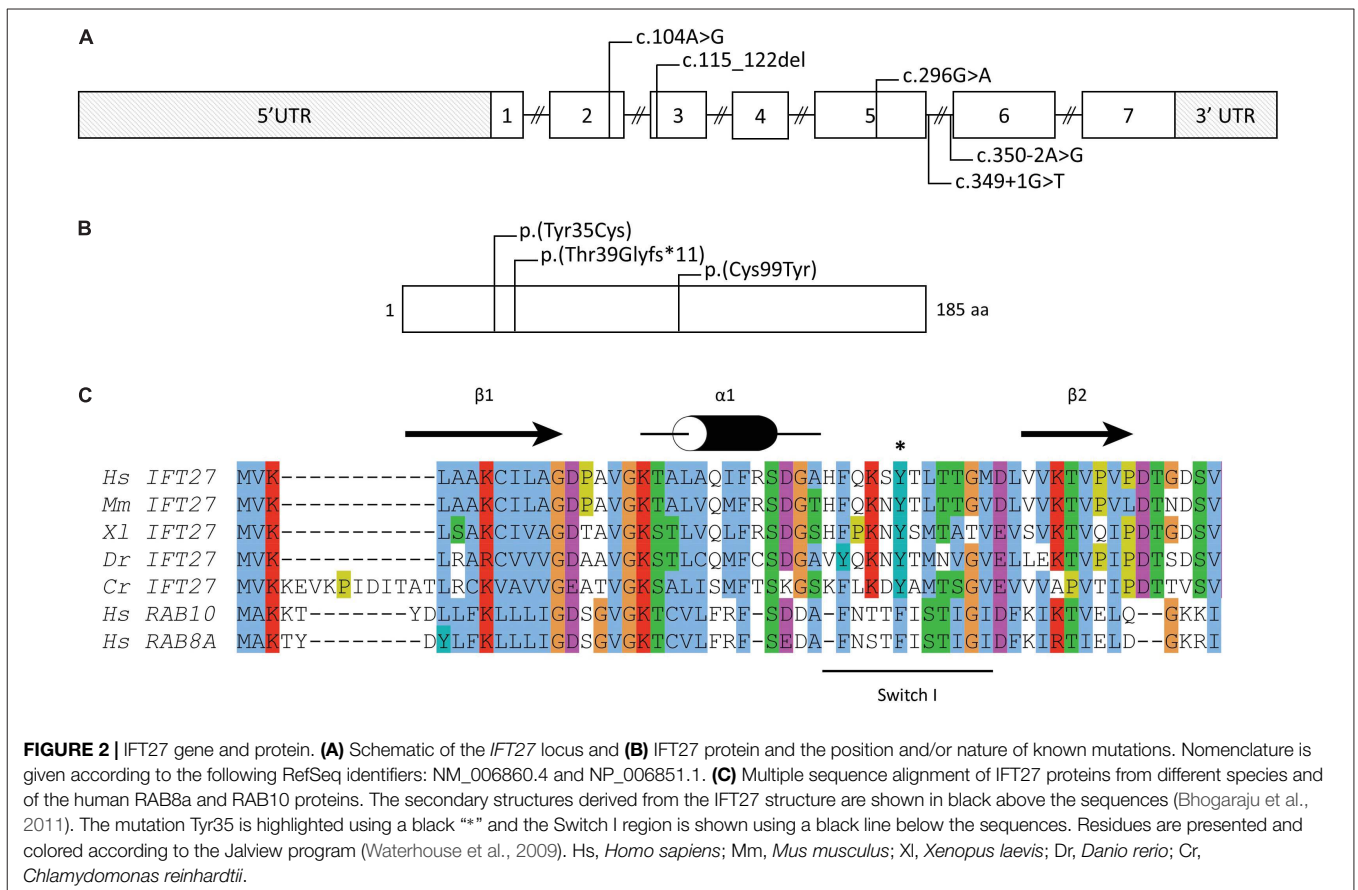
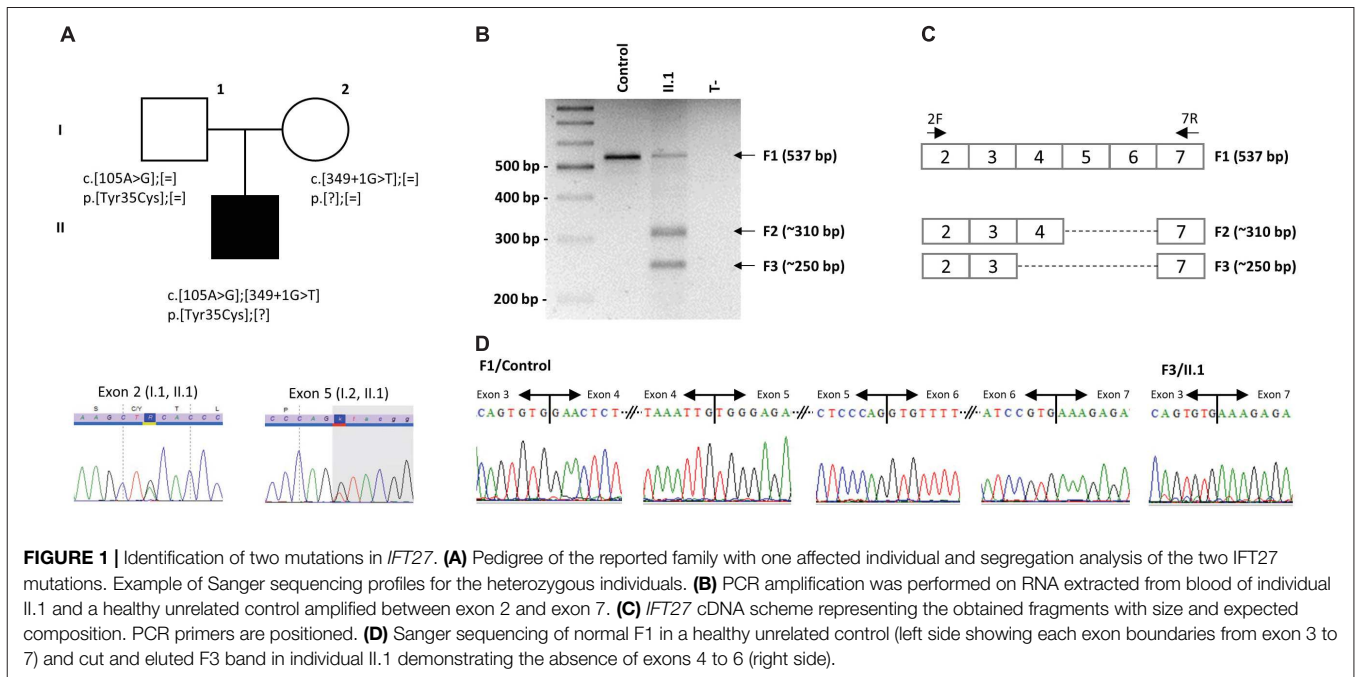
### RNA Extraction, cDNA Synthesis, and Sanger Sequencing

RNA was extracted from the patient’s blood using the PAXgene Blood RNA Kit (PreAnalytiX GmbH, Hombrechtikon, Switzerland). Reverse transcription was performed on 200 ng of RNA using the BioRad iScript cDNA® Synthesis Kit (BioRad, Hercules, CA, United States). PCR was performed on the cDNA using the Mastercycler ep. Gradient S thermocycler (Eppendorf, Germany). Specific fragment analysis was done by cutting the agarose gel around the fragment of interest (e.g., F3) and eluted in Tris–EDTA buffer 10.1 overnight. The elution product (5 µL) was reamplified by PCR using the same conditions to obtain a unique fragment. Bidirectional sequencing of the purified PCR products was performed by the GATC Sequencing Facilities (Konstanz, Germany). Primers used are summarized in **Supplementary Table S2**.

## RESULTS

The patient is the only child of an unrelated couple without personal or familial medical history. The patient was born at 39 weeks of amenorrhea with the following parameters: weight at 3210 g, height at 51 cm and head circumference at 35 cm. Mesoaxial polydactyly of the right hand with a Y-shaped metacarpian and syndactyly between the 5th and the 6th fingers and postaxial polydactyly of the right foot was noticed at birth. Partial atrioventricular septal defect was also diagnosed and operated at 5 weeks old. Mitral insufficiency persisted after operation and was operated at 2 years old. Renal ultrasound was normal. At birth, cerebral ultrasound showed isolated thin corpus callosum, not confirmed on cerebral MRI. Secondly, the patient presented with delayed psychomotor development: he walked at 25-month-old and had delayed language (5 words at 2 years; 10 words at 3 years and sentences at 5 years). Audition was normal. The patient received specialized education. Progressively, he developed obesity: at 2 years old, 16.8 kg (+ 3SD) for 93 cm (+ 2SD) and normal head circumference (49 cm); at 3 years old, 22.9 kg (> +3SD) for 101.5 cm (+ 2SD). At 7 years old, his BMI was 25. Initially, ophthalmologic examination revealed myopia at 2 years old and alternate divergent strabismus at 3 years old. Myopia was scalable was noticed at 4 years old. A cone-rod dystrophy was diagnosed at 7 years old.

Several diagnostic hypotheses led to the non-conclusive exploration of *GLI3* (Pallister-Hall), *OFD1* (Orofaciodigital syndrome I) and the known BBS genes at that time (*BBS1–BBS18*) using Sanger and targeted exome sequencing (Muller et al., 2010; Redin et al., 2012). This also included array-CGH screening. Finally, WES was performed on the index case and led to the identification of 2 already reported variations in the *IFT27/BBS19* gene co-segregating in the family (**Figure 1A**). The first variation, c.104A > G (p.Tyr35Cys), is affecting a highly



conserved position in the protein localized in the switch I region of this Rab-like GTPase (Figures 2A,B). The variation is absent from the variation databases (i.e., 1000 genomes and gnomAD)

and predicted as affecting the protein function (Supplementary Table S3). Given its relatively close position toward the end of exon 2 (Figure 2A), a possible splice effect was assessed on



the patient's RNA extracted from blood and failed to reveal any alteration (**Supplementary Figure S1**).

Given its localization, the second variation, c.349 + 1G > T, is predicted to abolish the donor splice site of exon 5 (**Figure 2A** and **Supplementary Table S3**). It is already described in gnomAD with 21 heterozygous alleles in the European (Non-Finnish) cohort leading to an allele frequency of 0.0085% compatible with the BBS rare condition. Analysis of the patient's RNA extracted from blood revealed a mix of alternatively spliced isoforms not found in controls (**Figures 1B–D**). This includes removal of exons 5 + 6 or 4 + 5 + 6 that are predicted to cause an in frame deletion of a significant part of the protein (76 or 96aa out of 185aa).

## DISCUSSION

Bardet-Biedl syndrome is an emblematic ciliopathy characterized by a clinical and genetic heterogeneity. Twenty-two BBS genes have been identified so far and mutations are found in about 80% of the known BBS patients (Redin et al., 2012; Khan et al., 2016). The last genes identified account only for a small fraction of the mutation load, emphasizing the importance to screen more families and increase our knowledge for these genes (e.g., type of mutations and clinical manifestations).

In our study, we identified two compound heterozygous mutations in *IFT27* (NM\_006860.4: c.[104A > G];[349 + 1G > T]) by WES in a child presenting with the classical BBS signs. *IFT27* is part of the Rab like GTPase that is predicted to have low GTPase activity. The gene is composed of 7 exons coding for a 186 aa protein.

Our first identified mutation, the c.104A > G is located in exon 2 leading to the drastic change of a tyrosine to a cysteine (p.Tyr35Cys). Localized in the connecting region between  $\alpha$ -helix 1 and  $\beta$ -sheet 2 in the so called "switch I" region (**Figure 2C**), a highly flexible region upon nucleotide binding pocket not crystallized (Bhogaraju et al., 2011) with an important role in GDP/GTP exchange and with interacting partners (Pylypenko et al., 2018), the missense is highly susceptible to affect the function of *IFT27*. The second mutation, c.349 + 1G > T, located in the donor splice site of exon 5 was predicted to affect the splicing of *IFT27* that we have demonstrated with a skipping of several exons of the gene. This mutation has been already reported in a fetus with a severe phenotype but the effect could not be assessed at that time (Quelin et al., 2018).

Initially the first identified mutations in *IFT27* were reported in a consanguineous BBS family presenting with RP, obesity, polydactyly of all extremities, mild intellectual disability, renal failure and hypogenitalism. The patients carried a homozygous missense (c.296G > A, p.Cys99Tyr) predicted to affect the stability of the protein (Aldahmesh et al., 2014).

Very recently, another BBS case with RP (cone rod dystrophy), obesity, polydactyly, maturation and learning delay and chronic renal failure has been reported but they could not to assess the segregation and the effect on the protein (c.[104A>G];[350-2A>G], p.[Tyr35(;?)]) (Sanchez-Navarro et al., 2018).

Lastly, 2 loss of function mutations (c.[115\_122del];[349+1G > T], p.[Thr39Glyfs\*11];[?]) have been reported in a fetus

presenting with a severe ciliopathy with short ribs polydactyly type II (SRPII or Majewski syndrome) and/or Pallister-Hall syndrome (Quelin et al., 2018). Interestingly, Pallister-Hall syndrome was also evoked for our patient given the association of cardiopathy and polydactyly with Y-Shaped metacarpian. The authors made the hypothesis that the loss of function mutations could explain the severe phenotype.

Comparing the reported patients including the fetus (**Supplementary Table S4**), all had postaxial polydactyly. All living patients had RP, learning difficulties and obesity. Moreover, one patient of the initial BBS family has deafness and hypogenitalism. The same family and the fetus have renal anomalies (renal hypoplasia to renal agenesis) contrasting with our patient who has no renal involvement. Finally, our patient and one patient of the initial family presented with a cardiopathy, and unexpectedly not found in the more severe fetal form. Interestingly some patients have unusual clinical presentation including mesoaxial polydactyly with Y-shaped metacarpian associated with postaxial polydactyly (our patient) or preaxial (fetus). The Y-shaped metacarpian has been already correlated with mutation in *BBS17/LZTFL1* (Schaefer et al., 2014) that is now extended to *IFT27*. One can notice also that 2 patients had cone-rod dystrophy (our patient and family 2), a rare condition already described in BBS (Scheidecker et al., 2015). These features show the clinical variability associated to *IFT27* mutations from severe lethal forms to classical BBS that might be revised when new patients will be described in the future.

Mutations in *IFT* genes were recently reported in several BBS patients (Bujakowska et al., 2015; Lindstrand et al., 2016; Schaefer et al., 2016). *IFT27*, *IFT74*, and *IFT172* encode components of the *IFT*-B complex required for anterograde transport of ciliary proteins (Taschner and Lorentzen, 2016). Mutations in *IFT* genes are implicated in different ciliopathies, principally in skeletal ciliopathies (Jeune asphyxiating thoracic dystrophy, Sensenbrenner syndrome and Mainzer-Saldino syndrome) but also in isolated RP and Senior-Løken syndrome (associating RP to nephronophthisis). This report confirms the *IFT*-pathway as a new pathomechanism for BBS.

## CONCLUSION

We identified the third BBS family mutated in *IFT27*. We confirmed *IFT27* as the 19th BBS gene delineating the typical and complete BBS phenotype for those rarely mutated patients and the implication of *IFT*-pathway in the BBS pathogenesis. This report illustrated the usefulness of WES sequencing to identify mutations in highly heterogeneous genetic disorders. Finally, our observation is a good illustration of the clinical and molecular continuum between the different ciliopathies.

## DATA AVAILABILITY

All variants have been submitted to ClinVar using the following range of accessions numbers SCV000839871 and SCV000839872 (<https://www.ncbi.nlm.nih.gov/clinvar/>).

## AUTHOR CONTRIBUTIONS

HD and JM conceived the study. CM-D, JG, and MH-E collected the clinical information. CD, CS, ES, and LM analyzed the data. VG conducted the bioinformatics analysis. ES and JM drafted the manuscript. HD and JM supervised the entire study. All the authors approved the final version of manuscript.

## FUNDING

CD was supported by a FRM grant (ECO20170637509). Whole exome sequencing was performed by the IGBMC Microarray and Sequencing platform, a member of the 'France Génomique' consortium (ANR-10-INBS-0009) and funded by "La Fondation Maladie Rare" through the programme «High throughput sequencing and rare diseases».

## REFERENCES

- Aldahmesh, M. A., Li, Y., Alhashem, A., Anazi, S., Alkuraya, H., Hashem, M., et al. (2014). IFT27, encoding a small GTPase component of IFT particles, is mutated in a consanguineous family with Bardet-Biedl syndrome. *Hum. Mol. Genet.* 23, 3307–3315. doi: 10.1093/hmg/ddu044
- Backenroth, D., Homsy, J., Murillo, L. R., Glessner, J., Lin, E., Brueckner, M., et al. (2014). CANOES: detecting rare copy number variants from whole exome sequencing data. *Nucleic Acids Res.* 42:e97. doi: 10.1093/nar/gku345
- Bhogaraju, S., Taschner, M., Morawetz, M., Basquin, C., and Lorentzen, E. (2011). Crystal structure of the intraflagellar transport complex 25/27. *EMBO J.* 30, 1907–1918. doi: 10.1038/emboj.2011.110
- Bujakowska, K. M., Zhang, Q., Siemiatkowska, A. M., Liu, Q., Place, E., Falk, M. J., et al. (2015). Mutations in IFT172 cause isolated retinal degeneration and Bardet-Biedl syndrome. *Hum. Mol. Genet.* 24, 230–242. doi: 10.1093/hmg/ddu441
- DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 43, 491–498. doi: 10.1038/ng.806
- Genomes Project, C., Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., et al. (2015). A global reference for human genetic variation. *Nature* 526, 68–74. doi: 10.1038/nature15393
- Geoffroy, V., Herenger, Y., Kress, A., Stoetzel, C., Piton, A., Dollfus, H., et al. (2018). AnnotSV: an integrated tool for structural variations annotation. *Bioinformatics* 34, 3572–3574. doi: 10.1093/bioinformatics/bty304
- Geoffroy, V., Pizot, C., Redin, C., Piton, A., Vasli, N., Stoetzel, C., et al. (2015). VaRank: a simple and powerful tool for ranking genetic variants. *PeerJ* 3, e796. doi: 10.7717/peerj.796
- Khan, S. A., Muhammad, N., Khan, M. A., Kamal, A., Rehman, Z. U., and Khan, S. (2016). Genetics of human Bardet-Biedl syndrome, an updates. *Clin. Genet.* 90, 3–15. doi: 10.1111/cge.12737
- Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536, 285–291. doi: 10.1038/nature19057
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324
- Lindstrand, A., Frangakis, S., Carvalho, C. M., Richardson, E. B., McFadden, K. A., Willer, J. R., et al. (2016). Copy-number variation contributes to the mutational load of bardet-biedl syndrome. *Am. J. Hum. Genet.* 99, 318–336. doi: 10.1016/j.ajhg.2015.04.023

## ACKNOWLEDGMENTS

We would like to thank the patient's family for their participation. We would like also to acknowledge the members of the diagnostic laboratories (molecular and cytogenetic units, in particular Manuela Antin, Anne-Sophie Leuvrey and Elsa Nourisson) at the University Hospital in Strasbourg for technical assistance. The computing resources for this work were provided by the BICS and BISTRO bioinformatics platforms in Strasbourg.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2019.00021/full#supplementary-material>

- MacDonald, J. R., Ziman, R., Yuen, R. K., Feuk, L., and Scherer, S. W. (2014). The database of genomic variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res.* 42, D986–D992. doi: 10.1093/nar/gkt958
- Muller, J., Stoetzel, C., Vincent, M. C., Leitch, C. C., Laurier, V., Danse, J. M., et al. (2010). Identification of 28 novel mutations in the Bardet-Biedl syndrome genes: the burden of private mutations in an extensively heterogeneous disease. *Hum. Genet.* 127, 583–593. doi: 10.1007/s00439-010-0804-9
- Nevers, Y., Prasad, M. K., Poidevin, L., Chennen, K., Allot, A., Kress, A., et al. (2017). Insights into ciliary genes and evolution from multi-level phylogenetic profiling. *Mol. Biol. Evol.* 34, 2016–2034. doi: 10.1093/molbev/msx146
- O'Leary, N. A., Wright, M. W., Brister, J. R., Ciuffo, S., Haddad, D., McVeigh, R., et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* 44, D733–D745. doi: 10.1093/nar/gkv1189
- Pylpyenko, O., Hammich, H., Yu, I. M., and Houdusse, A. (2018). Rab GTPases and their interacting protein partners: structural insights into Rab functional diversity. *Small GTPases* 9, 22–48. doi: 10.1080/21541248.2017.1336191
- Quelin, C., Loget, P., Boutaud, L., Elkhartoufi, N., Milon, J., Odent, S., et al. (2018). Loss of function IFT27 variants associated with an unclassified lethal fetal ciliopathy with renal agenesis. *Am. J. Med. Genet. A* 176, 1610–1613. doi: 10.1002/ajmg.a.38685
- Redin, C., Le Gras, S., Mhamdi, O., Geoffroy, V., Stoetzel, C., Vincent, M. C., et al. (2012). Targeted high-throughput sequencing for diagnosis of genetically heterogeneous diseases: efficient mutation detection in Bardet-Biedl and Alstrom syndromes. *J. Med. Genet.* 49, 502–512. doi: 10.1136/jmedgenet-2012-100875
- Sanchez-Navarro, R. J. I., da Silva, L., Blanco-Kelly, F., Zurita, O., Sanchez-Bolivar, N., Villaverde, C., et al. (2018). Combining targeted panel-based resequencing and copy-number variation analysis for the diagnosis of inherited syndromic retinopathies and associated ciliopathies. *Sci. Rep.* 8:5285. doi: 10.1038/s41598-018-23520-1
- Schaefer, E., Lauer, J., Durand, M., Pelletier, V., Obringer, C., Claussmann, A., et al. (2014). Mesoaxial polydactyly is a major feature in Bardet-Biedl syndrome patients with LZTFL1 (BBS17) mutations. *Clin. Genet.* 85, 476–481. doi: 10.1111/cge.12198
- Schaefer, E., Stoetzel, C., Scheidecker, S., Geoffroy, V., Prasad, M. K., Redin, C., et al. (2016). Identification of a novel mutation confirms the implication of IFT172 (BBS20) in Bardet-Biedl syndrome. *J. Hum. Genet.* 61, 447–450. doi: 10.1038/jhg.2015.162

- Scheidecker, S., Hull, S., Perdomo, Y., Studer, F., Pelletier, V., Muller, J., et al. (2015). Predominantly cone-system dysfunction as rare form of retinal degeneration in patients with molecularly confirmed Bardet-Biedl syndrome. *Am. J. Ophthalmol.* 160, 364.e1–372.e1. doi: 10.1016/j.ajo.2015.05.007
- Taschner, M., and Lorentzen, E. (2016). The Intraflagellar Transport Machinery. *Cold Spring Harb. Perspect. Biol.* 8:a028092. doi: 10.1101/cshperspect.a028092
- Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., and Barton, G. J. (2009). Jalview version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191. doi: 10.1093/bioinformatics/btp033

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Schaefer, Delvallée, Mary, Stoetzel, Geoffroy, Marks-Delesalle, Holder-Espinasse, Ghomid, Dollfus and Muller. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

## SUPPLEMENTARY DATA

	SNV	indel	SV
Total number of variants	92351	14676	12
After exclusion of variants with an allele frequency >1% (gnomAD, 1000G, internal exome database or DGV) (1-3)	1847	284	6
After exclusion of SNV/indel found at the homozygous state in gnomAD and in our internal exome database	1531	255	6
After exclusion of SNV/indel in 5' and 3'UTR, downstream, upstream, intron and synonymous locations without local splice effect prediction	559	84	6
After exclusion of missense without SIFT, PPH2 or PhastCons deleterious effect	500	84	6
After exclusion of variants absent from ciliary gene list (4)	84	8	0
After selection of variants consistent with recessive transmission (compound heterozygous, homozygous variants)	17 compounds heterozygous in: <i>ACADM</i> , <i>BOLA3</i> , <i>DAWI</i> , <i>DIMT1</i> , <i>DNAH8</i> , <i>EPB41L1</i> , <i>IFT27</i> , <i>KDM4B</i> , <i>KNTC1</i> , <i>LRRC56</i> , <i>LRTOMT</i> , <i>PPP3CC</i> , <i>SEC63</i> , <i>SMARCA5</i> , <i>SYNE2</i> , <i>TCTNI</i> and <i>USP25</i>		

**Supplementary Table 1.** Summary of the whole exome sequencing results. SNV: single nucleotide variation, indel: gain or loss of up to 50 nucleotides at a single locus, SV: structural variation. Exclusion of SV with a DGV frequency >1% is done only with studies of more than 1000 individuals.

Application	Exon	Forward (5'-3')	Reverse (5'-3')	Size
DNA	2	CATAGGGCGTGTTTCAGTGGGAAGA	CCAAGTTGCTTGTGGCCATTCTG	600 bp
	5	TCTACCGAGTTGGGAGGTTTCACA	TCCCTGGGGTTAGCTTTCTGGAAT	710 bp
cDNA	2 to 7	CAGATCTTCCGCAGTGATGG	TCTGTCTTCTCCGGTTGTGC	537 bp
	1 to 3 4	CCATGGTGAAGCTGGCAG	GTTCCACACTGTCTCCCGT	177 bp

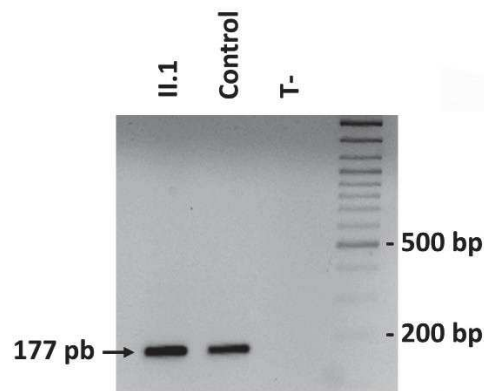
**Supplementary Table 2.** List of primers used in this study (*IFT27*: NM\_006860.4).

	Exon	MaxEntScan	NNSplice	SSF	Align-GVGD	PolyPhen-2	SIFT
c.104A>G, p.Y35C	2				Deleterious (Class 65)	Probably damaging (1.000)	Affect protein function (0.00)
c.349+1	5	-100%	-100%	-100%			

**Supplementary Table 3.** Prediction of the effect of the two variations on the protein function. Variant effect on the nearest splice site was predicted using MaxEntScan (5), NNSplice (6) and Splice Site Finder (7) by calculating score change between the wild type and the mutated sequences expressed as a percent differences. Missense effect was assessed by Align-GVGD (8), PolyPhen-2 (9) and SIFT (10).

	Ethnicity	Consanguinity	Segregation	Genotype	RP	Obesity	Polydactyly	CI	Renal defect	Genital defect	Anosmia	Others	Reference
Family 1	Saudi Arabia	Y	Y	c.[296G>A];[296G>A] p.[(Cys99Tyr)];[(Cys99Tyr)]	Y	Y	Y	Y	Y	Y	Y	Deafness, atopy, congenital heart disease, fatty liver	(11)
	Saudi Arabia		Y	c.[296G>A];[296G>A] p.[(Cys99Tyr)];[(Cys99Tyr)]	Y	Y	Y	Y	N	N	Y	Atopy, fatty liver	(11)
Family 2		N	N	c.[104A>G(;);350-2A>G] p.[Tyr35Cys(;);?]	Y <sup>b</sup>	Y	Y	Y	Y	N	N		(12)
Family 3 <sup>a</sup>		N	Y	c.[115_122del];[349+1G>T] p.[Thr39Glyfs*11];[?]			Y <sup>c</sup>		Y	Y		Short ribs, imperforate anus, laryngeal anomalies	(13)
Family 4	Caucasian	N	Y	c.[104A>G];[349+1G>T] p.[(Tyr35Cys)];[?]	Y <sup>b</sup>	Y	Y <sup>c</sup>	Y	N	N	N	Atrioventricular septal defect, Y-shaped metacarpian	This study

**Supplementary Table 4.** Genotype and phenotype summary of the reported *IFT27* patients. RP: retinitis pigmentosa, CI: cognitive impairment, <sup>a</sup>this case is a fetus, <sup>b</sup>cone-rod dystrophy, <sup>c</sup>mesoaxial.



**Supplementary Figure 1.** Cys35Tyr possible splice effect assessment. PCR amplification was performed on RNA extracted from blood of individual II.1 and a healthy unrelated control amplified between exon 1 and exon 3/4.

## SUPPLEMENTARY REFERENCES

1. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. *Nature* (2015) 526(7571):68-74. doi: 10.1038/nature15393. PubMed PMID: 26432245; PubMed Central PMCID: PMC4750478.
2. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* (2016) 536(7616):285-91. doi: 10.1038/nature19057. PubMed PMID: 27535533; PubMed Central PMCID: PMC45018207.
3. MacDonald JR, Ziman R, Yuen RK, Feuk L, Scherer SW. The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res* (2014) 42(Database issue):D986-92. doi: 10.1093/nar/gkt958. PubMed PMID: 24174537; PubMed Central PMCID: PMC3965079.
4. Nevers Y, Prasad MK, Poidevin L, Chennen K, Allot A, Kress A, et al. Insights into Ciliary Genes and Evolution from Multi-Level Phylogenetic Profiling. *Mol Biol Evol* (2017) 34(8):2016-34. Epub 2017/05/02. doi: 10.1093/molbev/msx146. PubMed PMID: 28460059; PubMed Central PMCID: PMC5850483.
5. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *Journal of computational biology : a journal of computational molecular cell biology* (2004) 11(2-3):377-94. doi: 10.1089/1066527041410418. PubMed PMID: 15285897.
6. Reese MG, Eeckman FH, Kulp D, Haussler D. Improved splice site detection in Genie. *Journal of computational biology : a journal of computational molecular cell biology* (1997) 4(3):311-23. Epub 1997/10/01. PubMed PMID: 9278062.
7. Shapiro MB, Senapathy P. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res* (1987) 15(17):7155-74. PubMed PMID: 3658675; PubMed Central PMCID: PMC306199.
8. Tavtigian SV, Deffenbaugh AM, Yin L, Judkins T, Scholl T, Samollow PB, et al. Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. *J Med Genet* (2006) 43(4):295-305. Epub 2005/07/15. doi: 10.1136/jmg.2005.033878. PubMed PMID: 16014699; PubMed Central PMCID: PMC1563222.
9. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods* (2010) 7(4):248-9. Epub 2010/04/01. doi: 10.1038/nmeth0410-248. PubMed PMID: 20354512; PubMed Central PMCID: PMC2855889.

10. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* (2009) 4(7):1073-81. Epub 2009/06/30. doi: 10.1038/nprot.2009.86. PubMed PMID: 19561590.
11. Aldahmesh MA, Li Y, Alhashem A, Anazi S, Alkuraya H, Hashem M, et al. IFT27, encoding a small GTPase component of IFT particles, is mutated in a consanguineous family with Bardet-Biedl syndrome. *Hum Mol Genet* (2014) 23(12):3307-15. Epub 2014/02/04. doi: 10.1093/hmg/ddu044. PubMed PMID: 24488770; PubMed Central PMCID: PMC4047285.
12. Sanchez-Navarro I, L RJdS, Blanco-Kelly F, Zurita O, Sanchez-Bolivar N, Villaverde C, et al. Combining targeted panel-based resequencing and copy-number variation analysis for the diagnosis of inherited syndromic retinopathies and associated ciliopathies. *Sci Rep* (2018) 8(1):5285. Epub 2018/03/29. doi: 10.1038/s41598-018-23520-1. PubMed PMID: 29588463; PubMed Central PMCID: PMC5869593.
13. Quelin C, Loget P, Boutaud L, Elkhartoufi N, Milon J, Odent S, et al. Loss of function IFT27 variants associated with an unclassified lethal fetal ciliopathy with renal agenesis. *Am J Med Genet A* (2018). Epub 2018/04/29. doi: 10.1002/ajmg.a.38685. PubMed PMID: 29704304.

Le projet II

portant sur *BBS1*



## I. Les rétrotransposons.

Parmi les  $3.10^9$  pb du génome humain, 2/3 des séquences correspondent à des séquences d'éléments mobiles insérés (MEI) qui peuvent être de 2 types : les transposons et les rétrotransposons. Les transposons sont des séquences d'ADNg qui sont capables de s'extraire d'un locus pour s'insérer ailleurs dans le génome : on parle d'un processus de « couper-coller » (Medina et Perron 2017). Les rétrotransposons sont des séquences d'ADNg issues d'ADNc insérées dans le génome humain. Celles-ci sont transcrites en ARN, qui subit une rétrotransposition en ADNc, qui est intégré ailleurs dans le génome (Martin 2018). Pour les rétrotransposons, on parle d'un processus de « copier-coller ».

Parmi les rétrotransposons, certains codent pour leur propre machinerie : LINE-1 (Long INterspersted Element 1) par exemple, est considéré comme une classe de rétrotransposons autonomes de grande taille (plusieurs kilobases). En comparaison, les SVA (SINE-VNTR-Alu) et les Alu dépendent de la machinerie LINE-1 pour se propager dans le génome : ce sont des rétrotransposons non autonomes (Hancks et Kazazian 2016). Contrairement aux SVA, les séquences Alu font parties des SINE (Short INterspersted Element) : des rétrotransposons caractérisés par des séquences inférieures à 500pb. Les Alu proviennent de la dimérisation d'ARN 7SL après la délétion de sa séquence centrale. L'ARN 7SL est un ARN non codant qui appartient au complexe « particule de reconnaissance du signal » impliqué dans la synthèse des protéines par le ribosome. LINE-1, SVA et *Alu* sont les 3 classes de rétrotransposons encore actives dans le génome.

LINE-1 est responsable des insertions *de novo* identifiées dans le génome. La séquence LINE-1 est transcrite par l'ARN polymérase II, le transcrit est bi-cistronique et code pour ORF1p (une protéine de liaison au ARNm, RNP) et ORF2p (une rétro-transposase (RT) et une endonucléase (EN)). Cette enzyme reconnaît le motif 5' TTTT-AA 3' (site consensus de reconnaissance) et insère la séquence d'intérêt à cet endroit. Une fois insérée, la séquence est flanquée par un TSD (Target Site Duplication) spécifique de chaque insertion. ORF1p et ORF2p forment un complexe RNP ribonucléoprotéique en se liant à l'ARN LINE-1 (action en *cis*) ou bien aux ARN SVA ou Alu (action en *trans*) (Burns 2020) (Figure 37).

Les rétrotransposons occupent, au total, 40% du génome humain. Au sein de ce groupe d'éléments mobiles, 85% (+/-2%) correspondent à des séquences Alu dont le taux d'insertion atteint 1 pour 20 naissances, 12% (+/- 2%) sont des LINE-1 dont le taux d'insertion s'élève à 1 pour 270 naissances et 2,5% (+/-1%) sont des SVA dont le taux d'insertion est plus bas : 1 pour 916 naissances (C. Stewart et al. 2011; Bragg et al. 2017; Buckley et al. 2017).

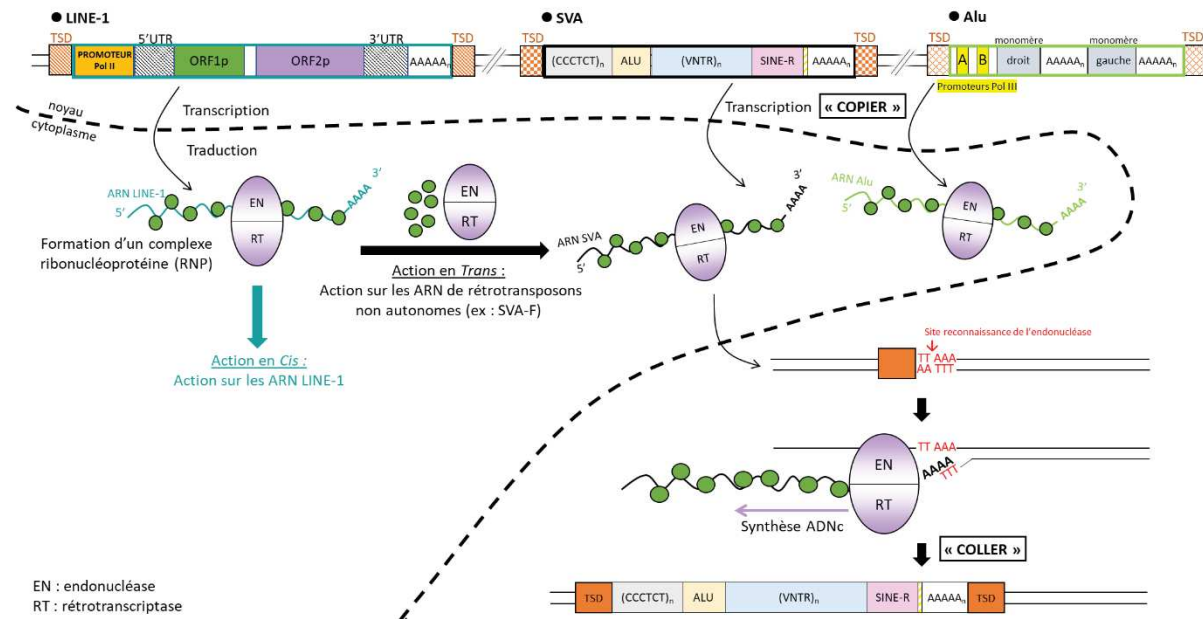
Les insertions se font dans les parties non codantes du génome humain pour la plupart (98,5% du génome) et ont donc peu d'impact sur les fonctions des cellules. Néanmoins, lorsque l'insertion se produit dans/à proximité d'un gène, elle peut avoir un impact sur la fonction du gène (niveau d'expression, interruption de la séquence codante...) et donc être associée à un phénotype. Depuis

1988, on dénombre dans la littérature 124 insertions (LINE 1, *Alu*, SVA confondus) impliquées dans des pathologies humaines. La première décrit l'insertion d'un LINE-1 dans le gène codant pour le facteur de coagulation 8 qui provoque une Hémophilie A (Kazazian et al. 1988a).

Les SVA sont les rétrotransposons les plus jeunes d'un point de vue évolutif, ils sont exclusivement retrouvés dans le génome des hominidés. Il existe 6 familles de SVA répertoriées (notées de A à F) en plus du SVA-F1. Chez l'Homme, on retrouve spécifiquement les types E et F (H. Wang et al. 2005). Globalement, environ 3600 copies de SVA ont été identifiées dans le génome humain de référence. Dans 1/3 des cas, un SVA s'insère au niveau d'une région génique : 16 insertions de SVA ont été associées à des maladies (Bragg et al. 2017; Kherraf et al. 2018; Tavares et al. 2019a; Torene et al. 2020). Selon les cas, l'insertion d'un SVA peut provoquer la délétion d'une partie du gène concerné (Vogt et al. 2014; Takasu et al. 2007), perturber le processus d'épissage (Conley et al. 2005; Hassoun et al. 1994) ou entraîner une exonisation : incorporation d'une partie de la séquence du SVA (non codante) dans la séquence codante du gène d'intérêt (Kobayashi et al. 1998; Wilund et al. 2002; Akman et al. 2010a).

Les SVA-F mesurent en moyenne 1,7 kb et se composent de 5 domaines. Ils sont flanqués à chaque extrémité d'une séquence TSD (<20 pb). Ils sont composés en 5' d'une séquence répétée d'hexamères (CCCTCT)<sub>n</sub>. Une séquence *Alu*-like précède une séquence hautement répétée appelée VNTR (Variable number of tandem repeat) dont la taille varie en fonction du nombre de répétitions des motifs de 40 pb, riches en GC. Enfin, un domaine SINE-R (le R indique l'origine rétrovirale du SINE) est suivi d'une queue polyA (~200pb) (Figure 37). La majorité des SVA insérés présente l'ensemble de 5 domaines mais il est possible qu'ils perdent une partie de leurs extrémités 5' ou 3' lors de l'insertion (Ostertag et al. 2003a). Leur richesse en répétitions présage des potentielles variations au sein d'un même type de SVA. D'ailleurs, il a été montré dans le cas d'un SVA inséré dans le gène *TAF1* que le nombre de répétitions du domaine (CCCTCT)<sub>n</sub> impacte la sévérité du phénotype observé : plus le nombre de répétitions augmente plus les premiers symptômes se déclarent tôt. Cet événement mutationnel est responsable d'une dystonie-parkinson liées à l'X (Burns 2020).

Notre laboratoire a déjà montré que l'insertion d'une séquence *Alu* dans le gène *BBS12* pouvait être responsable d'un BBS (Redin et al. 2012a). Ce projet s'intéresse à la mise en évidence d'un nouvel événement mutationnel fréquent dans le gène *BBS1* avec la mise en évidence de l'insertion d'un SVA de type F.



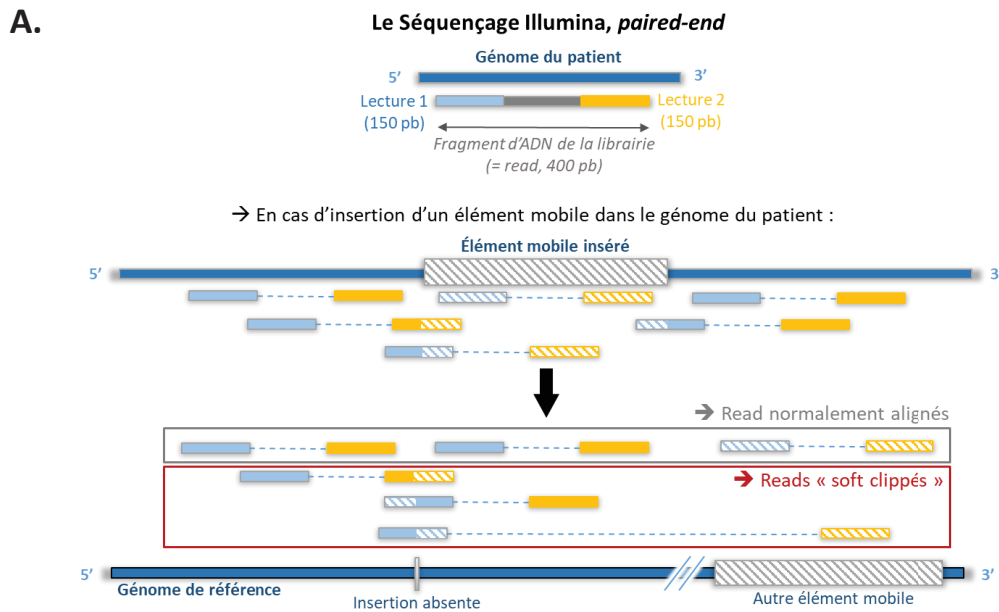
**Figure 37 : La rétrotransposition dépendante de LINE-1.** Représentation du mécanisme de rétrotransposition dépendante de LINE-1 dans le cas d'un SVA.

## II. Le synopsis

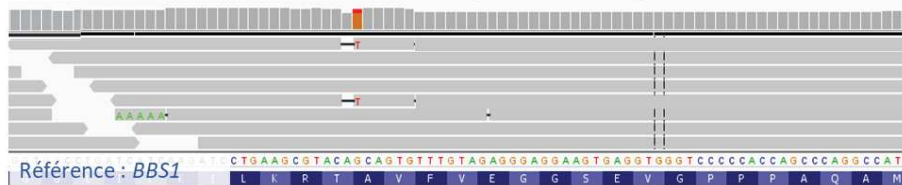
Il s'agit d'un projet débuté au cours de mon stage de M1 au laboratoire et poursuivi durant ma thèse. Le patient index présentant un phénotype typique de BBS est issu d'une famille caucasienne non consanguine. La recherche de mutation a d'abord permis la mise en évidence de la mutation récurrente du gène *BBS1* c.1169T>G, p.M390R à l'état hétérozygote. Au début du projet, il existait dans la cohorte strasbourgeoise, un excès de patients BBS hétérozygotes pour la M390R pour lesquels la seconde mutation était manquante. A l'issue du projet, nous avons résolu l'ensemble de ces cas. Pour une famille, la seconde mutation correspond à une large délétion des exons 4 à 11 de *BBS1*. Pour les autres patients, la seconde mutation correspond à l'insertion d'un rétrotransposon de type SVA-F au niveau de l'exon 13 du gène *BBS1*.

L'identification de cette mutation a été complexe. Nous n'avons pas appliqué d'outils bioinformatiques dédiés à la détection des MEI lors de la première analyse par séquençage haut débit car ils étaient peu développés à l'époque. Seul un séquençage à très haute profondeur (1000x) a permis de détecter ce que l'on croyait être une délétion/insertion de quelques bases alors qu'il s'agissait finalement d'une partie du SVA F mal alignée. Comme l'explique la Figure 38, lors d'un séquençage Illumina en paired-end, la distance séparant la lecture 1 en 5' et la lecture 2 en 3' correspond normalement à 400 pb : l'ensemble forme un read. Dans le cas où la région séquencée ne contient pas la séquence du rétrotransposon : les reads générés sont correctement alignés sur le génome de référence. Dans le cas où le read généré à une séquence correspondant dans sa totalité à la séquence de l'insertion alors il est aligné sur de nombreuses séquences de rétrotransposons

existants dans le génome de référence. Par contre, dans le cas où le read contient une partie de la séquence de référence et une partie de la séquence du rétrotransposon, lors de l'alignement, une partie du read est alignée correctement alors que la partie correspondant au rétrotransposon est mal alignée ou alignée ailleurs. La partie mal alignée est alors étiquetée comme de mauvaise qualité et le read complet est « soft clippé » car considéré comme non conforme à la référence (ces reads sont cachés par défaut dans le logiciel de visualisation). C'est en faisant apparaître les reads « soft clippés »

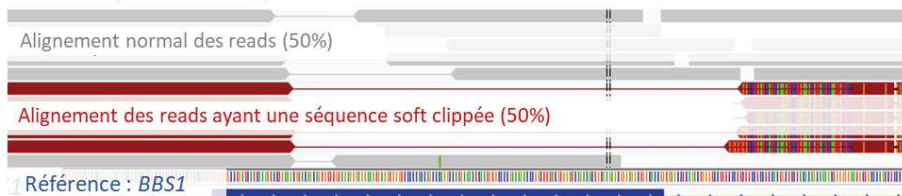


**B.** → 1<sup>ère</sup> visualisation des données issues du TES du patient avec le logiciel IGV :



→ 2<sup>ème</sup> visualisation des données issues du TES du patient avec le logiciel IGV :

✓ Les séquences soft clippées sont visibles

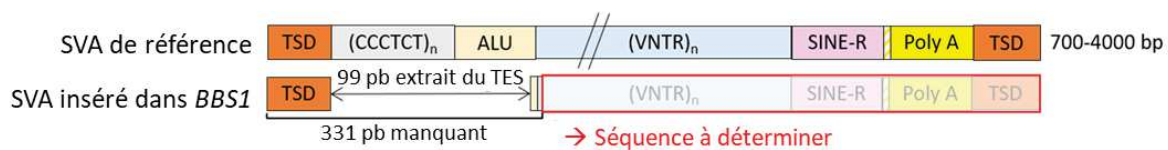


du cas index dans le logiciel IGV que l'insertion a été mise en évidence.

**Figure 38 : Les « soft clips ».** A. Explication de la génération des read « soft clippés » dans le cas où le patient est porteur d'une insertion dans son génome. B. Visualisations dans IGV (Integrative Genomic Viewer) des reads de patient sans/avec ceux « soft clippés ».

L'objectif principal de ce projet consistait à déterminer la taille, la séquence et l'impact fonctionnel de cette insertion. De plus, en raison des difficultés rencontrées lors de la détection, il était possible que d'autres patients soient porteurs de cette insertion. Il fallait donc évaluer sa fréquence. En fonction de celle-ci et du contexte génomique, il s'agissait également de déterminer si cet évènement pouvait se reproduire ou si cela restait un évènement isolé.

Une séquence de 99 pb avait été extraite des reads « soft clippés » et un alignement en utilisant le logiciel « Repeat masker » avait permis d'identifier un SVA (SINE R-VNTR-Alu) ayant une séquence 5' tronquée (Figure 39).

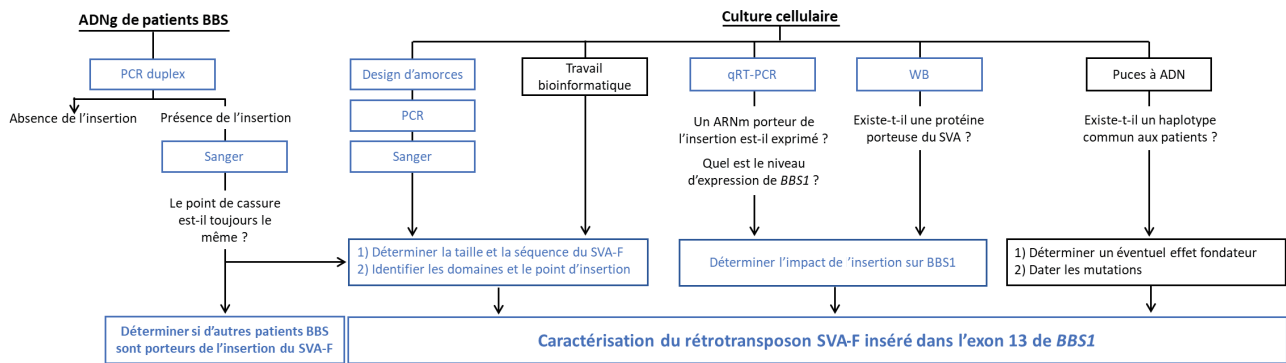


**Figure 39 : Composition des SVA-F de référence.** Les SVA-F sont normalement composés de 6 domaines. Pour le SVA-F inséré dans *BBS1*, l'extrémité 5' est manquante et la majeure partie de l'extrémité 3' doit être déterminée. TSD : Target Site Duplication, ALU : Alu like domain, VNTR : Variable number tandem repeat, SINE-R : Short interspersed nucleotide Element

Le logigramme d'expérience de la Figure 40 montre la démarche scientifique que nous avons mise en place pour répondre aux problématiques. J'ai toujours travaillé sur le matériel biologique extrait des cellules en culture du cas index.

Parallèlement à mon travail de caractérisation moléculaire, un travail bioinformatique de reconstitution de séquences à partir des lectures extraites du TES/WGS a été mis en œuvre par un stagiaire M2, Samuel Nicaise. Il est parvenu à étendre la séquence initiale de 99 à 350 pb.

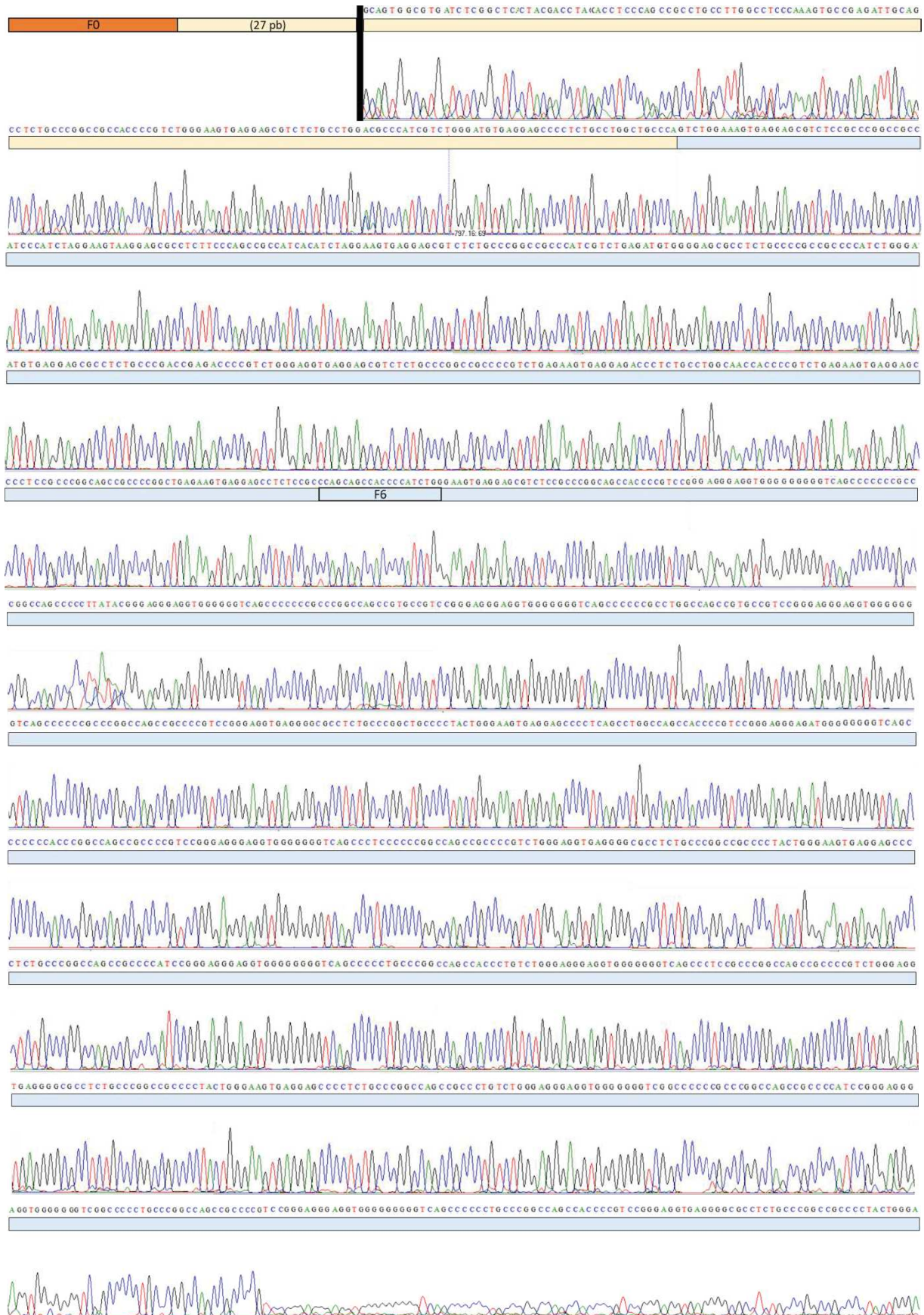
A partir de ces séquences et de la séquence consensus du SVA de Kwon et al. 2013, j'ai « marché » sur le génome du patient et déterminé la séquence complète du rétrotransposon (2435 pb) qui est disponible en Figure 41. La difficulté de cette région (à savoir la richesse en GC, la présence de nombreuses répétitions et la queue polyA) a demandé de nombreuses optimisations des protocoles PCR. J'ai identifié 4 des 5 domaines connus ainsi que le site consensus de reconnaissance de la transposase LINE-1. J'ai confirmé l'impact fonctionnel de cette insertion sur l'expression du gène par qRT-PCR et WB. Un ARNm porteur de l'insertion est exprimé chez le patient. Le niveau d'expression de *BBS1* est réduit de moitié chez le patient : la quasi-totalité des ARNm sont porteurs de la M390R. Au niveau protéique, j'ai montré qu'il n'existait pas de protéine plus courte porteuse de l'insertion et j'ai confirmé que la quantité protéique de *BBS1* du patient est réduite de moitié par rapport aux contrôles.

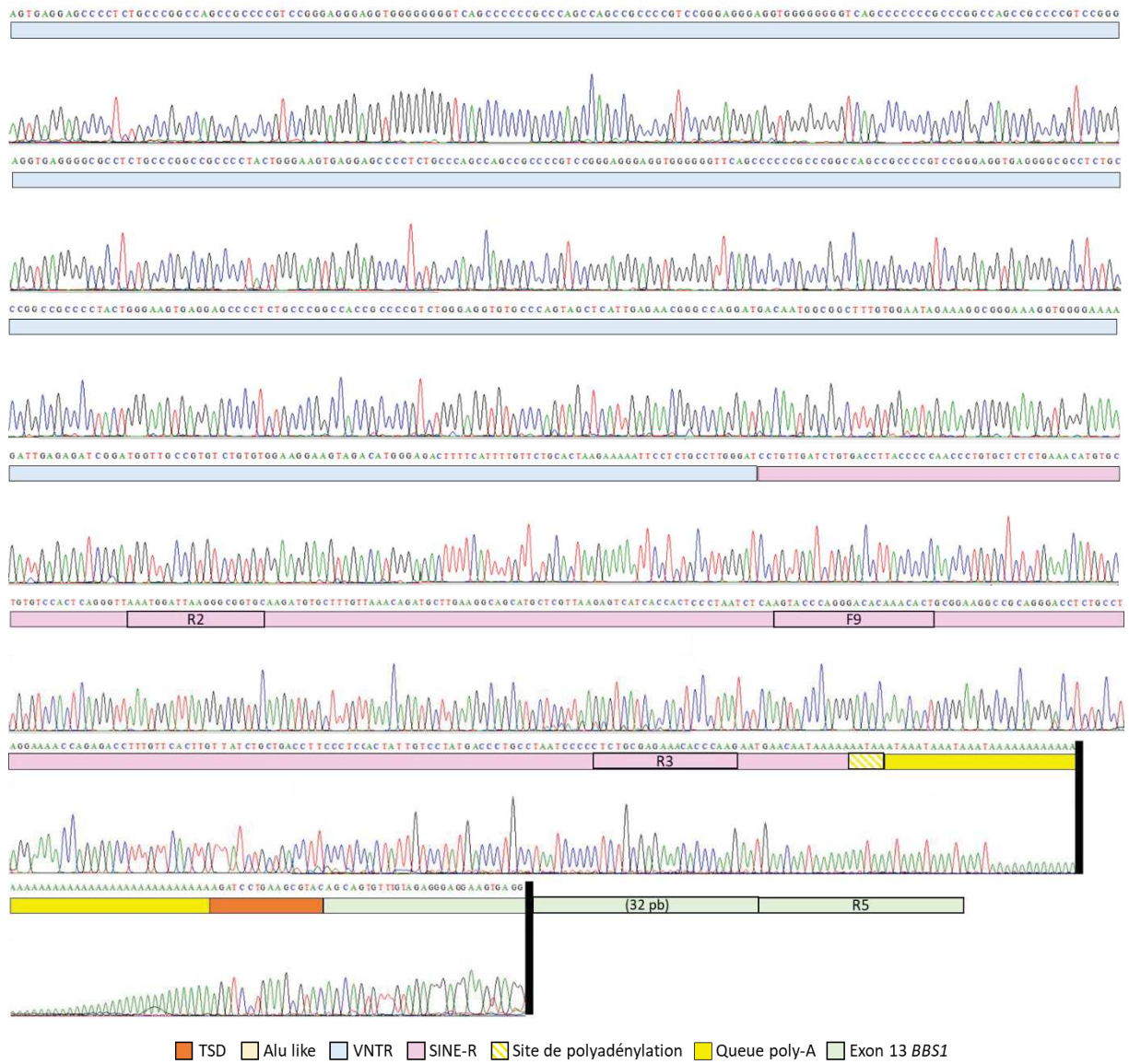


**Figure 40 : Logigramme.** Résumé des expériences mises en œuvre pour répondre aux questions posées et aux objectifs du projet. Mon travail est identifié en **bleu**.



### III. Les données supplémentaires





**Figure 41 :** La séquence Sanger totale du SVA-F.



## IV. La conclusion et les perspectives.

Ce projet de caractérisation moléculaire nous a permis de conclure plus précisément sur la nature du SVA-F inséré dans *BBS1* et sur son impact fonctionnel. La mise au point d'une PCR duplex pour détecter rapidement cet évènement nous a permis d'explorer la cohorte de patients négatifs de Strasbourg. Nous avons finalement résolu le cas de 6 familles au total. Les patients présentent le même génotype avec exactement le même point d'insertion.

Pour le cas index, nous avons déterminé la taille totale de 2435 pb de l'insertion et identifié 4 des 5 domaines canoniques composant ce type de rétrotransposons. Nous avons montré qu'un ARNm porteur de l'insertion est exprimé mais aucune protéine porteuse de l'insertion n'a été mise en évidence par WB. Le NMD (*Nonsense mediated mRNA decay*) est un mécanisme cellulaire qui empêche la traduction de protéines tronquées en captant et dégradant les ARNm porteurs de codons stops prématurés. L'insertion du SVA-F cause l'apparition d'un codon stop prématuré, la protéine résultante serait tronquée de son extrémité C-terminale. Dans ce cas, il serait intéressant de bloquer ce système de dégradation en utilisant un inhibiteur comme l'émétine par exemple. Une fois le NMD bloqué, il faudrait refaire le WB afin de déterminer si une protéine tronquée porteuse de l'insertion est synthétisée chez le patient, dans cet état. Une hypothèse post transcriptionnelle impliquant une dégradation via le protéasome de la protéine ne peut pas non plus être écartée.

Ce projet a également été propice pour la mise en place de collaborations avec d'autres équipes de recherche. Par exemple, la collaboration établie avec Erica Davis a révélé qu'il existe, dans leur cohorte, 2 familles de patients BBS résolus (porteurs de 2 allèles mutés dans un premier gène *BBS*) qui sont également porteurs d'un 3<sup>ème</sup> allèle correspondant à l'insertion du SVA-F dans *BBS1*. Badano et al. 2003 a montré l'implication de *BBS1* dans l'hérédité tri-allélique en rapportant une mutation sur un seul allèle associée avec 2 mutations dans *BBS2* ou *BBS7*. Dans ce cas, lorsque l'apparition du phénotype est associée à des mutations dans plus de 2 gènes, on parle d'un mode de transmission oligogénique. Ce 3<sup>ème</sup> allèle exercerait un effet modificateur qui permettrait d'expliquer les variabilités inter et intra-familles mises en évidence lors des évaluations phénotypiques. Ce 3<sup>ème</sup> allèle modifierait la sévérité ou l'âge d'apparition des premiers signes cliniques. Dans le cadre de ce projet, nous voudrions confirmer et comprendre précisément l'effet d'un tel allèle. L'identification de familles supplémentaires est nécessaire. Il serait donc intéressant d'explorer la cohorte de BBS résolus de Strasbourg pour identifier d'autres patients présentant le même génotype. Ensuite, il serait intéressant de comparer le phénotype des patients résolus porteurs de ce 3<sup>ème</sup> allèle à celui des patients n'ayant pas cet allèle pour déterminer si effectivement il module le phénotype.

Le test par PCR duplex présenté dans le manuscrit est maintenant appliqué systématiquement au laboratoire de diagnostic génétique (HUS) face à de nouveaux cas BBS irrésolus. Un protocole

bioinformatique de détection des insertions d'éléments répétés au sein du génome, basé sur le logiciel Mobster, a également été implémenté. Cet ajout s'inscrit dans la continuité de 2 études récentes ayant rapporté plusieurs identifications d'insertion d'éléments mobiles (MEI) dans les exomes de patients (Gardner et al. 2019; Torene et al. 2020). Un bon de 0,15% du taux de diagnostic dans le cas de la 2<sup>ème</sup> étude citée a été observé et confirme la nécessité de mettre en place des systèmes de détection des MEI dans le cadre du diagnostic génétique.

Parmi les 6 familles, 5 sont originaires de Bretagne et une seule est issue des pays scandinaves. En prenant en considération l'origine commune des familles, nous avons émis l'hypothèse que l'apparition de cette mutation soit liée à un effet fondateur. Les résultats des puces à ADN (SNP arrays) ont confirmé cette hypothèse. Nous avons défini 2 haplotypes et sommes parvenus à dater l'apparition de la mutation M390R et de l'insertion SVA-F en collaborant l'équipe d'Emmanuelle Genin, de Brest. Le premier haplotype est de grande taille (>1,6Mb). Il est commun aux patients porteurs de la mutation M390R, à l'état homozygote. Cette mutation serait apparue entre il y a 2325 ans [1225;4575]. Le deuxième haplotype est de plus petite taille (>56kb). Il correspond au plus petit haplotype commun entre les patients porteurs de l'insertion. Cette mutation serait apparue il y a 1850 ans [825;4750]. L'insertion du SVA-F est apparue plus récemment que la mutation M390R.

Ces résultats sont récapitulés dans un article qui a été soumis dans le journal *Human genetics* le 20/07/20 et est actuellement en révision.

## V. Le manuscrit

### An SVA F retrotransposon insertion is a frequent cause of *BBS1* mutation.

Delvallée C.<sup>1</sup>, Nicaise S.<sup>1</sup>, Antin M.<sup>2</sup>, Leuvrey AS.<sup>2</sup>, Nourisson E.<sup>2</sup>, Leitch C. C.<sup>3</sup>, Kellaris G.<sup>3</sup>, Stoetzel C.<sup>1</sup>, Geoffroy V.<sup>1</sup>, Scheidecker S.<sup>1,2</sup>, Keren B.<sup>4,5</sup>, Depienne C.<sup>4,6</sup>, Klar J.<sup>7</sup>, Dahl N.<sup>7</sup>, Deleuze JF<sup>8</sup>, Génin E.<sup>9</sup>, Redon R.<sup>10</sup>, Demurger F.<sup>11</sup>, Devriendt K.<sup>12</sup>, Mathieu-Dramard M.<sup>13</sup>, Poitou-Bernert C.<sup>14</sup>, Odent S.<sup>15,16</sup>, Katsanis N.<sup>3,17</sup>, Mandel JL.<sup>2,18</sup>, Davis E. E.<sup>3,17</sup>, Dollfus H.<sup>1,19,20</sup>, Muller J.<sup>1,2</sup>.

<sup>1</sup>Laboratoire de Génétique Médicale, Institut de génétique médicale d'Alsace IGMA, INSERM U1112, Fédération de Médecine Translationnelle de Strasbourg (FMTS), Université de Strasbourg UMRS\_1112, Strasbourg, France.

<sup>2</sup>Laboratoires de Diagnostic Génétique, Hôpitaux Universitaires de Strasbourg, Strasbourg, France.

<sup>3</sup>Advanced Center for Translational and Genetic Medicine (ACT-GeM), Stanley Manne Children's Research Institute, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL, USA

<sup>4</sup>Institut du Cerveau et de la Moelle épinière (ICM), Sorbonne Université, UMR S 1127, Inserm U1127, CNRS UMR 7225, F-75013 Paris, France.

<sup>5</sup>AP-HP, Hôpital de la Pitié-Salpêtrière, Département de Génétique, F-75013, Paris, France.

<sup>6</sup>Institute of Human Genetics, University Hospital Essen, University of Duisburg-Essen, Hufelandstraße 55, 45147 Essen, Germany

<sup>7</sup>Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, Uppsala, Sweden.

<sup>8</sup>Centre National de Recherche en Génomique Humaine (CNRGH), Institut de biologie François Jacob, CEA, 91000, Evry, France.

<sup>9</sup>Inserm UMR1078, CHRU Brest, Univ Brest, Brest, France

<sup>10</sup>Université de Nantes, CNRS, INSERM, l'institut du thorax, F-44000 Nantes, France.

<sup>11</sup>Service de Génétique Médicale, Centre Hospitalier Bretagne Atlantique, Vannes, France.

<sup>12</sup>Center for Human Genetics, University Hospital Leuven and KU Leuven, Leuven, Belgium.

<sup>13</sup>Centre d'activité de génétique clinique, CLAD nord de France, CHU Amiens, Amiens, France.

<sup>14</sup>Assistance Publique Hôpitaux de Paris, Nutrition Department Pitié-Salpêtrière Hospital; Sorbonne Université, INSERM, NutriOmics Research Unit, Paris, France;

<sup>15</sup>Service de Génétique Clinique, CHU Rennes, Rennes, France.

<sup>16</sup>CNRS, IGDR (Institut de Génétique et Développement de Rennes) UMR 6290, Université de Rennes, Rennes, France.

<sup>17</sup>Department of Pediatrics, Feinberg School of Medicine, Northwestern University, Chicago, IL USA

<sup>18</sup>Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS UMR 7104, INSERM U964, Université de Strasbourg, Dept Transl Med and Neurogenetics Illkirch, France

<sup>19</sup>Service de Génétique Médicale, Hôpitaux Universitaires de Strasbourg, Strasbourg, France.

<sup>20</sup>Filière SENSGENE, Centre de Référence pour les affections rares en génétique ophtalmologique, CARGO, Hôpitaux Universitaires de Strasbourg, Strasbourg, France.

Corresponding Author: Jean Muller

Email address: [jeanmuller@unistra.fr](mailto:jeanmuller@unistra.fr)

ORCID: 0000-0002-7682-559X

**ABSTRACT**

Bardet-Biedl syndrome (BBS) is an emblematic ciliopathy characterized by retinitis pigmentosa, obesity, polydactyly, cognitive impairment and renal failure. Mutations in 24 genes account for the molecular basis of >80% of cases. Toward saturated discovery of the mutational basis of the disorder, we applied targeted exome sequencing and mobile element insertion detection to 217 BBS samples. We identified a recurrent hominid-specific SINE-R/VNTR/*Alu* type F (SVA F) insertion in exon 13 of *BBS1* in 3 families. Coincidentally, these three pedigrees all carry the common c.1169T>G, p.M390R variant *in trans*, unmasking *BBS1* as their recessive locus. Whole genome sequencing, *de novo* assembly and SNP array analysis were performed to characterize the genomic event. This insertion is extremely rare in the general population (>1,300 samples screened and >10,800 individuals from gnomAD-SV) and due to a founder effect. Its 2,435 bp sequence contains hallmarks of LINE1 mediated retrotransposition. Functional studies with patient-derived cell lines confirmed that the *BBS1* SVA-F is deleterious as evidenced by a significant depletion of both mRNA and protein levels. Next, we performed PCR-based screening of two additional BBS cohorts: (1) 192 cases pre-selected to not have the common *BBS1* or *BBS10* mutations in homozygosity; or (2) 126 cases with no pre-selection bias. Together, we describe 8 new families with the SVA F insertion in *BBS1*, some of which are primary causal alleles (n=6), and some are oligogenic sites in addition to other recessive BBS loci (n=2). Such findings highlight the importance of dedicated bioinformatics pipelines to identify all types of variation.

**Keywords:** Bardet-Biedl syndrome, BBS1, Mobile element insertion, SVA F, Founder effect.

## INTRODUCTION

Bardet-Biedl syndrome (BBS, MIM# 209900) is an increasingly well-characterized ciliopathy with clinical diagnostic features including retinitis pigmentosa, postaxial polydactyly, obesity, cognitive impairment and renal dysfunction. BBS is a rare autosomal recessive condition affecting 1/160,000 live births in populations of northern European origin (Waters et Beales 2011). The disease prevalence is higher (1/13,500 to 1/43,500) in consanguineous populations or in isolated geographic regions (Farag et Teebi 1989; Gouronc et al. 2020). Recent studies have delineated the antenatal and postnatal presentation of BBS patients (Mary et al. 2019; Niederlova et al. 2019) with mutations in 24 genes accounting for the molecular basis of a majority of cases (Shamseldin et al. 2020a). Diagnostic yield in BBS can be high, ranging from 65% to 100% depending on the cohort, the inclusion criteria and the molecular test used (Mary et al. 2019; Shamseldin et al. 2020a). These diagnostic approaches include direct sequencing of the predominant *BBS1* and *BBS10* founder variants (Mykytyn et al. 2002b; Stoetzel et al. 2006, 10), large panel sequencing or whole exome (WES) and genome (WGS) sequencing (Redin et al. 2012b; Shaheen et al. 2016; Shamseldin et al. 2020a). Most of the causal genes reported since *BBS13/MKS1* have been described in few cases or even a single family (Yıldız Bölükbaşı et al. 2018; Wormser et al. 2019), suggesting that the causal loci yet to be discovered will be rare in BBS populations, and/or that pathogenic variation may be intractable to standard sequencing paradigms.

Among the genetic events arising in the human genome, mobile element insertions (MEI), although rarely described in the literature, occur ~1,000 times per genome according to the 1,000 Genomes project (Auton et al. 2015). These mutations may implicate diverse classes of elements (SINE, LINE, retrovirus and processed pseudogenes), with a small fraction of such elements that are still active (0.05%) in the human genome (Mills et al. 2007). They can move using several mechanisms and integrate at a different locus than initially identified (Ha, Loh, et Xing 2016). Integration can result in no functional consequences (Chatron et al. 2019) but several MEI have been reported to cause pathology (van der Klift et al. 2012; Nakamura et al. 2015; Taniguchi-Ikeda et al. 2011; Akman et al. 2010b; Kazazian et al. 1988b; J.-M. Chen et al. 2008; Miki et al. 1996; Redin et al. 2012b). With only ~2,700 copies found in the human genome (Hancks et Kazazian 2012), SINE-R/VNTR/*Alu* (SVA) is the youngest family of retrotransposons. Composed of 6 types (A to F), SVA E and F are specific to the human lineage (H. Wang et al. 2005). SVA are non-autonomous retrotransposons requiring the L1 machinery to move and integrate into another location (Raiz et al. 2012a). To date, 16 SVA insertions have been associated with human diseases (Hancks et Kazazian 2016; Brett et al. 2017, 1; Kherraf et al. 2018, 66; Tavares et al. 2019b; Torene et al. 2020).

Among the known BBS genes, *BBS1* contributes the highest fraction of pathogenic alleles to BBS patients with pathogenic variants in ~20% of all cases (Khan et al. 2016). In particular, this observation is largely due to a pathogenic founder variant (c.1169T>G, p.M390R) (Mykytyn et al. 2003). Among the unresolved cases in our cohort, several patients were heterozygous carriers of this specific missense change suggesting two different possibilities: (1) a missed second primary driver allele in *BBS1*; or (2) participation of this variant as a second-site modifier in the presence of a different primary BBS driver locus (Philip L. Beales et al. 2003; Stoetzel et al. 2006). Notably, due to the high frequency of this variant in the general population (0.27% in non-Finnish Europeans; gnomAD) this variant could be detected coincidentally.

In this study, we report the identification of a MEI of SVA F type in 8 families, a portion of which are *in trans* to the recurrent c.1169T>G (p.Met390Arg) pathogenic variant in *BBS1* (n=6), and a subset as secondary sites in addition to a known recessive BBS locus (n=2). Already described once in 2019 (Tavares et al. 2019b), here we estimate the frequency of this variant in BBS (131 families) identifying this variant as the second most common pathogenic variant in *BBS1*. We also determined its full-length sequence using WGS and performed functional analysis to understand its consequences. In our cohorts, the insertion always occurs at the same insertion location and its recurrence in BBS populations is the result of a founder effect.

## MATERIALS AND METHODS

### Subjects

Written informed consent was obtained from each participant and/or parent or legal guardian. The study protocols were approved by local Institutional Review Boards (IRB): “Comité de Protection des Personnes” (EST IV, N°DC-20142222), or the Lurie Children’s Hospital IRB (IRB 2019-3057 or IRB 2019-2950). Our research complies with the Declaration of Helsinki. All patients were diagnosed with Bardet-Biedl syndrome according to Beales’ clinical diagnostic criteria (P. L. Beales et al. 1999b) (Supplementary Table 1). Genomic DNA was extracted from peripheral blood leukocytes using QIAamp DNA mini kit (QIAGEN, Hilden, Germany) according to the manufacturer’s instructions. Two cohorts (cohort 1: n=217 unrelated cases; and cohort 2; 213 cases from 192 families) were pre-selected to exclude individuals who harbor homozygous recurrent pathogenic mutations in *BBS1* and *BBS10* (c.1169T>G and c.271dupT, respectively). Cohort 1 has also been explored using Targeted Exome Sequencing (see below). A third cohort (cohort 3; 134 cases from 126 families) excluded families with recessive *BBS1* mutations, but was otherwise not subjected to pre-selection based on mutational status. The last cohort is composed of *BBS1* positive cases (e.g. carrier of two pathogenic variants) of European ancestry (cohort 4: n=131 unrelated cases).

### Targeted Exome Sequencing

Following the French diagnostic recommendations (<https://anpgm.fr>) and the recent updates (Muller et al. 2010b; Mary et al. 2019; Gouronc et al. 2020), targeted exome sequencing (TES) was performed after negative screening of the recurrent pathogenic variants in *BBS1* and *BBS10* (cohort 1). TES of genes related to BBS included either 31 genes as described previously (Redin et al. 2012b) or an updated version of 58 genes (Supplementary Table 2). Exon regions of DNA samples were captured with in-solution enrichment methodology (Agilent QXT SureSelect custom panel) and sequenced with an Illumina NextSeq 550 instrument (paired-end sequencing, 2x150 bases, 48 libraries per lane). Following alignment to the reference human genome (GRCh37/hg19) using BWA v0.7.5a (H. Li et Durbin 2009), SNVs and indels were called with the Genome Analysis Toolkit v.3.4.46 (DePristo et al. 2011) using our in house pipeline (STARK) and following GATK best practices. Annotation and ranking was performed by VaRank (Geoffroy et al. 2015) configured using Alamut Batch standalone version (Interactive Biosoftware, Rouen). Structural variations were called using either CANOES (Backenroth et al. 2014) or Mobster (Thung et al. 2014) and annotated using AnnotSV 1.0 (Geoffroy, Herenger, et al. 2018). Dedicated SVA read analysis was performed using RepeatMasker (Smit, Hubley, et Green 2013) and BLAT (Kent 2002).

### Whole genome sequencing

Sequencing libraries were prepared from 1 µg of DNA using the TruSeq PCR-free DNA sample preparation kit (cat#FC-121-3001/3002, Illumina) targeting an insert size of 350 bp. The library was performed according to the manufacturer’s instructions (guide#15036187, revC). Sequencing was performed on a HiSeqX using V2 sequencing chemistry on 3 lanes. Bioinformatics analysis was performed according to the piper 1.2.0 framework ([www.github.com/NationalGenomicsInfrastructure/piper](http://www.github.com/NationalGenomicsInfrastructure/piper)). A mean coverage of 28.44X was obtained from the mapping of the 608,447,810 reads (paired-end sequencing, 2x150 bases) onto the reference human genome (GRCh37/hg19) using BWA v0.7.5a. The raw alignments were de duplicated, recalibrated and cleaned using GATK 3.3.0 and samtools 0.1.19 (H. Li 2011). Quality control information was obtained using Qualimap 2.0 (Okonechnikov, Conesa, et García-Alcalde 2015).

### WGS using Linked-reads technology (Chromium)

Libraries were prepared using the Chromium Gel Bead and Library Kit (10X Genomics, USA) and the Chromium instrument (10X Genomics), according to the manufacturer's instructions. For both strategies, libraries were sequenced on the Illumina HiSeq X5 system. Sequencing data were analyzed with LongRanger (v2.1.4) using default parameters. The algorithm performed alignment (GRCh37/hg19), haplotyping, and SV detection. The Loupe program (v2.1.1) from 10X Genomics was used to visualize the results as well as IGV (Thorvaldsdottir, Robinson, et Mesirov 2013). A mean depth of coverage of 69.6X was obtained from the mapping of the 1,616,790,902 reads (paired-end sequencing, 2x150 bases) onto the reference human genome (GRCh37/hg19). This represented 95.95% of mapped reads. 68.7% of the analyzed molecules were of length >20kb.

### Characterizing the SVA F sequence using WGS

Split reads carrying the SVA sequence were extracted from the alignment using their position and Concise Idiosyncratic Gapped Alignment Report (CIGAR) field. All reads from discordant pairs in a +/- 400 bp (insert size) window around the SVA insertion were queried. All mate reads were extracted from the alignment for both split

reads and reads from discordant pairs. Reads from 5' and 3' ends of the insertion were assembled separately using Velvet 1.2.10 (Zerbino et Birney 2008) with the maximum possible k-mer size (31), automatic expected coverage settings and minimal values of coverage cutoff (-cov\_cutoff 2 with WGS data).

#### Genomic DNA amplification and sequencing

Amplification was performed with a Mastercycler ep. Gradient S (Eppendorf, Germany) on SVA F carrier gDNA with several primers distributed inside the insertion to reconstitute it by overlapping amplicons. The SVA F reference sequence of Kwon *et al* (2013) was used for primer design. DMSO and Betaïne were added to the PCR mix to optimize amplification for this difficult fragment (repeated and GC rich sequence), and a touchdown cycling protocol was applied. The duplex PCR was performed using 3 primers (1F, 1R, 2R) with normal PCR conditions. Sequencing of PCR products was performed by the GATC Sequencing Facilities (Konstanz, Germany) using BigDye Terminator sequencing chemistry. Primers used are summarized in Supplementary Table 3.

#### qPCR

Absolute quantification was performed using the SyberGreen Mastermix Quantitect (Qiagen, Hilden, Germany) measured on the LightCycler 480 (Roche, Madison, WI, USA). Amplicons were designed in exons 8 and 9 of *BBS1* and compared to two reference genes (*RPPH1* and *GAPDH*). Each sample was analyzed in duplicate and standard deviation between technical replicates was <0.2. Standard curves were established using five different DNA quantities (50, 25, 10, 5 and 1 ng) and PCR efficiency was assessed for each amplicon ( $1.9 < \text{PCR efficiency} < 2.0$ ). The ratio between the mean concentrations from the amplicon of interest compared with reference genes was calculated. The mean values for both control genes were calculated for each patient and plotted. Primers used are summarized in Supplementary Table 3.

#### Cell culture

Primary dermal fibroblasts from a BBS case (A.II-3) and control individuals were obtained by skin biopsy as described (Scheidecker *et al.* 2014). The skin fibroblasts from A.II-3 and sex-matched healthy control were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum (FCS) and 1% penicillin–streptomycin–glutamine (PSG). To induce ciliation, cells were serum deprived for 24 hours in DMEM with 1% PSG but only 0.1% FCS.

#### RNA analysis and RT-qPCR

Total RNA was extracted from fibroblast cultures using a QIAshredder Kit (Qiagen Ref 79654) and RNeasy Kit (Qiagen Ref 74104). A DNase treatment was performed with TURBO DNA-free™. RNA integrity was confirmed by gel electrophoresis and RNA concentration was determined using the spectrophotometer/fluorometer DenoVIX<sup>R</sup> (DS-11 FX). Reverse transcription was performed on 1 µg of RNA using the BioRad iScript™ cDNA Synthesis Kit (BioRad, Hercules, CA). qPCR was performed with a BioRad CFX96™ Real-Time System with the iQ™ SYBR® Green Supermix. Primers used are summarized in Supplementary Table 3. The fold expression of the target gene was calculated using the relative quantification method by normalizing to both *GAPDH* and *HPRT* reference gene expression, using the CFX Manager Software Version 1.5 and calculation in Microsoft Excel.

#### Protein extraction and Western Blot

Total proteins were extracted from cell pellets using a RIPA protocol and quantification was performed by Bradford method with an Eppendorf BioPhotometer Plus spectrophotometer. 40 ng of protein per well was migrated on a 10% acrylamide gel (Mini-PROTEAN<sup>R</sup> TGX™ Gels cat.#456-1034, Biorad). Blotting transfer to a PVDF membrane (Trans-Blot<sup>R</sup> Turbo™ Transfer pack, cat.#170-4156, Biorad) was performed using the Transblot<sup>R</sup> Turbo™ Transfer System (Biorad). Blocking was conducted with TBST-milk before overnight incubation at 4°C with mouse polyclonal BBS1 primary antibody (1:500 dilution, ref: H00000582-A01, Abnova) or mouse monoclonal β-tubulin primary antibody (1:2500 dilution, ref: TUB-2A2, Euromedex). Secondary antibodies were horse anti-mouse HRP linked (1:2000 dilution, ref: 7076, Cell Signaling) and donkey anti-rabbit HRP linked (1:2000, ref: 7074, Cell signaling). Chemiluminescent detection was performed using a Chemidoc™ MP Imaging System (Biorad) with SuperSignal<sup>R</sup> West Femto Maximum Sensitivity Substrate (prod #34095, ThermoFisher) or Pierce™ ECL Western (ref 32209, ThermoFisher). Stripping was done using the Restore™ Western Blot Stripping Buffer™ (prod #21059). The β-tubulin immunoblotting was used as reference and normalization was performed using Image Lab software 6.0.1 (Biorad).



**SNP arrays**

Nine samples, including families A to D (A.I-1, A.II-3, B.II-1, C.II-3, D.I-1 and D.II-2) and 3 controls (homozygous for the c.1169T>G, p.Met390Arg) were processed on Illumina OmniExpress-24 v1 arrays using the Infinium assay as described (Stemers et al. 2006), and results were analyzed using Illumina GenomeStudio software (Supplementary Table 4). Reference standards were set with the Illumina Genetrain 2.0 algorithm using the 96 samples processed during the same run. Regions of homozygosity were called with the Illumina cnvPartition 3.1.6 algorithm, with a minimal region size of 1 Mb.

**Principal Component Analysis (PCA) of SNP array genotype data**

The four SVA F carriers were compared to European non-Finnish individuals from the 1000 Genomes project (Auton et al. 2015) and French controls from the FranceGenRef panel. For that purpose, we selected SNPs that were present in all three panels, merged the three datasets and performed PCA analysis on the observed genotypes at 118,577 SNPs (these were the SNPs remaining after LD-pruning using plink1.9 --indep-pairwise 50 5 0.2) (Chang et al. 2015).

**Haplotype reconstruction**

Haplotypes of the 4 SVA-F carriers, their parents and three *BBS1*-M390R homozygous carriers were reconstructed for the entire chromosome 11. For that purpose, the software Hapi-UR (Williams et al. 2012) was used with a window size of 100. Five runs were done and the consensus was obtained by using the best vote rule as suggested by the authors and implemented in the package.

**Estimation of the age of the Most Recent Common Ancestor (MRCA) of SVA F carriers**

The age of the MRCA of the SVA-F carriers was estimated by using the maximum likelihood method implemented in Estiage (Genin et al. 2004). Haplotypes shared by the four SVA F carriers were compared to find the position where recombination events are likely to have happened. Following Mathieson I & McVean G (Mathieson et McVean 2014), we scanned on the left and right side along the genome of each individual on each side of the mutation, until we reach a point where individuals have inconsistent homozygous genotypes. This method gives an estimate of the distance to the first recombination breaking the ancestral haplotype.

***BBS1* genomic context**

The genes and their corresponding sequences were extracted from the UCSC Table browser (Group: "Gene and gene predictions", Track: "NCBI RefSeq", Table: "UCSC RefSeq (refGene)") (Lee et al. 2019). The longest coding RefSeq transcript for each gene (O'Leary et al. 2015) was selected defining its genomic region on the human genome (GRCh37). Based on these coordinates, GC content was computed for each gene. Similarly, the mobile elements map has been downloaded from the UCSC Table browser (Group: "Repeats", Track: "RepeatMasker" (Smit, Hubley, et Green 2013)) and an overlap with a minimum of 1 bp was used to define the landscape of *Alu* and L1 insertions. Based on this, a density of L1 and *Alu* elements is defined for each gene.



## RESULTS

### **Identification of an SVA F insertion in BBS1**

According to our tiered diagnostic paradigm, we performed initial screening of recurrent BBS gene variants (*BBS1* p.M390R and *BBS10* p.C91Lfs\*5) in a first cohort of BBS cases with mixed ethnic origin. 217 individuals did not harbor homozygous variants at either mutation site (cohort 1) and advanced for targeted exome sequencing (TES). We identified seven individuals who carry heterozygous p.M390R changes, all of which were confirmed by TES and were without a second pathogenic *BBS1* variant identified *in trans*. We and others have shown previously that copy number variants can unmask a missing recessive allele in *trans* with the recurrent p.M390R variant, and hypothesized that structural variation might contribute to the disease in this cohort (Lindstrand et al. 2016; Tavares et al. 2019b). Accordingly, in family I, we identified a novel heterozygous deletion encompassing exons 4 to 11 (NM\_024649.4(*BBS1*): c.160-?\_1110+?del), p.(Leu54\_Pro370del)) (Supplementary Figure 1). This variant has not been observed in any of the following publicly available datasets: the Deciphering Developmental Disorders (DDD) project (845 individuals), the 1,000 genomes (2,504 individuals) or the Genome Aggregation Database (gnomAD SV) (10,847 individuals) (Firth et al. 2009; Sudmant et al. 2015; R. L. Collins et al. 2020). Predicted to remove more than 300 AA from the BBS1 protein, it can be classified as pathogenic according to the ACMG criteria (PVS1, PM2, PM3, PM4) (Q. Li et Wang 2017).

Among the remaining six p.M390R mutation-bearing cases, we leveraged a high depth of coverage to identify another putative pathogenic variant in exon 13 of *BBS1* with the deletion/insertion of 2 bases, c.1215\_1216delinsT (p.(Ala406Glnfs\*4)) in family A.II-3 (Supplementary Figure 2A). The variant was found in only 216 out of the 1108 reads at this position (19%). Careful analysis of the genomic region using the Integrative Genomics Viewer (IGV) software (Thorvaldsdottir, Robinson, et Mesirov 2013) revealed a more complicated event leading to incorrect variant calling (Supplementary Figure 2B). Indeed, soft clipped reads part (82 bp) revealed the insertion of an SVA F retrotransposon according to RepeatMasker with a possible truncation of 311 bases 5' of the consensus sequence (Supplementary Figure 2C). To assess the allelic status of both variants, we designed a duplex PCR with an amplicon specific to the SVA insertion (see Methods and Supplementary Figure 3). Co-segregation analysis in family A confirmed compound heterozygosity of both variants (Figure 1A) and the exact insertion point (Figure 1B).

### **Determination of the full length SVA F insertion using WGS data and genome walking**

The SVA F insertion is composed of 6 domains including: (1) hexameric CCCTCT repeats (Figure 1C), (2) followed by an *Alu*-like domain, (3) a Variable Number Tandem Repeats (VNTR) region, (4) a SINE-R corresponding to a retrotransposon coming from retroviral origin, (5) a poly(A) tail and (6) a target site duplication (TSD) on both extremities. The size of a typical SVA F full length sequence is reported to range from 700 bp to 4,000 bp (Kwon et al. 2013b; Hancks et Kazazian 2010; H. Wang et al. 2005; Ostertag et al. 2003b). The insertion site based on a consensus recognition motif or target site (5'-TTTT/AA-3') results in this TSD. To understand the consequences of this insertion, we determined the full-length sequence of this SVA F. For this, we first applied WGS coupled with a bioinformatics protocol and second, Sanger sequencing based on our data and previous reports.

Given the repeated sequence of the SVA F MEI, alignment of short reads onto the reference human genome will generate artifacts (multimapped reads, split reads, discordant read pairs) that can be rescued using a dedicated bioinformatics pipeline. The initial TES aligned reads allowed us to identify 99 bp from the SVA F (Figure 2 and Supplementary Figure 2). We generated WGS data from individual A.II-3 and assembled a 387 bp contig using 14 reads in 5' of the SVA F composed of *BBS1* (38 bp before the breakpoint in exon 13), a TSD (17 bp) and 330 bp that matched the SVA F consensus sequence according to RepeatMasker (5.5% divergence). We found that the *BBS1* SVA F sequence is missing the first 311 bases of the consensus indicating a truncation during the insertion process. Reads in the 3' end of the SVA F could not be assembled through to the poly(A) region (Figure 2). We attempted to lengthen the 5' sequence by performing a strict realignment (bwa\_mem with high mismatch, gap opening and extension penalties) of all reads using the 330 bp SVA sequence as a reference. Although the alignment step could catch more reads, the assembled sequence was only lengthened on the *BBS1* side and did not allow extension of the SVA F sequence further. Performing the same steps using the TES data from the same individual (A.II-3) confirmed the results based on the WGS but did not allow extension of the SVA F sequence either, potentially because of a shorter read length compared to the WGS (100 bp vs 150 bp). In conclusion, WGS data enabled accurate characterization of an additional 250 bp SVA F sequence (Figure 2).

Despite several long-range PCR attempts in various conditions, we determined the full sequence of the SVA F using multiple overlapping Sanger sequencing amplicons with optimized PCR protocols drafted to cross GC rich, repeated and poly(A) regions (Figure 2). Primers were designed based on the assembled patients sequences and

the published consensus SVA F sequence (Kwon et al. 2013b). Amplified fragments sequenced from both side and overlapping regions were assembled to reveal a final inserted sequence of 2,435 bp (Supplementary Figure 4). We identified hallmarks of L1 mediated retrotransposition including: (1) a target site similar at consensus recognition site (TCTT/GA) (Supplementary Figure 5), (2) the 17 bp TSD flanking each extremities of the insertion, (3) a VNTR domain (1,755 bp), (4) a SINE-R region (501 bp) and (5) a polyadenylation signal (AATAA) followed by a poly(A) tail of ~100 nt. The main difference compared to the consensus SVA F sequence, is the missing 311 bases in the 5' part. The variant can now be described as c.1214\_1215ins[MT113356], p.(Ala406Glnfs\*47).

#### **Functional impact of the SVA F insertion on BBS1**

Next, we investigated the functional impact of *the* BBS1 SVA F insertion using primary cells from BBS individual A.II-3. The allele carrying the SVA F insertion is predicted to create a premature stop codon 47 codons after the insertion site leading potentially to a shorter version of the BBS1 protein (452 vs 593 AA) or to a reduced or absent mRNA due to nonsense-mediated decay (NMD). We extracted total RNA from BBS case and sex-matched control fibroblasts, and we tested the expression of the corresponding *BBS1* alleles. qRT-PCR data demonstrated significantly reduced expression of *BBS1* corresponding *a priori* to the loss of the SVA F allele (Figure 3A). By comparison, patients homozygous for the c.1169T>G variant have, as expected, a normal or slightly reduced level of *BBS1* mRNA (Figure 3A). However, using a combination of primers, we demonstrated the expression of a transcript carrying the SVA allele (Supplementary Figure 6). We hypothesize that the expression level of this allele is minor compared to the one carrying the p.M390R.

To quantify mutant BBS1 protein, we carried out western blot analysis on protein extracted from patient A.II-3 primary fibroblasts and the matched control using an antibody targeting the BBS1 N-terminal region (amino acid 187 to 295 from Abnova). Only a single band corresponding to the length of the normal BBS protein could be observed around 50 kDa (Figure 3B). No other band corresponding to a smaller molecular weight (<50 kDa) was detected on the blot demonstrating the absence of a putative BBS1 polypeptide carrying the SVA F. This is also corroborated by the 50% reduced level of BBS1 protein compared to control. As expected, protein from the homozygous p.M390R cells has the same molecular weight and protein quantity compared to wild type.

#### **Screening in a large cohort of BBS patients**

We reanalyzed available TES data from cohort 1 with emphasis on *BBS1* exon 13. For this, we used an upgraded bioinformatics pipeline including Mobster to detect mobile element insertion (MEI) (Thung et al. 2014). Among the 217 samples already sequenced (TES), the pipeline detected on average 0.6 MEI per patient of which the *Alu* elements were the predominant category (50%) among the other MEI (L1 or SVA). Manual assessment of those events using IGV revealed multiple false positives mainly due to repeated regions either within targeted intronic genes or due to off target reads. Mobster could (re)detect the *BBS1* SVA F in family A as a control but also in 2 other samples (B.II-1 and C.II-1). Analysis of the sequencing data revealed the exact same insertion point and a lower ratio of alternate read sequencing data respectively 9% (11/121) and 13% (81/639) compared to family A (Supplementary Figure 7).

To screen additional patients not yet available by TES (cohort 2; 213 cases from 192 families), or an independent group with no pre-selection bias for BBS mutation other than exclusion of *BBS1* recessive cases (cohort 3; 134 cases from 126 families), we used the duplex PCR including the SVA insertion point (Supplementary Figure 1). This led to the identification of 5 additional cases (family D to G). Co-segregation analysis and Sanger sequencing confirmed the biallelic status of the variants from 3 individuals in cohort 2 (coincidentally, all *in trans* to the c.1169T>G, p.M390R change). Additionally, we detected heterozygous *BBS1* SVA F insertions in 2 families of European descent with known primary recessive BBS loci: one pedigree with compound heterozygous *BBS10* mutations p.Cys91Leufs\*5 and p.Arg49Trp (Stoetzel et al. 2006, 10), and another family with recessive *BBS2* changes p.Gln59\* and p.Tyr24\* plus a secondary site variant in *MKKS/BBS6*, p.Gln148\* (Katsanis et al. 2001a). For all additional families Sanger sequencing revealed the exact same insertion point (Figure 4). Together, we identified 8 families who harbor the SVA F insertion.

To estimate the overall incidence of the SVA F as a contributor of the *BBS1* mutational burden, we performed a retrospective analysis of *BBS1* positive cases (e.g. carrier of two pathogenic variants, cohort 4). Data from 131 unrelated cases of European ancestry corresponding to 262 alleles in total could be identified and account for 44 different pathogenic variations. Among those, the c.1169T>G represents 72.1% (189 alleles) while the second most frequent variant is the SVA F with a frequency of 2.3% (6 alleles).

**Founder effect or recurrent mechanism?**

Given the multiple carriers of this insertion in our combined cohorts (8 families), we wondered whether this MEI occurred as an independent (recurrent) mechanism or as the same mutational event derived from common ancestry thus suggesting a founder effect. To estimate the occurrence of such an ultra-rare type of variation, we screened healthy individuals from the general population (European ancestry). Using the duplex PCR described above, 437 control samples were genotyped but no SVA F carrier was identified. Second, we screened WGS data from a French cohort of 862 healthy individuals (FranceGenRef consortium dataset), for which visual inspections of the bam files (IGV batch scripting) did not reveal any SVA F alleles. These data are consistent with a previous report in which the *BBS1* SVA F was shown to be absent from 2,504 controls from 1,000 Genomes (Tavares et al. 2019b) and it is also absent from the recently published gnomAD SV 2.1 dataset comprising 14,891 controls from various origins (R. L. Collins et al. 2020).

Then, we compared the haplotype at the *BBS1* locus from the different patients for which we performed SNP arrays. We analyzed 4 SVA F haplotypes originating from 4 different families, along with 3 patients carrying the c.1169T>G change in the homozygous state (Supplementary Table 4). PCA analysis confirmed the geographical distribution of the patients compared to the 1000 Genomes and the FranceGenRef dataset with central European ancestry (Family A, B and D) and one of Sweden/Denmark ancestry (Family D) (Supplementary Figure 8). We observed the large known common haplotype in patients bearing the homozygous c.1169T>G pathogenic variant (~1.614 Mb). Interestingly, we also identified a common haplotype for patients sharing the SVA F variant but of a much smaller size (~56.6kb) (Supplementary Table 4). Haplotypes shared by the four SVA F carriers were compared to find the position where recombination events are likely to have happened. The estimation of the most recent common ancestor age was 74 generations 95% CI [33;190]. If we assume that one generation lasts 25 years, this will indicate an age of 1850 years with a 95% confidence interval ranging from 825 to 4750 years. Following this, we took the opportunity to estimate the age of the “M390R” haplotype. Taking into account the M390R carriers from family A, B and D together with the 3 homozygous (M390R) controls, we could estimate the age of the most recent common ancestor to 93 generations 95% CI [49;183]. If we assume that one generation lasts 25 years, this will indicate an age of 2325 years with a 95% confidence interval ranging from 1225 to 4575 years. Overall, these data suggest that this MEI is an ultra-rare event that arose from a common ancestor on one of the most common haplotypes at this locus and more recently than the M390R haplotype.

## DISCUSSION

Although the diagnostic yield in BBS is relatively high, (ranging from 60% to 100%), a substantial proportion of cases remain unresolved (Shaheen et al. 2016; Mary et al. 2019; Shamseldin et al. 2020a). This could be related to the attributes of the cohort (sample size and geographic/ethnic origin), the clinical inclusion criteria (strict or more inclusive) or the genetic test used (panel, WES, WGS). Several additional diagnostic challenges are linked to such heterogeneous diseases. Indeed, the ciliopathies exemplify genetic and allelic heterogeneity combined with phenotypic variability (Zaghloul et Katsanis 2010; Mitchison et Valente 2017b). Other specific situations explaining differences might include phenocopy due to multiple implicated genes as demonstrated by us and others (Karaca et al. 2018; Estrada-Cuzcano et al. 2020a). Thus, broad genetic screening approaches such as WES and WGS are especially indicated for diagnosis of ciliopathies. Nonetheless, certain types of variants are not easily detected depending on the test used. Indeed, WES can miss structural variations such as tandem duplications in specific regions of the genomes (GC rich) (Geoffroy, Stoetzel, Scheidecker, Schaefer, Perrault, Bär, Kröll, Delbarre, Antin, Leuvrey, Henry, Blanché, Decker, Kloth, Klaus, Mache, Martin-Coignard, et al. 2018, 140) and WGS will require further studies to interpret deep intronic or regulatory variation (Kremer et al. 2017; Cummings et al. 2017).

In addition to the first case of Mobile Element Insertion (MEI) reported in BBS patients by our group in 2012 (Redin et al. 2012b) and the recent SVA F report in 2019 (Tavares et al. 2019b), here we report 8 additional families with such type of pathogenic variant. Although rarely described or even detected with dedicated screening, a number of MEI cases (>100) have been reported in various rare diseases (Taşkesen et al. 2012) or in cancer (Hayano et al. 2015; Crivelli et al. 2017; Hancks et Kazazian 2012). The majority of the cases are *Alu* insertions (Hancks et Kazazian 2012) and only a few SVA (16). Recently, large studies have evaluated the impact of detecting MEI in patients' genomic data. The Deciphering Developmental Disorders (DDD) project has identified 9 *de novo* MEI (0.04%) from 9,738 trios WES (Gardner et al. 2019). Another report involving >38,871 cases corresponding to 89,874 exomes revealed an increased diagnostic yield of 0.15% (13 verified cases) (Torene et al. 2020) demonstrating at a large scale the necessity to specifically detect those events. Several MEI detection tools have been reported (for review see (Goerner-Potvin et Bourque 2018) and (Torene et al. 2020)). Many false positive calls exist and visual inspection of the sequencing data using IGV (by activating the visualization of the soft clipped reads) is required. Automatic analysis can be done using SV annotation tools such as AnnotSV and requires the integration of specific resources such as polymorphic databases (C. Stewart et al. 2011; J. Wang et al. 2006). As in our case, variant callers can incorrectly detect those events as indels at the insertion point and the same recommendations apply using IGV. Focusing on SVA, only 16 cases have been reported so far with an implication in rare genetic diseases (Hancks et Kazazian 2016; Brett et al. 2017, 1; Kherraf et al. 2018, 66; Tavares et al. 2019b; Torene et al. 2020).

To determine the size, structure and the sequence of the inserted SVA F, we combined bioinformatics analysis with molecular biology experiments. Given the nature of the SVA F sequence, which contains highly repeated elements, GC rich content and a poly(A) tail, complete amplification was not easily performed. We thus used the breakpoints identified by the TES and WGS to walk across the patient's genome and determine a 2,435bp insertion in exon 13 of *BBS1*. This insertion sequence is identical to the one already described by Tavares *et al* (Tavares et al. 2019b). Unlike most of the SVA elements in the human genome (63%) (H. Wang et al. 2005), this inserted SVA F is incomplete lacking 311 bases in 5' and possibly as almost all SVA it is inactive (Hancks et Kazazian 2012). Notably, beside short read standard WGS we also performed linked-read sequencing (10x genomics) to see whether this technology could overcome any of the technical challenges encountered through WGS and analysis (Supplementary Figure 9). Consistent with other reports, we did not find any significant advantage (Uguen et al. 2020).

Interestingly sequencing of the insertion point revealed the exact same sequence for all patients and the genotyping of some individuals revealed a common shared haplotype. The 8 families carrying the SVA F are all of European ancestry and for the 6 for which more details were available, 4 originated from the western part of France (family A, C, E and F), one from the northern part of France (family B) and a last one (family D) originated from Denmark and Sweden. These observations were highly suggestive of a single founder genomic event that we could confirm and estimate the age of the last common ancestor close to 1850 years.

*BBS1* is often the target of structural variations as demonstrated by the frequent reports of copy number variants (Redin et al. 2012b; Lindstrand et al. 2014; 2016; Shaheen et al. 2016; Tavares et al. 2019b; Mary et al. 2019). It is already known, that SVA F are an L1 machinery dependent retrotransposon (Raiz et al. 2012a). Young retrotransposon insertions preferentially occur at "AT" rich genomic regions lacking genes at L1 sequence consensus sites (Gasior et al. 2007). Older retrotransposons are maintained and accumulate in GC rich genomic regions close to genes. Gene deserts host recombination events and insertions will disappear whereas gene rich regions are less affected by these events (Medstrand, Lagemaat, et Mager 2002). *BBS1* is part of the 25% of all

human genes with a high *Alu* and L1 insertion rate (Supplementary Figure 10) and with a higher GC rate (49.9%) compared to the rest of the genome (41%) (Byeon et Kovalchuk 2016). We also identified an L1 recognition site near the TSD sequence with some differences compared to the known consensus (Gasior et al. 2007). These elements might explain the increased number of structural variations at the *BBS1* locus.

In conclusion, we identified the recurrence of a specific SVA F element in exon 13 of *BBS1* in multiple BBS families highlighting the importance of detecting such genetic events. We determined the exact full length sequence of the inserted SVA F, show its pathogenicity and demonstrated the founder effect. This rare allele is the second most frequent cause of *BBS1* after the predominant *BBS1* pathogenic allele, c.1169T>G (p.Met390Arg). We have also developed a duplex PCR surrounding the SVA insertion point, offering a simple and inexpensive way to detect this variant. Taking this into account we have now modified our BBS diagnostic strategy (Supplementary Figure 11).

#### **CONFLICT OF INTEREST STATEMENT**

NK is a significant shareholder in Rescindo Therapeutics. The other authors declare no conflict of interest.

#### **AVAILABILITY OF DATA AND MATERIAL**

Data generated or analyzed during this study are included in the published article and the corresponding supplementary data. The raw sequencing data generated in the course of this study are not publicly available due to the protocol and the corresponding consents used that did not include such information. All variants have been submitted to ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>) (submission SUB7790521). The SVA F full sequence is available in GenBank under the following accession MT113356.

#### **FUNDING INFORMATION**

This work has been published under the framework of the IdEX Unistra and benefits from a funding from the state managed by the French National Research Agency as part of the investments for the future program (JM). This work was supported by US National Institutes of Health grants HD042601 (NK), GM121317 (NK), DK072301 (NK and EED). NK is a Distinguished Valerie and George D. Kennedy Professor. WGS were supported by Fondation APLM and CREGEMES. CD is supported by an FRM grant (ECO20170637509).

#### **AUTHORS CONTRIBUTION**

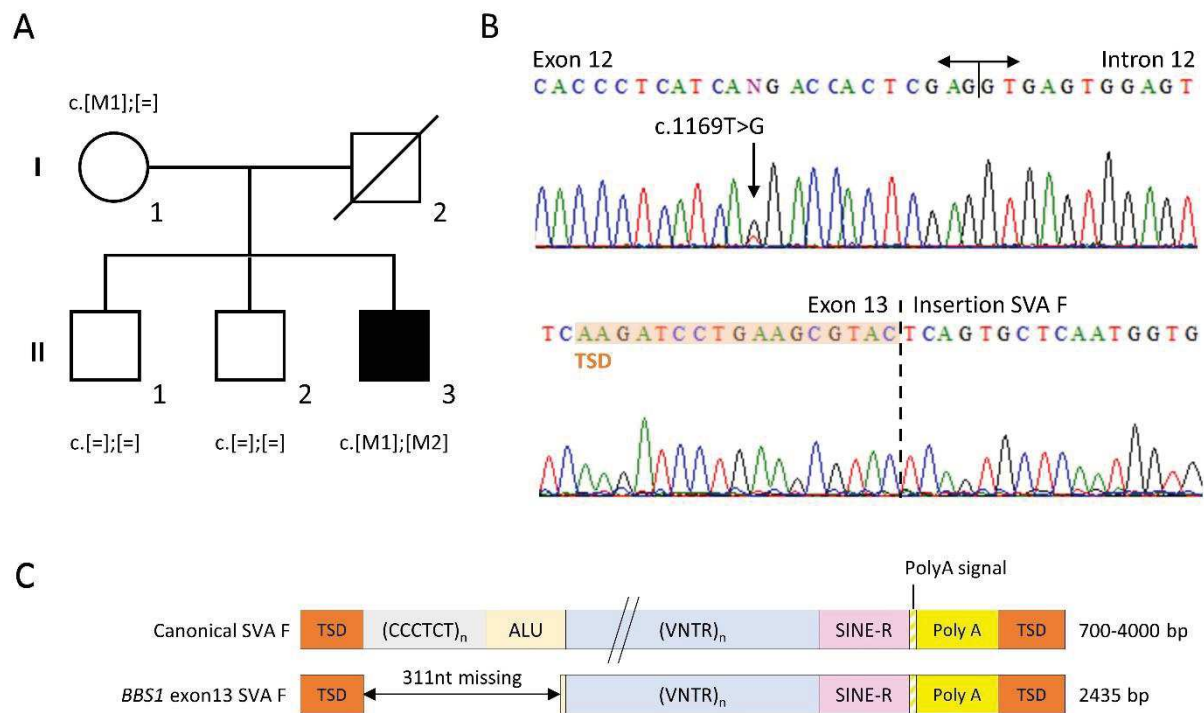
SS, FD, KD, MMD, CPB, SO and HD gathered data from patients and performed clinical investigations; MA, ASL, EN, CCL, GK, CS, JK, ND, JFD, NK, JLM, EED and JM gathered the numerous sequencing data (Sanger, panel, WGS) and performed analyses; CD and CS performed cell biology experiments and data analyses; BK and ChD performed the SNP array experiments and data analyses; EG performed the haplotype ancestry estimation; EG, RR, VG and JM performed the screening in the FranceGenRef dataset; SN, VG and JM performed bioinformatics and data analyses; CD, ED, HD, JM analysed the data and wrote the paper; EG, JLM EED, HD and JM contributed to manuscript writing; JLM, HD and JM provided direction for the project, conceived and designed the experiments.

#### **ACKNOWLEDGEMENTS**

We thank the patients and their family members for their participation, and the French Bardet-Biedl association for their support. We also thank the members of the diagnostic laboratories (molecular and cytogenetic units) at the University Hospital in Strasbourg, France, for technical assistance. Part of TES was performed by the Plateforme GenomEast (IGBMC, Illkirch-Graffenstaden, France) and for this, we thank Serge Vicaire, Murielle Philipps and Bernard Jost. We would like to thank Amélie Piton for scientific discussion and acknowledge Pierre Lindenbaum and every member of the FranceGenRef consortium.



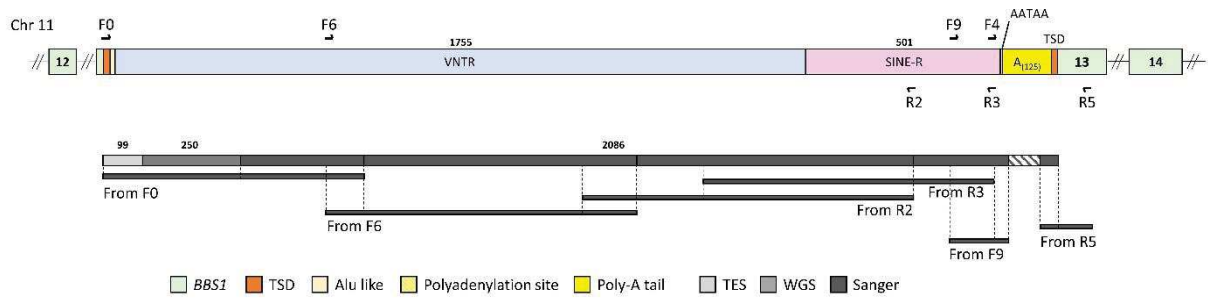
## Figures Legends

**Figure 1. Pedigree of family A and variant analysis**

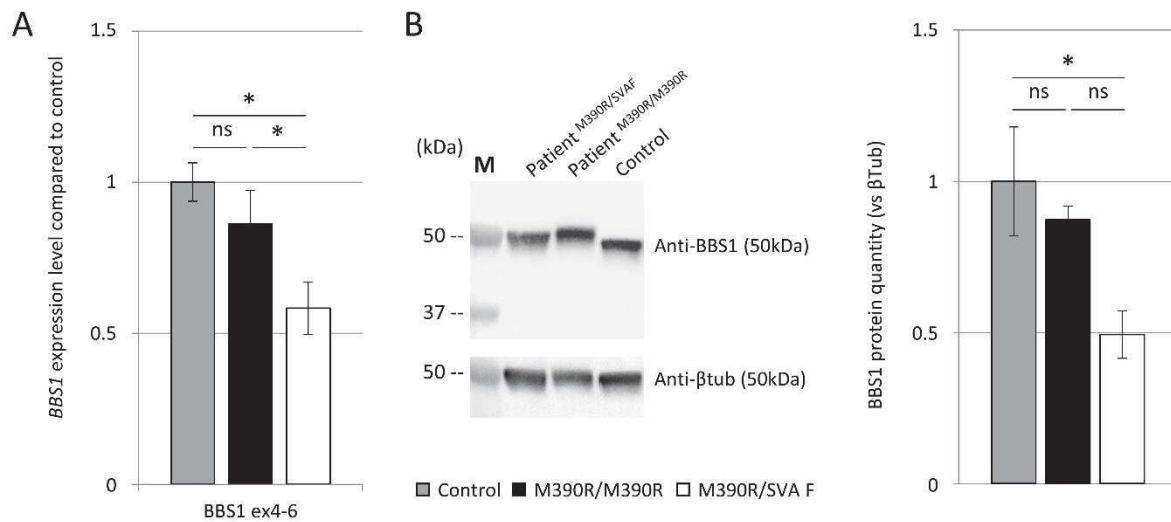
(A) Pedigree of family A. Individual II-3 carries biallelic pathogenic variants: c.1169T>G, p.M390R (M1) and the insertion of an SVA F (M2). Circles, females; squares, males; unfilled shapes, unaffected individuals; filled shape, affected individual; diagonal line, deceased; [=], wild type.

(B) Segregation analysis was performed for all members of the family. Example Sanger sequencing chromatogram for M1 (top) and M2 (bottom).

(C) Schematic of a typical SVA F (top) and variant SVA F in *BBS1* (bottom). A typical SVA F is composed of 6 parts including: (1) hexameric CCCTCT repeats followed by (2) an *Alu*-like domain, (3) a Variable Number Tandem Repeats (VNTR) domain, (4) a Short Interspersed Nucleotide Elements (SINE-R) domain, (5) a poly(A) tail and (6) the target site duplication (TSD) at both extremities.



**Figure 2. Schematic of the inserted SVA F element in exon 13 of *BBS1*.** Primer locations and amplification products show how each experiment contributed to the resolution of the full length sequence.

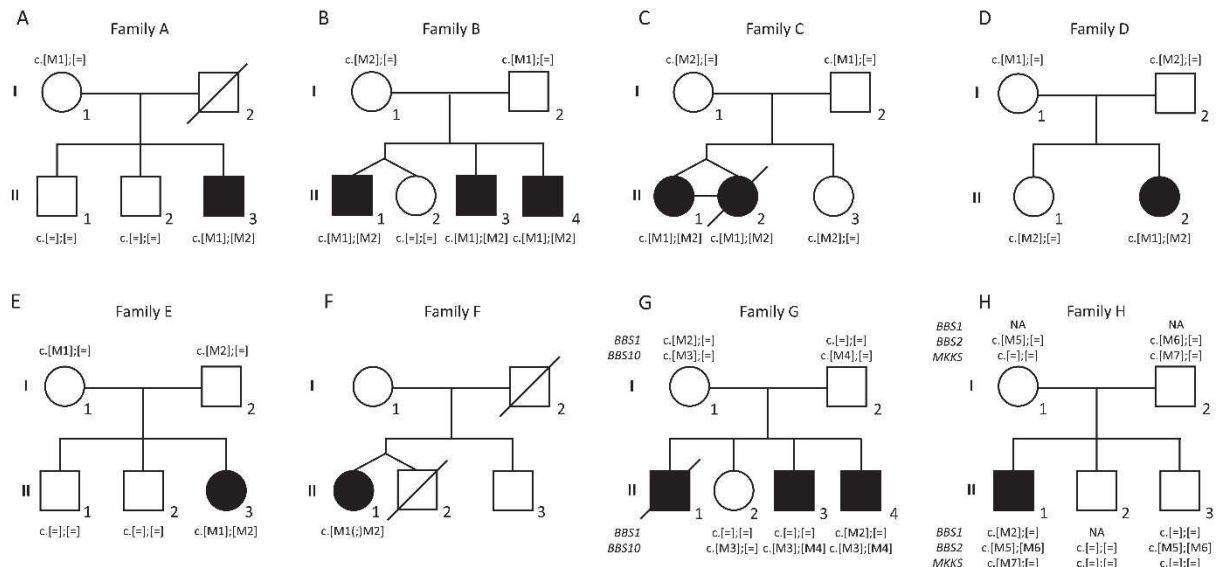


**Figure 3. Quantification of BBS1 mRNA and protein levels in the context of the BBS1 SVA F.**

(A) Analysis of *BBS1* mRNA expression by RT-qPCR. The *BBS1* expression level is significantly reduced in patient A.II-3 compared to control or homozygous p.M390R patient. The major fraction of *BBS1* expression likely corresponds to the p.M390R allele. Normalization was performed using both *GAPDH* and *HPRT* as reference genes. Mean of both results are shown. Controls are in gray (n=3), homozygous p.M390R in black (n=2), and patient A.II-3 (p.M390R and SVA F) in white (n=2). Bars show the mean of 2 independent experiments +/- SEM (n=2, t-test \*: p<0.05, ns: not significant).

(B) Analysis of the BBS1 protein level by western blot (n=3 replicates). No truncated version of BBS1 could be detected. The BBS1 protein detected in fibroblast cells from A.II-3 is half of the quantity compared to control. Bars show mean of 3 independent experiments +/- SEM (n=3, t-test \*: p<0.05, ns: not significant). M: molecular weight ladder (kDa).





**Figure 4.** Pedigrees of families with affected individuals carrying the SVA F in *BBS1*. M1: *BBS1*:c.1169T>G (p.M390R), M2: *BBS1*:c.1214\_1215ins[MT113356] (p.(Ala406Glnfs\*47), insertion of SVA F), M3: *BBS10*:c.271dupT (p.Cys91Leufs\*5), M4: *BBS10*:c.145C>T, (p.Arg49Trp), M5:*BBS2*:c.72C>T (p.Tyr24\*), M6: *BBS2*:c.175C>T (p.Gln59\*), M7: *MKKS/BBS6*:c.442C>T, (p.Gln148\*). Circles, females; squares, males; unfilled shapes, unaffected individuals; filled shape, affected individual; diagonal line, deceased; [=], wild type.

## REFERENCES

- Akman, Hasan O., Guido Davidzon, Kurenai Tanji, Emma J. MacDermott, Louann Larsen, Mercy M. Davidson, Ronald G. Haller, et al. 2010a. « Neutral Lipid Storage Disease with Subclinical Myopathy Due to a Retrotransposal Insertion in the PNPLA2 Gene ». *Neuromuscular Disorders* 20 (6): 397-402. <https://doi.org/10.1016/j.nmd.2010.04.004>.
- Auton, Adam, Gonçalo R. Abecasis, David M. Altshuler, Richard M. Durbin, Gonçalo R. Abecasis, David R. Bentley, Aravinda Chakravarti, et al. 2015. « A global reference for human genetic variation ». *Nature* 526 (7571): 68-74. <https://doi.org/10.1038/nature15393>.
- Backenroth, Daniel, Jason Homsy, Laura R. Murillo, Joe Glessner, Edwin Lin, Martina Brueckner, Richard Lifton, Elizabeth Goldmuntz, Wendy K. Chung, et Yufeng Shen. 2014. « CANOES: Detecting Rare Copy Number Variants from Whole Exome Sequencing Data ». *Nucleic Acids Research* 42 (12): e97. <https://doi.org/10.1093/nar/gku345>.
- Badano, Jose L., Jun Chul Kim, Bethan E. Hoskins, Richard Alan Lewis, Stephen J. Ansley, David J. Cutler, Claudio Castellán, Philip L. Beales, Michel R. Leroux, et Nicholas Katsanis. 2003. « Heterozygous Mutations in *BBS1*, *BBS2* and *BBS6* Have a Potential Epistatic Effect on Bardet–Biedl Patients with Two Mutations at a Second *BBS* Locus ». *Human Molecular Genetics* 12 (14): 1651-59. <https://doi.org/10.1093/hmg/ddg188>.
- Beales, P. L., N. Elcioglu, A. S. Woolf, D. Parker, et F. A. Flinter. 1999. « New Criteria for Improved Diagnosis of Bardet-Biedl Syndrome: Results of a Population Survey ». *Journal of Medical Genetics* 36 (6): 437-46.

- Beales, Philip L., Jose L. Badano, Alison J. Ross, Stephen J. Ansley, Bethan E. Hoskins, Brigitta Kirsten, Charles A. Mein, et al. 2003. « Genetic Interaction of BBS1 Mutations with Alleles at Other BBS Loci Can Result in Non-Mendelian Bardet-Biedl Syndrome ». *The American Journal of Human Genetics* 72 (5): 1187-99. <https://doi.org/10.1086/375178>.
- Bragg, D. Christopher, Kotchaphorn Mangkalaphiban, Christine A. Vaine, Nichita J. Kulkarni, David Shin, Rachita Yadav, Jyotsna Dhakal, et al. 2017. « Disease Onset in X-Linked Dystonia-Parkinsonism Correlates with Expansion of a Hexameric Repeat within an SVA Retrotransposon in *TAF1* ». *Proceedings of the National Academy of Sciences* 114 (51): E11020-28. <https://doi.org/10.1073/pnas.1712526114>.
- Brett, Maggie, George Korovesis, Angeline H. M. Lai, Eileen C. P. Lim, et Ene-Choo Tan. 2017. « Intragenic Multi-Exon Deletion in the *FBN1* Gene in a Child with Mildly Dilated Aortic Sinus: A Retrotransposal Event ». *Journal of Human Genetics* 62 (7): 711-15. <https://doi.org/10.1038/jhg.2017.32>.
- Buckley, Reuben M., R. Daniel Kortschak, Joy M. Raison, et David L. Adelson. 2017. « Similar Evolutionary Trajectories for Retrotransposon Accumulation in Mammals ». *Genome Biology and Evolution* 9 (9): 2336-53. <https://doi.org/10.1093/gbe/evx179>.
- Burns, Kathleen H. 2020. « Our Conflict with Transposable Elements and Its Implications for Human Disease ». *Annual Review of Pathology: Mechanisms of Disease* 15 (1): 51-70. <https://doi.org/10.1146/annurev-pathmechdis-012419-032633>.
- Byeon, Boseon, et Igor Kovalchuk. 2016. « Non-Coding RNAs Match the Deleted Genomic Regions in Humans ». *Scientific Reports* 6 (1). <https://doi.org/10.1038/srep37452>.
- Chang, Christopher C, Carson C Chow, Laurent CAM Tellier, Shashaank Vattikuti, Shaun M Purcell, et James J Lee. 2015. « Second-generation PLINK: rising to the challenge of larger and richer datasets ». *GigaScience* 4 (s13742-015-0047-8). <https://doi.org/10.1186/s13742-015-0047-8>.
- Chatron, Nicolas, Kevin Cassinari, Olivier Quenez, Stéphanie Baert-Desurmont, Claire Bardel, Marie-Pierre Buisine, Eduardo Calpena, et al. 2019. « Identification of mobile retrocopies during genetic testing: Consequences for routine diagnosis ». *Human Mutation* 40 (11): 1993-2000. <https://doi.org/10.1002/humu.23845>.
- Chen, Jian-Min, Emmanuelle Masson, Milan Macek, Odile Raguénès, Tereza Piskackova, Brigitte Fercot, Libor Fila, David N. Cooper, Marie-Pierre Audrézet, et Claude Férec. 2008. « Detection of Two Alu Insertions in the *CFTR* Gene ». *Journal of Cystic Fibrosis* 7 (1): 37-43. <https://doi.org/10.1016/j.jcf.2007.04.001>.
- Collins, Ryan L., Harrison Brand, Konrad J. Karczewski, Xuefang Zhao, Jessica Alföldi, Laurent C. Francioli, Amit V. Khera, et al. 2020. « A structural variation reference for medical and population genetics ». *Nature* 581 (7809): 444-51. <https://doi.org/10.1038/s41586-020-2287-8>.
- Conley, Mary Ellen, Julie D. Partain, Shannon M. Norland, Sheila A. Shurtleff, et Haig H. Kazazian. 2005. « Two Independent Retrotransposon Insertions at the Same Site within the Coding Region of *BTK* ». *Human Mutation* 25 (3): 324-25. <https://doi.org/10.1002/humu.9321>.
- Crivelli, Louise, Virginie Bubien, Natalie Jones, Jennifer Chiron, Françoise Bonnet, Emmanuelle Barouk-Simonet, Patrice Couzigou, Nicolas Sevenet, Frédéric Caux, et Michel Longy. 2017. « Insertion of Alu Elements at a *PTEN* Hotspot in Cowden Syndrome ». *European Journal of Human Genetics : EJHG* 25 (9): 1087-91. <https://doi.org/10.1038/ejhg.2017.81>.

Cummings, Beryl B., Jamie L. Marshall, Taru Tukiainen, Monkol Lek, Sandra Donkervoort, A. Reghan Foley, Veronique Bolduc, et al. 2017. « Improving genetic diagnosis in Mendelian disease with transcriptome sequencing ». *Science Translational Medicine* 9 (386): eaal5209. <https://doi.org/10.1126/scitranslmed.aal5209>.

DePristo, Mark A, Eric Banks, Ryan Poplin, Kiran V Garimella, Jared R Maguire, Christopher Hartl, Anthony A Philippakis, et al. 2011. « A framework for variation discovery and genotyping using next-generation DNA sequencing data ». *Nature Genetics* 43 (5): 491-98. <https://doi.org/10.1038/ng.806>.

Estrada-Cuzcano, Alejandro, Christelle Etard, Clarisse Delvallée, Corinne Stoetzel, Elise Schaefer, Sophie Scheidecker, Véronique Geoffroy, et al. 2020. « Novel IQCE variations confirm its role in postaxial polydactyly and cause ciliary defect phenotype in zebrafish ». *Human Mutation* 41 (1): 240-54. <https://doi.org/10.1002/humu.23924>.

Farag, Talaat I., et Ahmad S. Teebi. 1989. « High incidence of Bardet Biedl syndrome among the Bedouin ». *Clinical Genetics* 36 (6): 463-64. <https://doi.org/10.1111/j.1399-0004.1989.tb03378.x>.

Firth, Helen V., Shola M. Richards, A. Paul Bevan, Stephen Clayton, Manuel Corpas, Diana Rajan, Steven Van Vooren, Yves Moreau, Roger M. Pettett, et Nigel P. Carter. 2009. « DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources ». *The American Journal of Human Genetics* 84 (4): 524-33. <https://doi.org/10.1016/j.ajhg.2009.03.010>.

Gardner, Eugene J., Elena Prigmore, Giuseppe Gallone, Petr Danecek, Kaitlin E. Samocha, Juliet Handsaker, Sebastian S. Gerety, et al. 2019. « Contribution of retrotransposition to developmental disorders ». *Nature Communications* 10 (1): 4630. <https://doi.org/10.1038/s41467-019-12520-y>.

Gasior, Stephen L., Graeme Preston, Dale J. Hedges, Nicolas Gilbert, John V. Moran, et Prescott L. Deininger. 2007. « Characterization of Pre-Insertion Loci of de Novo L1 Insertions ». *Gene* 390 (1-2): 190-98. <https://doi.org/10.1016/j.gene.2006.08.024>.

Genin, E, A Tullio-Pelet, F Begeot, S Lyonnet, et L Abel. 2004. « Estimating the age of rare disease mutations: the example of Triple-A syndrome ». *Journal of Medical Genetics* 41 (6): 445. <https://doi.org/10.1136/jmg.2003.017962>.

Geoffroy, Véronique, Yvan Herenger, Arnaud Kress, Corinne Stoetzel, Amélie Piton, Hélène Dollfus, et Jean Muller. 2018. « AnnotSV: An Integrated Tool for Structural Variations Annotation ». *Bioinformatics (Oxford, England)* 34 (20): 3572-74. <https://doi.org/10.1093/bioinformatics/bty304>.

Geoffroy, Véronique, Cécile Pizot, Claire Redin, Amélie Piton, Nasim Vasli, Corinne Stoetzel, André Blavier, Jocelyn Laporte, et Jean Muller. 2015. « VaRank: A Simple and Powerful Tool for Ranking Genetic Variants ». *PeerJ* 3: e796. <https://doi.org/10.7717/peerj.796>.

Geoffroy, Véronique, Corinne Stoetzel, Sophie Scheidecker, Elise Schaefer, Isabelle Perrault, Séverine Bär, Ariane Kröll, et al. 2018. « Whole-genome sequencing in patients with ciliopathies uncovers a novel recurrent tandem duplication in IFT140 ». *Human Mutation* 39 (7): 983-92. <https://doi.org/10.1002/humu.23539>.

Goerner-Potvin, Patricia, et Guillaume Bourque. 2018. « Computational tools to unmask transposable elements ». *Nature Reviews Genetics* 19 (11): 688-704. <https://doi.org/10.1038/s41576-018-0050-x>.

- Gouronc, Aurélie, Vincent Zilliox, Marie-Line Jacquemont, Françoise Darcel, Anne-Sophie Leuvrey, Elsa Nourisson, Manuela Antin, et al. 2020. « High prevalence of Bardet-Biedl syndrome in La Réunion Island is due to a founder variant in ARL6/BBS3 ». *Clinical Genetics* n/a (n/a). <https://doi.org/10.1111/cge.13768>.
- Ha, Hongseok, Jui Wan Loh, et Jinchuan Xing. 2016. « Identification of Polymorphic SVA Retrotransposons Using a Mobile Element Scanning Method for SVA (ME-Scan-SVA) ». *Mobile DNA* 7 (1). <https://doi.org/10.1186/s13100-016-0072-x>.
- Hancks, Dustin C., et Haig H. Kazazian. 2010. « SVA Retrotransposons: Evolution and Genetic Instability ». *Seminars in Cancer Biology* 20 (4): 234-45. <https://doi.org/10.1016/j.semcancer.2010.04.001>.
- Hancks, Dustin C, et Haig H Kazazian. 2012. « Active Human Retrotransposons: Variation and Disease ». *Current Opinion in Genetics & Development* 22 (3): 191-203. <https://doi.org/10.1016/j.gde.2012.02.006>.
- Hancks, Dustin C., et Haig H. Kazazian. 2016. « Roles for retrotransposon insertions in human disease ». *Mobile DNA* 7 (mai). <https://doi.org/10.1186/s13100-016-0065-9>.
- Hassoun, H., T. L. Coetzer, J. N. Vassiliadis, K. E. Sahr, G. J. Maalouf, S. T. Saad, L. Catanzariti, et J. Palek. 1994. « A Novel Mobile Element Inserted in the Alpha Spectrin Gene: Spectrin Dayton. A Truncated Alpha Spectrin Associated with Hereditary Elliptocytosis ». *The Journal of Clinical Investigation* 94 (2): 643-48. <https://doi.org/10.1172/JCI117380>.
- Hayano, Takahide, Shiro Yamada, Kazuyoshi Hosomichi, Hirofumi Nakaoka, Kosuke Yoshihara, Sosuke Adachi, Katsunori Kashima, Kenichi Tanaka, Takayuki Enomoto, et Ituro Inoue. 2015. « Identification of Novel Exonic Mobile Element Insertions in Epithelial Ovarian Cancers ». *Human Genome Variation* 2 (octobre): 15030-15030. <https://doi.org/10.1038/hgv.2015.30>.
- Karaca, Ender, Jennifer E. Posey, Zeynep Coban Akdemir, Davut Pehlivan, Tamar Harel, Shalini N. Jhangiani, Yavuz Bayram, et al. 2018. « Phenotypic expansion illuminates multilocus pathogenic variation ». *Genetics in Medicine* 20 (12): 1528-37. <https://doi.org/10.1038/gim.2018.33>.
- Katsanis, Nicholas, Stephen J. Ansley, Jose L. Badano, Erica R. Eichers, Richard Alan Lewis, Bethan E. Hoskins, Peter J. Scambler, William S. Davidson, Philip L. Beales, et James R. Lupski. 2001. « Triallelic Inheritance in Bardet-Biedl Syndrome, a Mendelian Recessive Disorder ». *Science* 293 (5538): 2256. <https://doi.org/10.1126/science.1063525>.
- Kazazian, Haig H., Corinne Wong, Hagop Youssoufian, Alan F. Scott, Deborah G. Phillips, et Stylianos E. Antonarakis. 1988a. « Haemophilia A Resulting from de Novo Insertion of L1 Sequences Represents a Novel Mechanism for Mutation in Man ». *Nature* 332 (6160): 164-66. <https://doi.org/10.1038/332164a0>.
- Kent, W James. 2002. « BLAT--the BLAST-like Alignment Tool ». *Genome Research* 12 (4): 656-64. <https://doi.org/10.1101/gr.229202>.
- Khan, Sher Alam, Noor Muhammad, Muzammil Ahmad Khan, Anwar Kamal, Zia Ur Rehman, et Saadullah Khan. 2016. « Genetics of Human Bardet-Biedl Syndrome, an Updates ». *Clinical Genetics*, janvier. <https://doi.org/10.1111/cge.12737>.

- Kherraf, Zine-Eddine, Amir Amiri-Yekta, Denis Dacheux, Thomas Karaouzène, Charles Coutton, Marie Christou-Kent, Guillaume Martinez, et al. 2018. « A Homozygous Ancestral SVA-Insertion-Mediated Deletion in WDR66 Induces Multiple Morphological Abnormalities of the Sperm Flagellum and Male Infertility ». *The American Journal of Human Genetics* 103 (3): 400-412. <https://doi.org/10.1016/j.ajhg.2018.07.014>.
- Klift, Heleen M. van der, Carli M. Tops, Frederik J. Hes, Peter Devilee, et Juul T. Wijnen. 2012. « Insertion of an SVA Element, a Nonautonomous Retrotransposon, in *PMS2* Intron 7 as a Novel Cause of Lynch Syndrome ». *Human Mutation* 33 (7): 1051-55. <https://doi.org/10.1002/humu.22092>.
- Kobayashi, K., Y. Nakahori, M. Miyake, K. Matsumura, E. Kondo-Iida, Y. Nomura, M. Segawa, et al. 1998. « An Ancient Retrotransposal Insertion Causes Fukuyama-Type Congenital Muscular Dystrophy ». *Nature* 394 (6691): 388-92. <https://doi.org/10.1038/28653>.
- Kremer, Laura S., Daniel M. Bader, Christian Mertes, Robert Kopajtich, Garwin Pichler, Arcangela Iuso, Tobias B. Haack, et al. 2017. « Genetic diagnosis of Mendelian disorders via RNA sequencing ». *Nature Communications* 8 (1): 15824. <https://doi.org/10.1038/ncomms15824>.
- Kwon, Yun-Jeong, Yuri Choi, Jungwoo Eo, Yu-Na Noh, Jeong-An Gim, Yi-Deun Jung, Ja-Rang Lee, et Heui-Soo Kim. 2013a. « Structure and Expression Analyses of SVA Elements in Relation to Functional Genes ». *Genomics & Informatics* 11 (3): 142. <https://doi.org/10.5808/GI.2013.11.3.142>.
- Lee, Christopher M, Galt P Barber, Jonathan Casper, Hiram Clawson, Mark Diekhans, Jairo Navarro Gonzalez, Angie S Hinrichs, et al. 2019. « UCSC Genome Browser enters 20th year ». *Nucleic Acids Research* 48 (D1): D756-61. <https://doi.org/10.1093/nar/gkz1012>.
- Li, Heng. 2011. « A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data ». *Bioinformatics* 27 (21): 2987-93. <https://doi.org/10.1093/bioinformatics/btr509>.
- Li, Heng, et Richard Durbin. 2009. « Fast and accurate short read alignment with Burrows–Wheeler transform ». *Bioinformatics* 25 (14): 1754-60. <https://doi.org/10.1093/bioinformatics/btp324>.
- Li, Quan, et Kai Wang. 2017. « InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines ». *The American Journal of Human Genetics* 100 (2): 267-80. <https://doi.org/10.1016/j.ajhg.2017.01.004>.
- Lindstrand, Anna, Erica E. Davis, Claudia M. B. Carvalho, Davut Pehlivan, Jason R. Willer, I.-Chun Tsai, Subhadra Ramanathan, et al. 2014. « Recurrent CNVs and SNVs at the NPHP1 Locus Contribute Pathogenic Alleles to Bardet-Biedl Syndrome. » *American Journal of Human Genetics* 94 (5): 745-54. <https://doi.org/10.1016/j.ajhg.2014.03.017>.
- Lindstrand, Anna, Stephan Frangakis, Claudia M.B. Carvalho, Ellen B. Richardson, Kelsey A. McFadden, Jason R. Willer, Davut Pehlivan, et al. 2016. « Copy-Number Variation Contributes to the Mutational Load of Bardet-Biedl Syndrome ». *The American Journal of Human Genetics* 99 (2): 318-36. <https://doi.org/10.1016/j.ajhg.2015.04.023>.
- Martin, Sandra L. 2018. « On the move ». *eLife* 7 (février): e34901. <https://doi.org/10.7554/eLife.34901>.



- Mary, Laura, Kirsley Chennen, Corinne Stoetzel, Manuela Antin, Anne Leuvrey, Elsa Nourisson, Elisabeth Alanio-Detton, et al. 2019. « Bardet-Biedl syndrome: Antenatal presentation of forty-five fetuses with biallelic pathogenic variants in known Bardet-Biedl syndrome genes ». *Clinical Genetics* 95 (3): 384-97. <https://doi.org/10.1111/cge.13500>.
- Mathieson, Iain, et Gil McVean. 2014. « Demography and the Age of Rare Variants ». *PLOS Genetics* 10 (8): e1004528. <https://doi.org/10.1371/journal.pgen.1004528>.
- Medina, Julie, et Hervé Perron. 2017. « Séquences provenant d'éléments génétiques mobiles, face cachée du génome humain ». *médecine/sciences* 33 (2): 151-58. <https://doi.org/10.1051/medsci/20173302010>.
- Medstrand, Patrik, Louie N. van de Lagemaat, et Dixie L. Mager. 2002. « Retroelement Distributions in the Human Genome: Variations Associated With Age and Proximity to Genes ». *Genome Research* 12 (10): 1483-95. <https://doi.org/10.1101/gr.388902>.
- Miki, Yoshio, Toyomasa Katagiri, Fujio Kasumi, Takamasa Yoshimoto, et Yusuke Nakamura. 1996. « Mutation Analysis in the BRCA2 Gene in Primary Breast Cancers ». *Nature Genetics* 13 (2): 245-47. <https://doi.org/10.1038/ng0696-245>.
- Mills, Ryan E., E. Andrew Bennett, Rebecca C. Iskow, et Scott E. Devine. 2007. « Which Transposable Elements Are Active in the Human Genome? » *Trends in Genetics* 23 (4): 183-91. <https://doi.org/10.1016/j.tig.2007.02.006>.
- Mitchison, Hannah M, et Enza Maria Valente. 2017. « Motile and non-motile cilia in human pathology: from function to phenotypes ». *The Journal of Pathology* 241 (2): 294-309. <https://doi.org/10.1002/path.4843>.
- Muller, Jean, C. Stoetzel, M. C. Vincent, C. C. Leitch, V. Laurier, J. M. Danse, S. Hellé, et al. 2010. « Identification of 28 novel mutations in the Bardet-Biedl syndrome genes: the burden of private mutations in an extensively heterogeneous disease ». *Human genetics* 127 (5): 583-93. <https://doi.org/10.1007/s00439-010-0804-9>.
- Mykytyn, Kirk, Darryl Y. Nishimura, Charles C. Searby, Gretel Beck, Kevin Bugge, Heidi L. Haines, Alberto S. Cornier, et al. 2003. « Evaluation of Complex Inheritance Involving the Most Common Bardet-Biedl Syndrome Locus (BBS1) ». *The American Journal of Human Genetics* 72 (2): 429-37. <https://doi.org/10.1086/346172>.
- Mykytyn, Kirk, Darryl Y. Nishimura, Charles C. Searby, Mythreyi Shastri, Hsan-jan Yen, John S. Beck, Terry Braun, et al. 2002. « Identification of the Gene (BBS1) Most Commonly Involved in Bardet-Biedl Syndrome, a Complex Human Obesity Syndrome ». *Nature Genetics* 31 (4): 435-38. <https://doi.org/10.1038/ng935>.
- Nakamura, Yuki, Moe Murata, Yuki Takagi, Toshihiro Kozuka, Yukiko Nakata, Ryo Hasebe, Akira Takagi, Jun-ichi Kitazawa, Midori Shima, et Tetsuhito Kojima. 2015. « SVA Retrotransposition in Exon 6 of the Coagulation Factor IX Gene Causing Severe Hemophilia B ». *International Journal of Hematology* 102 (1): 134-39. <https://doi.org/10.1007/s12185-015-1765-5>.
- Niederlova, Veronika, Martin Modrak, Oksana Tsyklauri, Martina Huranova, et Ondrej Stepanek. 2019. « Meta-analysis of genotype-phenotype associations in Bardet-Biedl Syndrome uncovers differences among causative genes ». *Human Mutation* 0 (ja). <https://doi.org/10.1002/humu.23862>.

Okonechnikov, Konstantin, Ana Conesa, et Fernando García-Alcalde. 2015. « Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data ». *Bioinformatics* 32 (2): 292-94. <https://doi.org/10.1093/bioinformatics/btv566>.

O'Leary, Nuala A., Mathew W. Wright, J. Rodney Brister, Stacy Ciuffo, Diana Haddad, Rich McVeigh, Bhanu Rajput, et al. 2015. « Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation ». *Nucleic Acids Research* 44 (D1): D733-45. <https://doi.org/10.1093/nar/gkv1189>.

Ostertag, Eric M., John L. Goodier, Yue Zhang, et Haig H. Kazazian. 2003a. « SVA Elements Are Nonautonomous Retrotransposons That Cause Disease in Humans ». *The American Journal of Human Genetics* 73 (6): 1444-51. <https://doi.org/10.1086/380207>.

Raiz, Julija, Annette Damert, Sergiu Chira, Ulrike Held, Sabine Klawitter, Matthias Hamdorf, Johannes Löwer, Wolf H. Strätling, Roswitha Löwer, et Gerald G. Schumann. 2012a. « The Non-Autonomous Retrotransposon SVA Is Trans -Mobilized by the Human LINE-1 Protein Machinery ». *Nucleic Acids Research* 40 (4): 1666-83. <https://doi.org/10.1093/nar/gkr863>.

Redin, Claire, Stéphanie Le Gras, Oussema Mhamdi, Véronique Geoffroy, Corinne Stoetzel, Marie-Claire Vincent, Pietro Chiurazzi, et al. 2012a. « Targeted high-throughput sequencing for diagnosis of genetically heterogeneous diseases: efficient mutation detection in Bardet-Biedl and Alström Syndromes ». *Journal of Medical Genetics* 49 (8): 502-12. <https://doi.org/10.1136/jmedgenet-2012-100875>.

Sayers, Eric W., Jeff Beck, J. Rodney Brister, Evan E. Bolton, Kathi Canese, Donald C. Comeau, Kathryn Funk, et al. 2019. « Database Resources of the National Center for Biotechnology Information. » *Nucleic Acids Research*, octobre. <https://doi.org/10.1093/nar/gkz899>.

Scheidecker, Sophie, Christelle Etard, Nathan W. Pierce, Véronique Geoffroy, Elise Schaefer, Jean Muller, Kirsley Chennen, et al. 2014. « Exome Sequencing of Bardet-Biedl Syndrome Patient Identifies a Null Mutation in the BBSome Subunit BBIP1 (BBS18) ». *Journal of Medical Genetics* 51 (2): 132-36. <https://doi.org/10.1136/jmedgenet-2013-101785>.

Shaheen, Ranad, Katarzyna Szymanska, Basudha Basu, Nisha Patel, Nour Ewida, Eissa Faqeih, Amal Al Hashem, et al. 2016. « Characterizing the morbid genome of ciliopathies ». *Genome Biology* 17 (1): 242. <https://doi.org/10.1186/s13059-016-1099-5>.

Shamseldin, Hanan E., Ranad Shaheen, Nour Ewida, Dalal K. Bubshait, Hisham Alkuraya, Elham Almardawi, Ali Howaidi, et al. 2020. « The morbid genome of ciliopathies: an update ». *Genetics in Medicine*, février. <https://doi.org/10.1038/s41436-020-0761-1>.

Smit, AFA, R Hubley, et P Green. 2013. « RepeatMasker Open-4.0 », 2015 2013. <http://www.repeatmasker.org>.

Stemers, Frank J, Weihua Chang, Grace Lee, David L Barker, Richard Shen, et Kevin L Gunderson. 2006. « Whole-genome genotyping with the single-base extension assay ». *Nature Methods* 3 (1): 31-33. <https://doi.org/10.1038/nmeth842>.

Stewart, Chip, Deniz Kural, Michael P Strömberg, Jerilyn A Walker, Miriam K Konkel, Adrian M Stütz, Alexander E Urban, et al. 2011. « A Comprehensive Map of Mobile Element Insertion Polymorphisms in Humans ». *PLoS Genetics* 7 (8): e1002236-e1002236. <https://doi.org/10.1371/journal.pgen.1002236>.



- Stoetzel, Corinne, Virginie Laurier, Erica E. Davis, Jean Muller, Suzanne Rix, José L. Badano, Carmen C. Leitch, et al. 2006. « BBS10 Encodes a Vertebrate-Specific Chaperonin-like Protein and Is a Major BBS Locus ». *Nature Genetics* 38 (5): 521-24. <https://doi.org/10.1038/ng1771>.
- Sudmant, Peter H., Tobias Rausch, Eugene J. Gardner, Robert E. Handsaker, Alexej Abyzov, John Huddleston, Yan Zhang, et al. 2015. « An integrated map of structural variation in 2,504 human genomes ». *Nature* 526 (7571): 75-81. <https://doi.org/10.1038/nature15394>.
- Takasu, M., R. Hayashi, E. Maruya, M. Ota, K. Imura, K. Kougo, C. Kobayashi, et al. 2007. « Deletion of Entire HLA-A Gene Accompanied by an Insertion of a Retrotransposon ». *Tissue Antigens* 70 (2): 144-50. <https://doi.org/10.1111/j.1399-0039.2007.00870.x>.
- Taniguchi-Ikeda, Mariko, Kazuhiro Kobayashi, Motoi Kanagawa, Chih-chieh Yu, Kouhei Mori, Tetsuya Oda, Atsushi Kuga, et al. 2011. « Pathogenic Exon-Trapping by SVA Retrotransposon and Rescue in Fukuyama Muscular Dystrophy ». *Nature* 478 (7367): 127-31. <https://doi.org/10.1038/nature10456>.
- Taşkesen, Mustafa, Gayle B Collin, Alexei V Evsikov, Ayşegül Güzel, R Köksal Özgül, Jan D Marshall, et Jürgen K Naggert. 2012. « Novel Alu Retrotransposon Insertion Leading to Alström Syndrome ». *Human Genetics* 131 (3): 407-13. <https://doi.org/10.1007/s00439-011-1083-9>.
- Tavares, Erika, Chen Yu Tang, Anjali Vig, Shuning Li, Gail Billingsley, Wilson Sung, Ajoy Vincent, Bhooma Thiruvahindrapuram, et Elise Héon. 2019a. « Retrotransposon Insertion as a Novel Mutational Event in Bardet-Biedl Syndrome ». *Molecular Genetics & Genomic Medicine* 7 (2): e00521. <https://doi.org/10.1002/mgg3.521>.
- The 1000 Genomes Project Consortium, Richard A. Gibbs, Eric Boerwinkle, Harsha Doddapaneni, Yi Han, Viktoriya Korchina, Christie Kovar, et al. 2015. « A Global Reference for Human Genetic Variation ». *Nature* 526 (7571): 68-74. <https://doi.org/10.1038/nature15393>.
- Thorvaldsdottir, H., J. T. Robinson, et J. P. Mesirov. 2013. « Integrative Genomics Viewer (IGV): High-Performance Genomics Data Visualization and Exploration ». *Briefings in Bioinformatics* 14 (2): 178-92. <https://doi.org/10.1093/bib/bbs017>.
- Thung, Djie Tjwan, Joep de Ligt, Lisenka EM Vissers, Marloes Steehouwer, Mark Kroon, Petra de Vries, Eline P. Slagboom, Kai Ye, Joris A. Veltman, et Jayne Y. Hehir-Kwa. 2014. « Mobster: accurate detection of mobile element insertions in next generation sequencing data ». *Genome Biology* 15 (10): 488. <https://doi.org/10.1186/s13059-014-0488-x>.
- Torene, Rebecca I., Kevin Galens, Shuxi Liu, Kevin Arvai, Carlos Borroto, Julie Scuffins, Zhancheng Zhang, et al. 2020. « Mobile element insertion detection in 89,874 clinical exomes ». *Genetics in Medicine*, janvier. <https://doi.org/10.1038/s41436-020-0749-x>.
- Uguen, Kévin, Claire Jubin, Yannis Duffourd, Claire Bardel, Valérie Malan, Jean-Michel Dupont, Laila El Khattabi, et al. 2020. « Genome Sequencing in Cytogenetics: Comparison of Short-Read and Linked-Read Approaches for Germline Structural Variant Detection and Characterization ». *Molecular Genetics & Genomic Medicine* 8 (3): e1114-e1114. <https://doi.org/10.1002/mgg3.1114>.
- Vogt, Julia, Kathrin Bengesser, Kathleen BM Claes, Katharina Wimmer, Victor-Felix Mautner, Rick van Minkelen, Eric Legius, et al. 2014. « SVA retrotransposon insertion-associated deletion represents a novel mutational mechanism underlying large genomic copy number changes with non-recurrent breakpoints ». *Genome Biology* 15 (6): R80. <https://doi.org/10.1186/gb-2014-15-6-r80>.

- Wang, Hui, Jinchuan Xing, Deepak Grover, Dale J. Hedges, Kyudong Han, Jerilyn A. Walker, et Mark A. Batzer. 2005. « SVA Elements: A Hominid-Specific Retroposon Family ». *Journal of Molecular Biology* 354 (4): 994-1007. <https://doi.org/10.1016/j.jmb.2005.09.085>.
- Wang, Jianxin, Lei Song, Deepak Grover, Sami Azrak, Mark A Batzer, et Ping Liang. 2006. « DbRIP: A Highly Integrated Database of Retrotransposon Insertion Polymorphisms in Humans ». *Human Mutation* 27 (4): 323-29. <https://doi.org/10.1002/humu.20307>.
- Waters, Aoife M., et Philip L. Beales. 2011. « Ciliopathies: An Expanding Disease Spectrum ». *Pediatric Nephrology (Berlin, Germany)* 26 (7): 1039-56. <https://doi.org/10.1007/s00467-010-1731-7>.
- Williams, Amy L., Nick Patterson, Joseph Glessner, Hakon Hakonarson, et David Reich. 2012. « Phasing of Many Thousands of Genotyped Samples ». *The American Journal of Human Genetics* 91 (2): 238-51. <https://doi.org/10.1016/j.ajhg.2012.06.013>.
- Wilund, Kenneth R., Ming Yi, Filomena Campagna, Marcello Arca, Giovanni Zuliani, Renato Fellin, Yiu-Keo Ho, J. Victor Garcia, Helen H. Hobbs, et Jonathan C. Cohen. 2002. « Molecular Mechanisms of Autosomal Recessive Hypercholesterolemia ». *Human Molecular Genetics* 11 (24): 3019-30. <https://doi.org/10.1093/hmg/11.24.3019>.
- Wormser, Ohad, Libe Gradstein, Yuval Yogev, Yonatan Perez, Rotem Kadir, Inna Goliand, Yair Sadka, et al. 2019. « SCAPER localizes to primary cilia and its mutation affects cilia length, causing Bardet-Biedl syndrome ». *European Journal of Human Genetics* 27 (6): 928-40. <https://doi.org/10.1038/s41431-019-0347-z>.
- Yıldız Bölükbaşı, Esra, Sara Mumtaz, Muhammad Afzal, Ute Woehlbier, Sajid Malik, et Aslihan Tolun. 2018. « Homozygous mutation in *CEP19*, a gene mutated in morbid obesity, in Bardet-Biedl syndrome with predominant postaxial polydactyly ». *Journal of Medical Genetics* 55 (3): 189. <https://doi.org/10.1136/jmedgenet-2017-104758>.
- Zaghloul, Norann A., et Nicholas Katsanis. 2010. « Functional modules, mutational load and human genetic disease ». *Trends in Genetics* 26 (4): 168-76. <https://doi.org/10.1016/j.tig.2010.01.006>.
- Zerbino, Daniel R., et Ewan Birney. 2008. « Velvet: Algorithms for de Novo Short Read Assembly Using de Bruijn Graphs. » *Genome Research* 18 (5): 821-29. <https://doi.org/10.1101/gr.074492.107>.

**SUPPLEMENTARY DATA**

Family	A	B		C		D	E	F	G		H	I	
Patient	A.II-3	B.II-1	B.II-3	B.II-4	C.II-1	C.II-2	D.II-2	E.II-3	F.II-1	G.II-3	G.II-4	H.II.1	I.II-1
<b>Sensorineural features</b>													
Retinal dystrophy	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Nystagmus		Y	Y	Y				Y					
<b>Abnormalities of the hands and feet</b>													
Postaxial polydactyly	Y	Y	Y	Y			Y	Y	Y	Y	N	Y	Y
-Hands		Y					Y		Y				Y
-Feet	Y		Y	Y				Y					Y
Brachydactyly		Y	Y	Y					Y				
Clinodactyly							Y						
<b>Body mass index</b>													
Obesity (BMI)		Y (47.3)	Y (31.1)	Y (43.3)	Y	Y	Y (29.3)	Y	Y (38.5)	Y	Y	Y	Y
<b>Development psychomotor</b>													
Intellectual disability	mild	mild	mild	Y						Y	Y	N	Moderate
Learning difficulty	Y	Y	Y	Y				Y				N	Y
Speech delay									Y			N	Severe
Slow ideation	Y	Y					Y		Y			N	Y
Motor development delay		Y	Y	Y					Y			N	Y
<b>Renal and genital abnormalities</b>													
Micropenis		Y	Y	Y									Y
Cryptorchidism		Y											N
Urogenital anomalies							Y			Y	Y	Y	Ptosis right kidney
<b>Other clinical features</b>													
		Facial dysmorphism (little mouth)			Cervical/rectal cancer		Hepatic steatosis, lumbar stenosis, iron deficiency	T2D, heart murmur	Asthma, hypothyroidism, scoliosis		Macrocephaly		Right pes cavus

**Supplementary Table 1.** Clinical data from BBS patients with pathogenic variants. Genotypes: families A to F are *BBS1*:c.[1169T>G];[1214\_1215ins[MT113356]] p.[Met390Arg];[Ala406Glnfs\*47] while family G and H carrying in addition to the *BBS1*:c. 1214\_1215ins[MT113356] p.Ala406Glnfs\*47 respectively *BBS10* pathogenic variants c.[271dupT];[145C>T], p.[Cys91Leufs\*5;];[Arg49Trp] or *BBS2*:c.[72C>T];[175C>T], p.[Tyr24\*];[Gln59\*] and *MKKS/BBS6*:c.442C>T, p.Gln148\*. Family I carries *BBS1*:c.[1169T>G];[160-?\_1110+?del], p.[M390R];[(Leu54\_Pro370del)]. T2D: Type 2 diabetes. Y, yes; N, No, a blank indicates data not recorded.

Gene name (HGNC)	BBS	NPHP	MKS	Joubert	Senior Løken	Alström
<i>BBS1</i>	<i>BBS1</i>					
<i>BBS2</i>	<i>BBS2</i>					
<i>ARL6</i>	<i>BBS3</i>					
<i>BBS4</i>	<i>BBS4</i>					
<i>BBS5</i>	<i>BBS5</i>					
<i>MKKS</i>	<i>BBS6</i>					
<i>BBS7</i>	<i>BBS7</i>					
<i>TTC8</i>	<i>BBS8</i>					
<i>BBS9</i>	<i>BBS9</i>					
<i>BBS10</i>	<i>BBS10</i>					
<i>TRIM32</i>	<i>BBS11</i>					
<i>BBS12</i>	<i>BBS12</i>					
<i>MKS1</i>	<i>BBS13</i>		<i>MKS1</i>	<i>JBTS27</i>		
<i>CEP290</i>	<i>BBS14</i>	<i>NPHP6</i>	<i>MKS4</i>	<i>JBTS5</i>	<i>SLSN6</i>	
<i>WDPCP</i>	<i>BBS15</i>					
<i>SDCCAG8</i>	<i>BBS16</i>	<i>NPHP10</i>			<i>SLSN7</i>	
<i>LZTFL1</i>	<i>BBS17</i>					
<i>BBIP1</i>	<i>BBS18</i>					
<i>IFT27</i>	<i>BBS19</i>					
<i>IFT172</i>	<i>BBS20</i>	<i>NPHP17</i>				
<i>ALMS1</i>						<i>ALMS1</i>
<i>NPHP1</i>		<i>NPHP1</i>		<i>JBTS4</i>	<i>SLSN1</i>	
<i>INVS</i>		<i>NPHP2</i>				
<i>NPHP3</i>		<i>NPHP3</i>	<i>MKS7</i>		<i>SLSN3</i>	
<i>NPHP4</i>		<i>NPHP4</i>			<i>SLSN4</i>	
<i>IQCB1</i>		<i>NPHP5</i>			<i>SLSN5</i>	
<i>GLIS2</i>		<i>NPHP7</i>				
<i>RPGRIP1L</i>		<i>NPHP8</i>	<i>MKS5</i>	<i>JBTS7</i>		
<i>NEK8</i>		<i>NPHP9</i>				
<i>TMEM67</i>		<i>NPHP11</i>	<i>MKS3</i>	<i>JBTS6</i>		

<i>TTC21B</i>	<i>NPHP12</i>	<i>JBTS11</i>	
<i>TMEM216</i>		<i>MKS2</i>	<i>JBTS2</i>
<i>AHI1</i>			<i>JBTS3</i>
<i>CCDC28B</i>			
<i>CC2D2A</i>		<i>MKS6</i>	<i>JBTS9</i>
<i>TCTN2</i>		<i>MKS8</i>	
<i>B9D1</i>		<i>MKS9</i>	<i>JBTS27</i>
<i>B9D2</i>		<i>MKS10</i>	
<i>TMEM231</i>		<i>MKS11</i>	<i>JBTS20</i>
<i>XPNPEP3</i>	<i>NPHPL1</i>		
<i>ZNF423</i>	<i>NPHP14</i>		<i>JBTS19</i>
<i>WDR19</i>	<i>NPHP13</i>		
<i>CEP164</i>	<i>NPHP15</i>		<i>SLSN8?</i>
<i>ANKS6</i>	<i>NPHP16</i>		
<i>INPP5E</i>			<i>JBTS1</i>
<i>ARL13B</i>			<i>JBTS8</i>
<i>OFD1</i>			<i>JBTS10</i>
<i>KIF7</i>			<i>JBTS12</i>
<i>TCTN1</i>			<i>JBTS13</i>
<i>TMEM237</i>			<i>JBTS14</i>
<i>CEP41</i>			<i>JBTS15</i>
<i>TMEM238</i>			<i>JBTS16</i>
<i>C5orf42</i>			<i>JBTS17</i>
<i>TCTN3</i>			<i>JBTS18</i>
<i>CSPP1</i>			<i>JBTS21</i>
<i>PDE6D</i>			<i>JBTS22</i>
<i>DCDC2</i>	<i>NPHP19</i>		
<i>CCDC41</i>	<i>NPHP18</i>		

**Supplementary Table 2.** Official genes names (and aliases) covered by the updated set of 58 BBS-related genes screened by targeted exome sequencing. BBS: Bardet-Biedl syndrome, NPHP: Nephronophthisis, MKS: Meckel syndrome, JBTS: Joubert syndrome

Application	Gene	Name	Exon	Forward (5'-3')	Name	Exon	Reverse (5'-3')	Size (bp)
Primers for Sanger on gDNA								
Duplex PCR for SVA detection	<i>BBS1</i>	1F	Int12	CACAGTACTCCACAAATAACTGCT	1R	SVA_F	CAGCCTGGGCACCATTGA	270
					2R	Int13	ATTCCCCAGCTTTGCTGT	440
M390R	<i>BBS1</i>	A	Int11	CCATCCCCTGTCTTGCTTTC	B	Int12	CATGCTTCATTCCACCTCC	189
SVA-F sequence determination	<i>BBS1</i>	F0	TSD	CATCAAGATCCTGAAGCGTACTC	R3	SINE-R	TCTTGGGTGTTTCTCGCAGA	~2300
		F4	VNTR	CTCTGCGAGAAACACCCAAG	R5	Ex13	GTCTTTGCGGGCACATTGAG	~250
		F9	SINE-R	AGTACCCAGGGACAACACT				~350
		F6	SINE-R	CAGCAGCCACCCCATCTG	R2	SINE-R	GCACCGCCCTTAATCCATT	
Primers for Sanger on cDNA								
M390R	<i>BBS1</i>	F10-11	10-11	CCCACAAGGGGAAGAAGC	R13	13	TCAGGATCTTGATGATCAGG	264
					R14	14	AGCACGTAGGCGCAGCA	446
Primers for qPCR on gDNA								
qPCR	<i>BBS1</i>	8F	8	TTCTCCCTCATGTGGCATTCTGG	8R	8	AGCATCCTCGTCAGCCAGGTTCC	137
		9F	9	CCCCGTCTCCTAGAGGTTTCTG	9R	9	GCCACCAGTTTTCTGAGATGTTCC	145
	<i>RPPH1</i>		1	TCTGGGAAATCACCATAAACGTGAA		1	GTGGCCCACTGATGAGCTT	110
	<i>GAPDH</i>		Int1	CTCCCACACATGCACTTA		Int1	CCTAGTCCCAGGGCTTGATT	99
Primers for RT qPCR on cDNA								
mRNA quantification	<i>BBS1</i>	4F	4	TGGAACCAGGCCAAAGAG	6R	6	AAGGCTCCTCTGCCGTCT	85
	<i>GAPDH</i>		4-5	GGAGCGAGATCCCTCAAAT		6	GGCTGTGCATACTTCTCATGG	197
	<i>HPRT</i>		1-2	CCTGGCGTCTGTATTAGTAT		2-3	AGACGTTCACTCTGTCCATAA	131

**Supplementary Table 3.** List and characteristics of *BBS1* primers used in this study.

Chr	Position	rsID	A.I-1		A.II-3		B.II-1		C.II-3		D.I-1		D.II-2		Control 1		Control 2		Control 3		
			A1	A2	A1	A2	A1	A2	A1	A2	A1	A2	A1	A2	A1	A2	A1	A2	A1	A2	A1
11	65564786	rs44205	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G	G	G	G	
11	65564987	rs12800057	T	T	T	T	G	G	T	G	T	G	T	G	T	G	T	T	T	G	T
11	65570405	rs7116758	T	T	T	T	C	C	T	C	T	C	T	C	T	C	T	T	T	C	T
11	65570517	rs7119612	T	T	T	C	T	C	C	C	C	T	C	T	T	T	T	T	T	C	T
11	65577516	rs10896050	G	G	G	G	T	G	G	T	G	G	G	G	G	G	G	G	G	G	G
11	65578672	rs11227306	A	A	A	A	A	C	A	A	C	A	C	A	A	A	A	A	A	A	A
11	65580638	rs12576996	T	T	T	T	G	T	T	T	T	T	C	T	G	T	T	T	T	T	T
11	65585257	rs11227311	C	C	C	C	C	A	C	C	C	C	C	A	C	C	C	C	C	C	C
11	65590713	rs601887	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G
11	65594820	rs7102454	G	G	G	G	A	G	G	G	A	G	A	G	A	G	G	G	G	G	G
11	65617324	rs654638	A	A	A	A	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A
11	65621995	rs635375	A	A	A	A	G	A	A	G	A	G	A	A	A	A	A	A	A	A	A
11	65623739	rs11227332	T	T	T	T	T	C	T	C	T	T	T	T	T	T	T	T	T	T	T
11	65626701	rs1939212	C	C	C	C	T	C	C	T	T	C	C	C	C	C	C	C	C	C	C
11	65628334	rs13817	T	T	T	T	T	C	T	C	T	C	T	T	T	T	T	T	T	T	T
11	65629932	rs659857	C	C	C	C	C	T	C	T	T	C	C	C	C	C	C	C	C	C	C
11	65631167	rs648732	T	T	T	T	T	C	C	T	C	T	T	T	T	T	T	T	T	T	T
11	65631973	rs630303	A	A	A	A	G	A	A	G	A	G	A	A	A	A	A	A	A	A	A
11	65632262	rs558114	T	T	T	T	T	G	T	G	T	G	T	T	T	T	T	T	T	T	T
11	65637076	rs570387	T	T	T	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T
11	65637273	rs501630	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
11	65640906	rs7940691	G	G	G	G	G	G	G	G	A	G	A	G	A	G	G	G	G	G	G
11	65644027	rs583887	C	C	C	C	C	T	C	T	C	C	C	C	C	C	C	C	C	C	C
11	65649984	rs687672	C	C	C	C	T	C	T	C	C	C	C	C	C	C	C	C	C	C	C
11	65653242	rs568617	G	G	G	G	A	G	A	G	G	G	G	G	G	G	G	G	G	G	G
11	65656282	rs656980	C	C	C	C	C	T	C	T	C	C	C	C	C	C	C	C	C	C	C
11	65656564	rs2231884	G	G	G	G	G	A	A	G	G	G	G	G	G	G	G	G	G	G	G
11	65661291	rs10791830	G	G	G	G	G	A	G	A	G	G	G	G	G	G	G	G	G	G	G
11	65663239	rs622614	C	C	C	C	T	C	T	C	C	C	C	C	C	C	C	C	C	C	C
11	65664006	rs614520	C	C	C	C	C	C	C	C	T	C	T	C	C	C	C	C	C	C	C
11	65665200	rs1151523	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
11	65676516	rs653914	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
11	65705108	rs528736	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G
11	65712413	rs624273	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
11	65724291	rs470192	A	A	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	A
11	65727301	rs491973	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G



11	65732651	rs550435	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	T
11	65732800	rs552130	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
11	65733393	rs688862	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
11	65746653	rs679581	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
11	65747057	rs754532	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
11	65749337	rs735942	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
11	65754808	rs1786175	A	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A
11	65756283	rs1151511	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
11	65758854	rs1151514	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
11	65762654	rs12574568	A	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A
11	65765551	rs610497	T	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T
11	65768093	rs3825068	A	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A
11	65783915	rs668210	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
11	65790527	rs2845570	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
11	65793454	rs1203998	A	A	A	A	G	A	A	A	A	A	A	A	A	A	A	A	A
11	65793797	rs1893316	T	T	T	T	C	T	T	T	C	T	T	T	T	T	T	T	T
11	65793942	rs947847	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C
11	65808467	rs2576	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C
11	65817307	rs574153	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G	G	G
11	65817592	rs571374	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
11	65819551	rs12803276	T	T	T	T	T	C	T	T	C	T	T	T	T	T	T	T	T
11	65819661	rs2241303	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
11	65820673	rs503156	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C
11	65824472	rs547404	T	T	T	T	T	C	T	T	T	T	T	T	T	T	T	T	T
11	65851902	rs472518	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G
11	65866187	rs537497	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C
11	65876570	rs873491	A	A	A	A	G	A	A	A	G	A	A	A	A	A	A	A	A
11	65886662	rs524281	T	T	T	T	T	T	T	T	G	T	T	T	T	T	T	T	T
11	65910129	rs540943	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C
11	65922589	rs801739	T	T	T	T	C	T	C	C	C	T	T	C	T	T	T	T	T
11	65934549	rs801733	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
11	65935189	rs513474	T	T	T	T	C	T	C	C	C	T	T	C	T	T	T	T	T
11	65943647	rs7123153	A	A	A	A	G	A	G	G	A	A	A	A	A	A	A	A	A
11	65957003	rs17307346	A	A	A	A	C	A	C	C	A	A	A	A	A	A	A	A	A
11	65985168	rs512421	T	T	T	T	C	C	C	C	C	T	C	T	T	T	T	T	T
11	65992008	rs9326370	A	A	A	A	G	A	G	G	A	A	A	A	A	A	A	A	A
11	65998757	rs559298	G	G	G	G	A	A	A	A	A	G	A	A	G	G	G	G	G
11	66001869	rs10896104	C	C	C	C	T	T	C	T	C	C	T	C	C	C	C	C	C

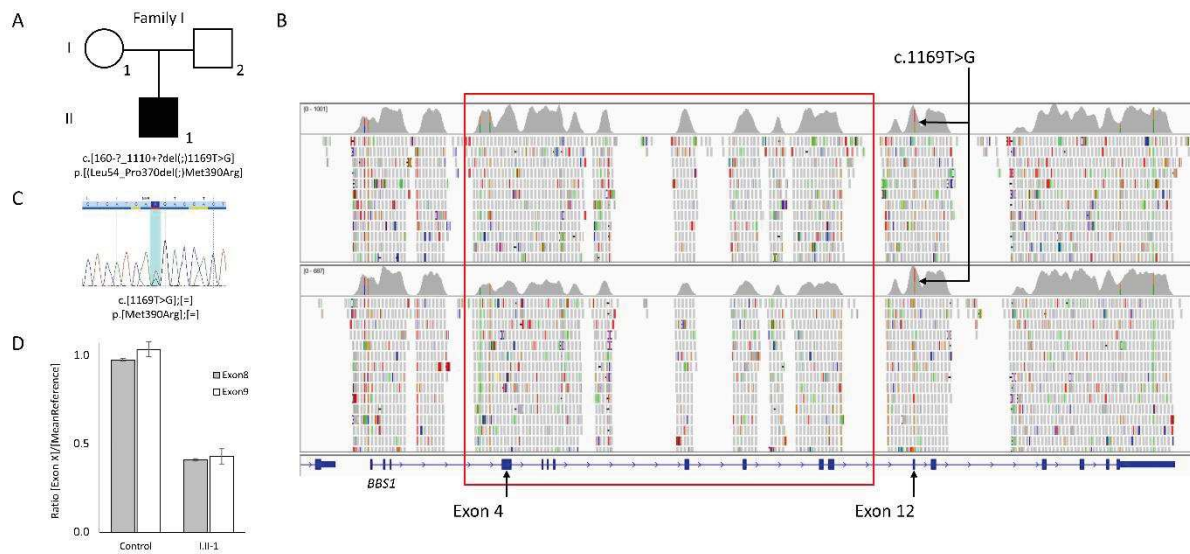
11	66011280	rs2298466	C	C	C	T	T	T	T	T	C	C	T	C	C	C	C	C	C	C	C	C
11	66031660	rs1151539	C	C	C	T	T	C	T	T	C	C	T	C	C	C	C	C	C	C	C	C
11	66035608	rs10791865	C	C	C	C	C	A	C	C	A	C	C	C	C	C	C	C	C	C	C	C
11	66041807	rs1151540	A	A	A	C	A	C	C	C	A	A	C	A	A	A	A	A	A	A	A	A
11	66050712	rs4073582	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
11	66058082	rs556595	A	A	A	C	C	A	C	C	A	A	C	A	A	A	A	A	A	A	A	A
11	66061233	rs484261	A	A	A	A	A	C	A	A	C	A	C	A	A	A	A	A	A	A	A	A
11	66062023	rs34837704	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
11	66064126	rs560577	T	T	T	T	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T
11	66064850	rs531784	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
11	66066349	rs580241	T	T	T	C	T	C	T	C	T	T	C	T	T	T	T	T	T	T	T	T
11	66066664	rs4930351	G	G	G	A	A	A	A	A	G	G	A	G	G	G	G	G	G	G	G	G
11	66070146	rs523583	C	C	C	A	A	A	A	A	A	C	A	C	C	C	C	C	C	C	C	C
11	66077203	rs4390351	T	T	T	T	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T
11	66080770	rs7924580	G	G	G	A	G	A	A	A	G	G	A	G	G	G	G	G	G	G	G	G
11	66083129	rs3741367	G	G	G	A	A	A	A	A	A	G	A	G	G	G	G	G	G	G	G	G
11	66083782	rs3741368	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
11	66090032	rs10219183	G	G	G	T	G	T	T	T	G	G	T	G	G	G	G	G	G	G	G	G
11	66091814	rs3016311	A	A	A	A	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A
11	66092730	rs1785631	T	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	T	T	T
11	66093612	rs1892938	T	T	T	G	G	G	G	G	T	T	G	T	T	T	T	T	T	T	T	T
11	66106725	rs948360	G	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G
11	66108660	rs947978	A	A	A	G	G	G	G	G	A	A	G	A	A	A	A	A	A	A	A	A
11	66109033	rs11537993	C	C	C	T	T	T	T	T	T	C	T	C	C	C	C	C	C	C	C	C
11	66114900	rs11537930	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C
11	66129442	rs4930178	G	G	G	C	C	C	C	C	C	G	G	C	C	G	G	G	G	G	G	G
11	66133336	rs2279861	C	C	C	T	T	T	T	T	C	C	T	C	C	C	C	C	C	C	C	C
11	66135179	rs2279862	A	A	A	G	G	G	G	G	A	A	G	A	A	A	A	A	A	A	A	A
11	66171621	rs6591216	T	T	T	G	G	G	G	G	G	T	G	T	T	T	T	T	T	T	T	T
11	66177952	rs2282530	G	G	G	A	A	A	A	A	G	G	A	G	G	G	G	G	G	G	G	G
11	66179063	rs12785651	C	C	C	T	T	T	T	T	C	C	T	C	C	C	C	C	C	C	C	C
11	66179556	rs12786563	G	G	G	A	A	A	A	A	G	G	A	G	G	G	G	G	G	G	G	G
11	66182417	rs12577917	T	T	T	C	C	C	C	C	T	T	C	T	T	T	T	T	T	T	T	T
11	66184140	rs2282529	G	G	G	A	A	A	G	A	G	G	A	G	G	G	G	G	G	G	G	G
11	66186882	rs7947391	T	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	T	T	T
11	66196265	rs17495838	T	T	T	T	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T
11	66203025	rs2279865	A	A	A	A	A	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A
11	66234023	rs7925108	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C

11	66240882	rs2277302	A	A	A	A	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A
11	66245851	rs2282640	A	A	A	A	A	C	A	A	A	C	A	A	A	A	A	A	A	A	A	A
11	66246337	rs11821155	A	A	A	A	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A
11	66254085	rs11550299	C	C	C	C	C	A	C	C	C	C	C	C	C	C	C	C	C	C	C	C
11	66262606	rs2511224	A	A	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	A	A
11	66272142	rs1671063	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C
11	66272237	rs2305535	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
11	66272976	rs10501397	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
11	66282095	rs2298806	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
11	66288123	rs11822028	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
11	66293652	BBS	Other	M390R	M390R	SVA	SVA	M390R	Other	SVA	Other	M390R	SVA	M390R	M390R	M390R	M390R	M390R	M390R	M390R	M390R	
11	66297363	rs3816492	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
11	66299515	rs8432	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C
11	66306355	rs3737525	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
11	66311011	rs3825065	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
11	66313026	rs11227527	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
11	66316557	rs597626	T	T	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	T	T
11	66317823	rs1791690	A	A	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	A	A
11	66319262	rs10791881	A	A	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	A	A
11	66324762	rs677488	G	G	G	A	G	A	A	G	G	G	G	A	G	G	G	G	G	G	G	G
11	66326581	rs2275998	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
11	66328095	rs1815739	A	A	A	A	G	A	G	G	A	A	A	A	G	A	A	A	A	A	A	A
11	66329732	rs540874	T	T	T	C	T	C	C	T	T	T	T	C	T	T	T	T	T	T	T	T
11	66331000	rs4576	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
11	66331458	rs572846	T	T	T	C	C	T	C	T	T	T	T	C	T	T	T	T	T	T	T	T
11	66333072	rs4630309	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
11	66348503	rs692892	A	A	A	A	A	G	A	A	G	A	A	A	A	A	A	A	A	A	A	A
11	66358126	rs486584	A	A	A	A	A	A	G	A	A	G	A	A	A	A	A	A	A	A	A	A
11	66369211	rs1791684	C	C	C	C	C	A	C	C	C	C	A	C	C	C	C	C	C	C	C	C
11	66381517	rs12795774	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C
11	66394387	rs15969	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
11	66399681	rs484262	A	A	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	A	A
11	66401373	rs566673	G	G	G	G	G	T	G	G	T	G	G	A	G	G	G	G	G	G	G	G
11	66409409	rs519380	G	G	G	G	G	A	G	G	T	G	G	G	G	G	G	G	G	G	G	G
11	66419554	rs481296	T	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	T	T	T
11	66438074	rs12577111	C	C	C	C	T	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C
11	66451178	rs3741364	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C
11	66468469	rs506028	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A

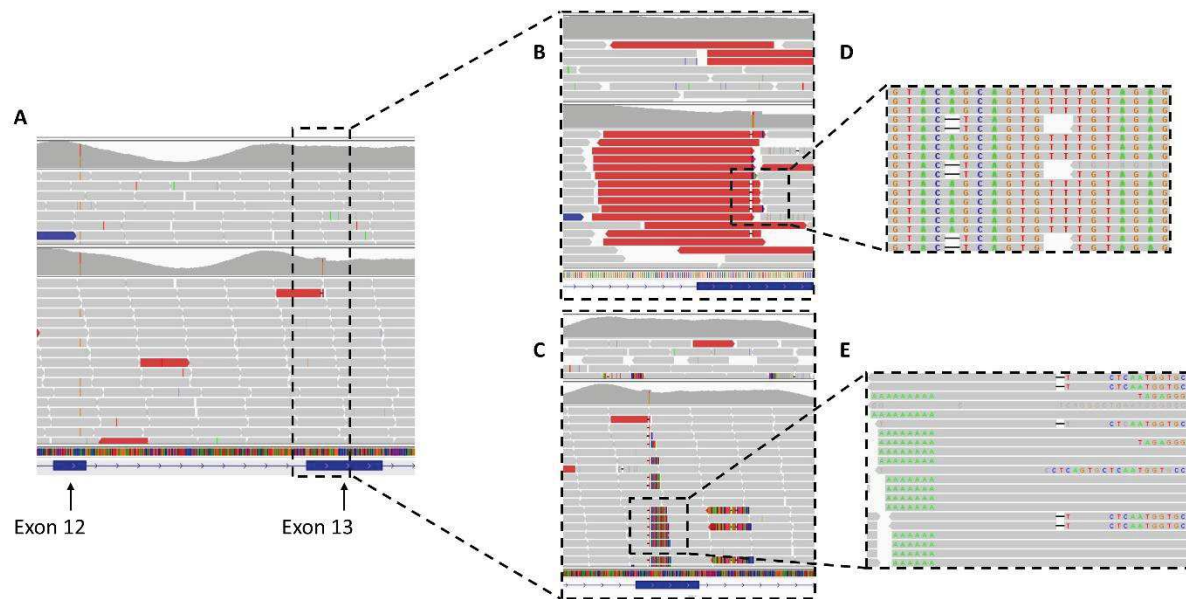
11	66469032	rs532439	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C
11	66493825	rs10791889	C	C	C	T	C	T	C	T	C	T	C	T	C	T	C	C	C
11	66560023	rs4930390	T	T	T	C	T	C	T	C	C	T	C	T	C	T	T	T	T
11	66568943	rs682842	G	G	G	A	A	G	G	A	A	G	A	G	A	G	G	G	G
11	66579572	rs11601241	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
11	66582210	rs10750790	A	A	A	G	A	G	A	G	G	A	G	A	G	A	A	A	A
11	66608521	rs1201342	G	G	G	G	C	G	G	G	G	G	G	G	G	G	G	G	G
11	66608991	rs594313	A	A	A	A	G	A	G	A	G	A	G	A	G	A	A	A	A
11	66621383	rs2077432	G	G	G	A	G	A	A	A	A	A	G	A	G	A	G	G	G
11	66649612	rs7119426	T	T	T	C	T	C	C	C	C	T	C	T	C	T	T	T	T
11	66650060	rs10896139	C	C	C	C	T	C	T	T	T	T	C	T	C	T	C	C	C
11	66662731	rs7122539	G	G	G	A	G	A	A	A	A	A	G	A	G	A	G	G	G
11	66673079	rs10160758	A	A	A	A	G	A	G	G	G	G	A	C	A	A	A	A	A
11	66675990	rs6591226	A	A	A	A	C	A	C	C	C	C	A	C	A	A	A	A	A
11	66676273	rs11227623	C	C	C	T	C	T	T	T	T	C	C	T	C	C	C	C	C
11	66676995	rs735498	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
11	66679738	rs10791894	A	A	A	A	G	A	G	G	G	G	A	C	A	A	A	A	A
11	66687863	rs7944999	T	T	T	C	T	C	C	C	C	T	C	G	T	T	T	T	T
11	66690454	rs7110302	T	T	T	C	T	C	C	C	C	T	C	T	C	T	T	T	T
11	66744119	rs4930187	C	C	C	T	C	T	C	T	C	C	C	T	C	C	C	C	C
11	66766754	rs12419173	T	T	T	G	T	G	T	G	T	G	T	G	T	G	T	T	T
11	66781263	rs896436	T	T	T	C	T	C	T	C	T	C	C	C	C	T	T	T	T
11	66782594	rs1253440	G	G	G	T	G	T	G	T	G	T	T	T	T	G	G	G	G
11	66784715	rs1253439	G	G	G	A	G	A	G	A	G	A	A	A	A	A	G	G	G
11	66822639	rs17148051	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
11	66826160	rs7112925	A	A	A	G	A	G	G	A	A	G	G	G	G	A	A	A	A
11	66832528	rs11227673	A	A	A	A	G	A	G	G	A	A	G	G	G	A	A	A	A
11	66843350	rs7932636	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	T
11	66853465	rs10791899	T	T	T	T	T	T	T	G	T	T	T	T	T	T	T	T	T
11	66889558	rs4478996	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C
11	66963425	rs674499	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C
11	67011765	rs12288048	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	A
11	67031968	rs12791853	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G
11	67033076	rs11605263	G	G	G	A	A	G	A	G	A	G	G	A	G	A	G	G	G
11	67034266	rs11227756	A	A	A	A	A	A	A	C	A	A	A	A	A	A	A	A	A
11	67050292	rs2071007	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	A	A
11	67065889	rs10896167	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
11	67068859	rs2298815	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	G

11	67070017	rs4930192	T	T	T	<b>T</b>	<b>T</b>	T	T	T	<b>C</b>	T	T	T	T	T	T	T
11	67070766	rs4255546	A	A	A	<b>A</b>	<b>A</b>	A	A	G	<b>A</b>	A	A	<b>A</b>	<b>A</b>	A	A	A
11	67076810	rs3737460	G	G	G	<b>G</b>	<b>G</b>	G	G	T	<b>G</b>	G	G	<b>G</b>	<b>G</b>	G	G	G
11	67088394	rs1894112	G	G	G	<b>G</b>	<b>G</b>	G	A	<b>G</b>	G	G	<b>G</b>	<b>G</b>	G	G	G	G
11	67124333	rs1638566	A	A	A	<b>A</b>	<b>A</b>	A	A	<b>C</b>	A	A	<b>A</b>	<b>A</b>	A	A	A	A
11	67125223	rs1638567	A	A	A	<b>A</b>	<b>A</b>	A	A	G	<b>A</b>	A	<b>A</b>	<b>A</b>	A	A	A	A
11	67130122	rs1790740	T	T	T	<b>T</b>	<b>T</b>	T	T	T	<b>C</b>	T	T	<b>T</b>	<b>T</b>	T	T	T
11	67137727	rs12788752	A	A	A	<b>A</b>	<b>A</b>	A	A	G	<b>A</b>	A	<b>A</b>	<b>A</b>	A	A	A	A
11	67147422	rs1638571	A	A	A	<b>A</b>	<b>A</b>	A	A	A	<b>C</b>	A	<b>A</b>	<b>A</b>	A	A	A	A
11	67171736	rs2514256	G	G	G	<b>G</b>	<b>G</b>	G	A	<b>G</b>	G	G	<b>G</b>	<b>G</b>	G	G	G	G
11	67173240	rs12805811	A	A	A	<b>A</b>	<b>A</b>	A	A	A	<b>G</b>	A	<b>A</b>	<b>A</b>	A	A	A	A
11	67174597	rs1638586	G	G	G	<b>G</b>	<b>G</b>	G	G	<b>A</b>	G	G	<b>G</b>	<b>G</b>	G	G	G	G
11	67175692	rs10896172	A	A	A	<b>A</b>	<b>A</b>	A	A	G	<b>A</b>	A	<b>A</b>	<b>A</b>	A	A	A	A
11	67188860	rs12290164	A	A	A	<b>A</b>	<b>A</b>	A	A	A	<b>G</b>	A	<b>A</b>	<b>A</b>	A	A	A	A
11	67191449	rs1790747	A	A	A	<b>A</b>	<b>A</b>	A	A	G	<b>A</b>	A	<b>A</b>	<b>A</b>	A	A	A	A
11	67192555	rs1626067	G	G	G	<b>A</b>	<b>G</b>	A	A	A	<b>A</b>	G	<b>A</b>	<b>A</b>	A	A	A	A
11	67196237	rs1476792	T	T	T	<b>C</b>	<b>T</b>	C	C	<b>C</b>	C	T	<b>C</b>	<b>C</b>	C	C	C	T
11	67202044	rs1790752	G	G	G	<b>A</b>	<b>G</b>	A	A	A	<b>A</b>	G	<b>A</b>	<b>A</b>	A	A	A	G
11	67202826	rs10274	C	C	C	<b>T</b>	<b>C</b>	T	T	T	<b>T</b>	C	<b>T</b>	<b>T</b>	T	T	T	C

**Supplementary Table 4.** Haplotype analysis for ~1.6Mb around *BBS1* (between rs44205 and rs10274) on chromosome 11. In total, 9 individuals including 3 affected and 3 non affected from 4 families carriers of the SVA F were analysed (family A, B, C and D) together with 3 affected “control” samples (homozygous for the c.1169T>G, p.M390R). After phasing each allele is represented in each column and alleles carriers of the SVA F are highlighted in bold (A.II-3, B.II-1, C.II-3 and D.II-2). Defined using 3 independent “control” *BBS1* patients homozygous for the c.1169T>G (p.M390R), the shared “M390R” haplotype is highlighted in green and spans ~1.614 Mb, *BBS1* region is highlighted in grey and the shared “SVA F” haplotype is highlighted in orange and spans ~56.6kb.

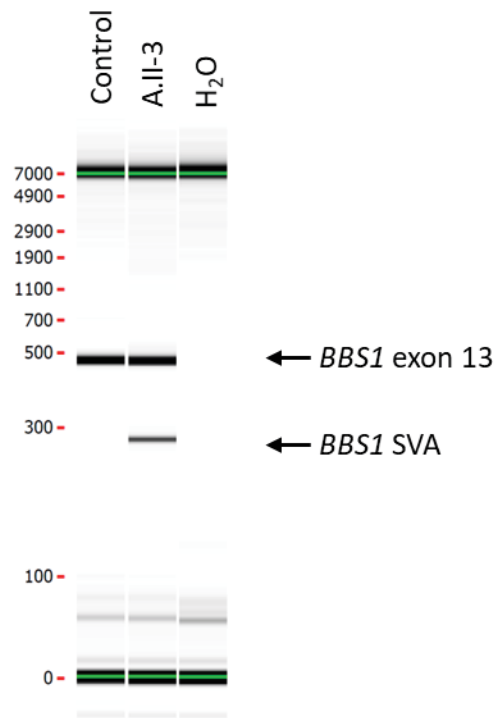


**Supplementary Figure 1.** (A) Pedigree of family I with 2 variants in *BBS1*. No parental segregation could be performed. (B) Next generation sequencing data surrounding the *BBS1* locus from patient A.II-3 (upper panel) and patient I.II-1 (lower panel) harboring a heterozygous deletion of exon 4 to 11 (c.160-?\_1110+?del). Both patients are also heterozygous for the *BBS1* c.1169T>G (p.M390R) pathogenic variant. (C) Sanger sequences from *BBS1* exon 12 showing the c.1169T>G at the heterozygous state in patient I.II-1. (D) Quantitative real-time PCR from exons 8 and 9 in *BBS1* from I.II-1 and one unrelated control. DNA quantity from the corresponding exons were compared with two reference genes (*RPPH1* and *GAPDH*) using the absolute quantification method. Ratios can be interpreted as following: homozygous deletion (ratio < 0.1), heterozygous deletion (0.3–0.7), wild type (0.8–1.2), duplication (1.3–1.6), and triplication (>0.8). Error bars represents standard deviation.



**Supplementary Figure 2.** (A) Next generation sequencing data displayed from the Integrative Genomics Viewer (Thorvaldsdottir, Robinson, et Mesirov 2013) surrounding the *BBS1* locus from patient A.II-3 (lower IGV track) and one control individual (upper IGV track) both harboring a heterozygous c.1169T>G (p.M390R) in exon 12. (B) Magnified view of the incorrectly called deletion insertion with the sorting and coloring of the reads according to the “insert size”. (C) Magnified view of the region of interest with the soft-clipped bases turned off revealing multiple reads with aberrant alignments on their right side and corresponding to an SVA F insertion. (D) The sequence reveals all reads ending at the exact same position with a deletion of AG and the insertions of a T (c.1215\_1216delinsT). (E) The spacing between the reads on the left (including the 3' end of the SVA F with the poly(A) tail) and the reads on the right (including the 5' end of the SVA F) corresponds to the Target Site Duplication (TSD).





**Supplementary Figure 3.** Duplex PCR profiles on a LabChip GX (Perkin Elmer) for a rapid detection of the SVA insertion (see Supplementary Table 3 for primer details). Wild type samples harbor a single band corresponding to *BBS1* exon 13 whereas heterozygous carriers for the SVA insertion show an additional band specific for the insertion breakpoint.

GTGGTGGCCTGATCATCAAGATCCTGAAGCGTACTCAGTGCTCAATGGTGCCAGGCTGGAGTGCAGT  
 GGCGTGATCTCGGCTCACTACAACCTACACCTCCAGCCGCTGCCTTGGCCTCCCAAAGTGCCGAGA  
 TTGCAGCCTCTGCCCAGCCGCCACCCCGTCTGGGAAGTGAGGAGCGTCTCTGCCTGGCCGCCATCGT  
 CTGGGATGTGAGGAGCCCCTCTGCCTGGCTGGCCAGTCTGGAAAGTGAGGAGCGTCTCCGCCCGGCCG  
CCATCCCATCTAGGAAGTAAGGAGCGCCTCTTCCAGCCGCCATCACATCTAGGAGTGAGGAGCGTC  
TCTGCCCGGCCGCCATCGTCTGAGATGTGGGGAGCGCCTCTGCCCGCCGCCCATCTGGGATGTGA  
 GGAGCGCCTCTGCCCAGCGAGACCCCGTCTGGGAGGTGAGGAGCGTCTCTGCCCGGCCGCCCGTCT  
 GAGAAGTGAGGAGACCCTCTGCCTGGCAACCACCCCGTCTGAGAAGTGAGGAGCCCCTCCGCCCGCA  
 GCCGCCCGGCTGAGAAGTGAGGAGCCTCTCCGCCAGCAGCCACCCCATCTGGGAAGTGAGGAGCGT  
 CTCCGCCCGGAGCCACCCCGTCCGGGAGGGAGGTGGGGGGGGTTCAGCCCCCGCCCGGCCAGCCC  
 CCTTATACGGGAGGGAGGTGGGGGGTTCAGCCCCCGCCCGGCCAGCCGTGCCGTCCGGGAGGGAGGT  
 GGGGGGGTTCAGCCCCCGCCTGGCCAGCCGTGCCGTCCGGGAGGGAGGTGGGGGGGTTCAGCCCCCGC  
 CCGGCCAGCCGCCCGTCCGGGAGGTGAGGGGCGCCTCTGCCCGGCTGCCCTACTGGGAAGTGAGGA  
 GCCCTCAGCCTGGCCAGCCACCCCGTCCGGGAGGGAGATGGGGGGGTTCAGCCCCCGCCCGGCCAG  
 CCGCCCCGTCCGGGAGGGAGGGAGGTGGGGGGGTTCAGCCCTCCCCCGGCCAGCCGCCCGTCTGGGA  
 GGTGAGGGGCGCCTCTGCCCGGCCGCCCTACTGGGAAGTGAGGAGCCCCTCTGCCCGGCCAGCCGCC  
 CCATCCGGGAGGGAGGTGGGGGGGTTCAGCCCCGTGCCCGGCCAGCCACCCGTCTGGGAGGGAGGT  
 GGGGGGTTCAGCCCTCCGCCCGGCCAGCCGCCCGTCTGGGAGGTGAGGGGCGCCTCTGCCCGGCCGCC  
 CCTACTGGGAAGTGAGGAGCCCCTCTGCCCGGCCAGCCGCCCTGTCTGGGAGGGAGGTGGGGGGTTCG  
 GCCCCCCGCCCGGCCAGCCGCCCATCCGGGAGGGAGGTGGGGGGGTTCGGCCCCGTGCCCGGCCAGCC  
 GCCCGTCCGGGAGGGAGGTGGGGGGGGGGTTCAGCCCCCGTCCCGGCCAGCCACCCCGTCCGGGAG  
 GTGAGGGGCGCCTCTGCCCGGCCGCCCTACTGGGAAGTGAGGAGCCCCTCTGCCCGGCCAGCCGCC  
 CGTCCGGGAGGGAGGTGGGGGGGGTTCAGCCCCCGCCAGCCAGCCGCCCGTCCGGGAGGGAGGTGG  
 GGGGGTTCAGCCCCCGCCCGGCCAGCCGCCCGTCCGGCAGGTGAGGGGCGCCTCTGCCCGGCCCGC  
 CCCTACTGGGAAGTGAGGAGCCCCTCTGCCAGCCAGCCGCCCGTCCGGGAGGGAGGTGGGGGGTTC  
 AGCCCCCGCCCGGCCAGCCGCCCGTCCGGGAGGTGAGGGGCGCCTCTGCCCGGCCGCCCTACTGG  
 GAAGTGAGGAGCCCCTCTGCCCGGCCACCGCCCCGTCTGGGAGGTGTGCCAGTAGCTCATTTGAGAAC  
 GGGCCAGGATGACAAATGGCGGCTTTGTGGAATAGAAAGCGGGAAAGGTGGGGAAAAGATTGAGAGAT  
 CGGATGGTTGCCGTGTCTGTGTGGAAGGAAGTAGACATGGGAGACTTTTTCATTTTGTCTGCACATAAG  
 AAAAATTCCTCTGCCCTGGGATCCTGTTGATCTGTGACCTTACCCCAACCCTGTGCTCTCTGAAACA  
 TGTGCTGTGTCCACTCAGGGTTAAATGGATTAAGGGCGGTGCAAGATGTGCTTTGTTAAACAGATGCT  
 TGAAGGCAGCATGCTCGTTAAGAGTCATCACCCTCCCTAATCTCAAGTACCCAGGGACAAACTGC  
 GGAAGGCCGAGGGACCTCTGCCTAGGAAAACCAGAGACCTTTGTTCACTTGTATCTGCTGACCTTC  
 CCTCCACTATTGTCCATGACCCTGCCTAATCCCCCTCTGCGAGAAACACCCAAGAATGAACAATAAA  
 AAAATAAAATAAATAAATAAATA [A]<sup>100</sup>AAGATCCTGAAGCGTACAGCAGTGTTTGTAGAGGGAGGAAG  
 TGAGGTGAGGTGGGTCCCCCACCAGCCAGGCCATGAAACTCAATGTGCCCGAAAGAC

**Supplementary Figure 4.** Full length sequence of the SVA F inserted in exon 13 of *BBS1*. The sequence is available in GenBank (Sayers et al. 2019) under the following accession MT113356. The underlined sequence has been extracted from WGS data. Nucleotides written in red correspond to changes between the first genome extracted sequence and Sanger sequence. Each domain of the SVA is labelled by color : **exon 13 of *BBS1***, **TSD**, **Alu like**, **VNTR**, **SINE-R**, **polyadenylation site** and **poly(A) tail**.

**A**

5' GTGGTGGCCTGATCATC AAGATCCTGAAGCGTACAGCAGTGTTTGTAGAGGG 3'  
 3' CACCACCGGACTAGTAG TTCTAGGACTTCGCATGTCGTCACAAACATGTCCC 5'

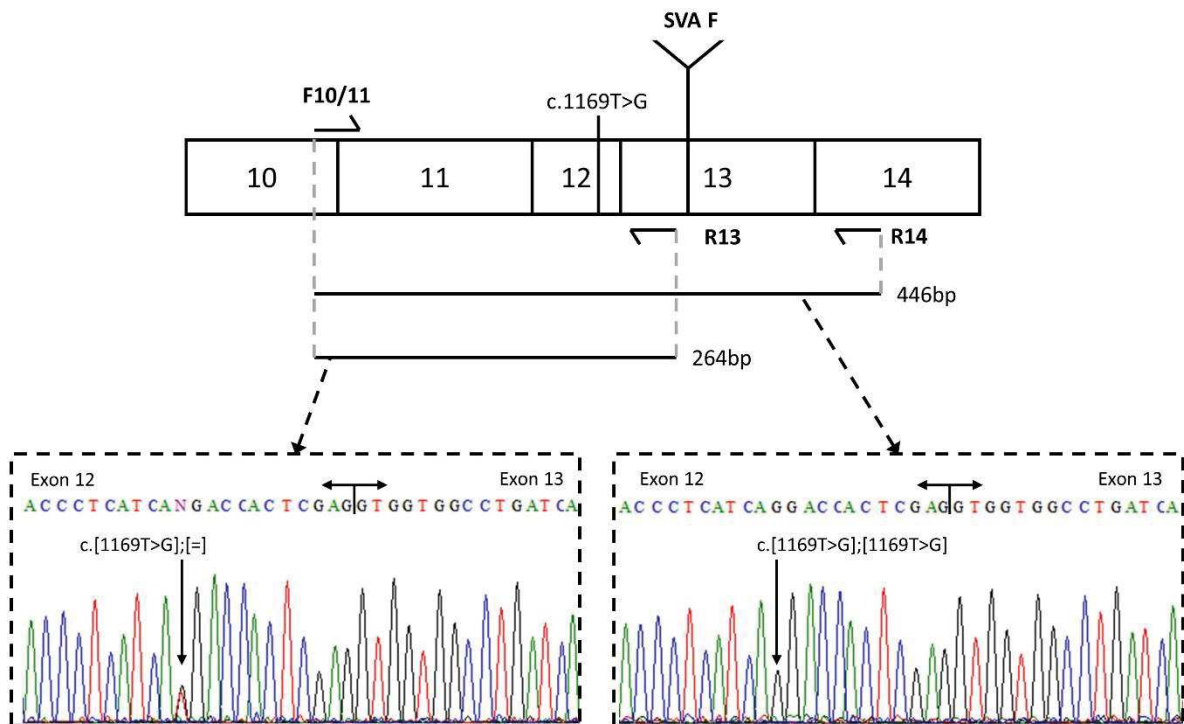
↑  
 Target site  
 for cleavage

**B**

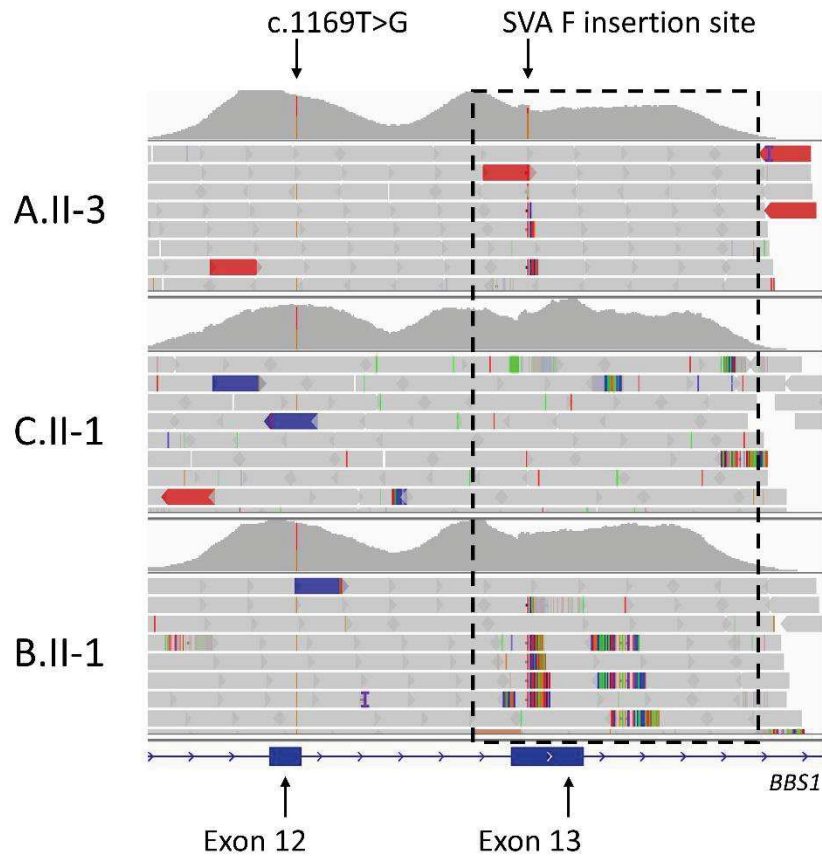
5' CATC AAGATCCTGAAGAGTAC TC----- (A)<sub>125</sub> AAGATCCTGAAGAGTACAGCAGT 3'

TSD sequence SVA-F BBS1 exon 13

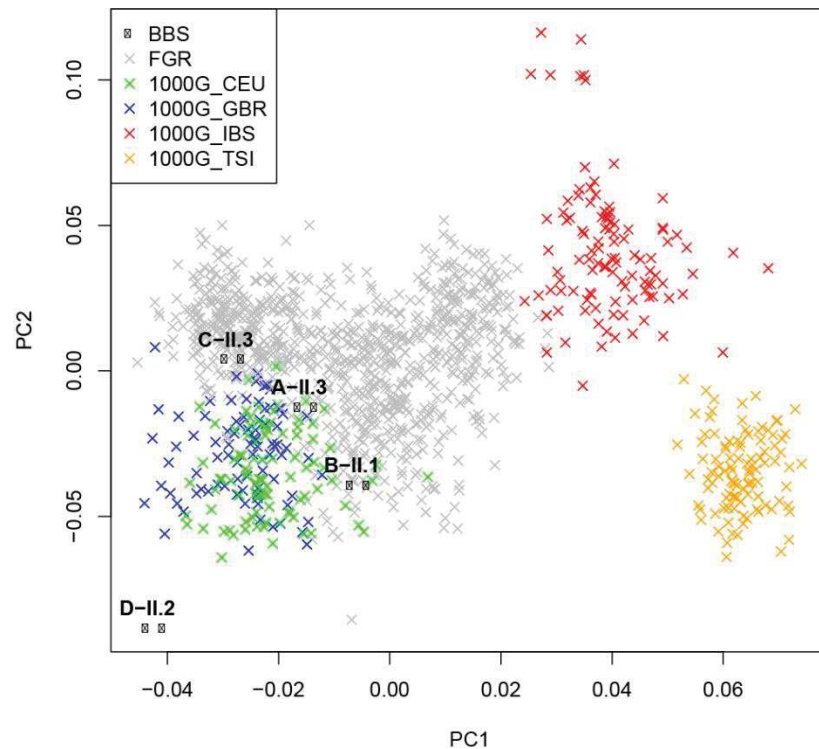
**Supplementary Figure 5.** Identification of the target sequence for the SVA F integration in exon 13 of *BBS1*. (A) The sequence of the integration site 5'-TCTT/GA-3' is highlighted in blue and differs slightly from the consensus sequence 5'-TTTT/AA-3' according to Raiz *et al.* (Raiz *et al.* 2012b) (B) The integration of the SVA F (in yellow) will occur in exon 13 of *BBS1* (in green) resulting in target site duplication (TSD, in orange).



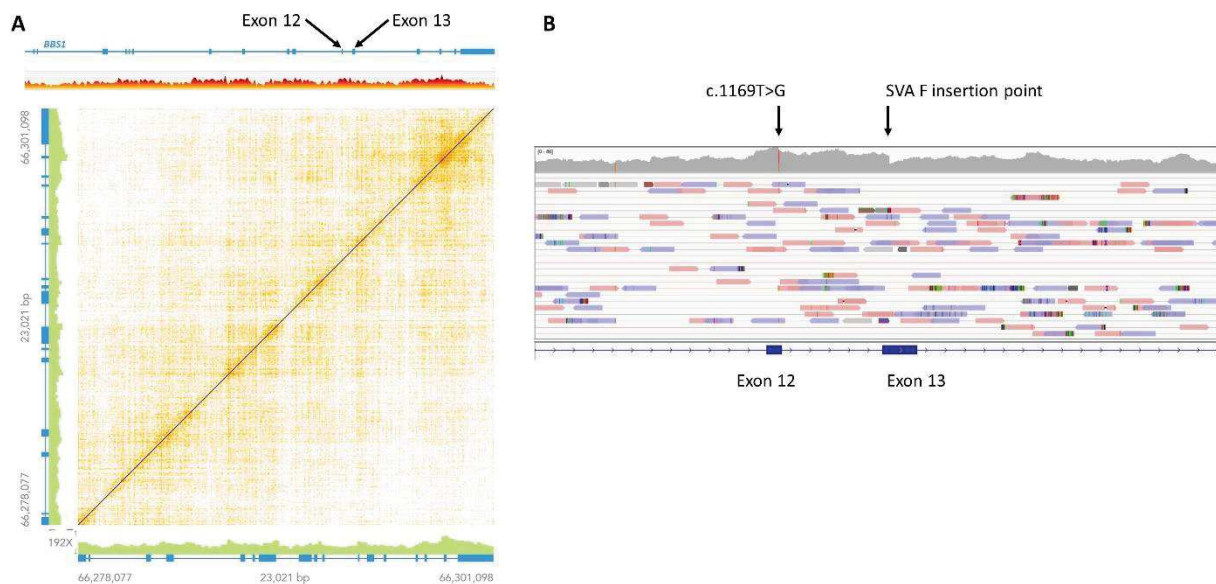
**Supplementary Figure 6.** RNA analysis surrounding the SVA insertion demonstrates the existence of an mRNA carrying the SVA insertion. Two amplicons have been designed around the insertion point of the SVA F that also integrate the second *BBS1* variant (c.1169T>G). Analysis of both sequences reveal that depending on the integration of the SVA F in the amplicon, the c.1169T>G variant is either observed in the heterozygous or homozygous state.



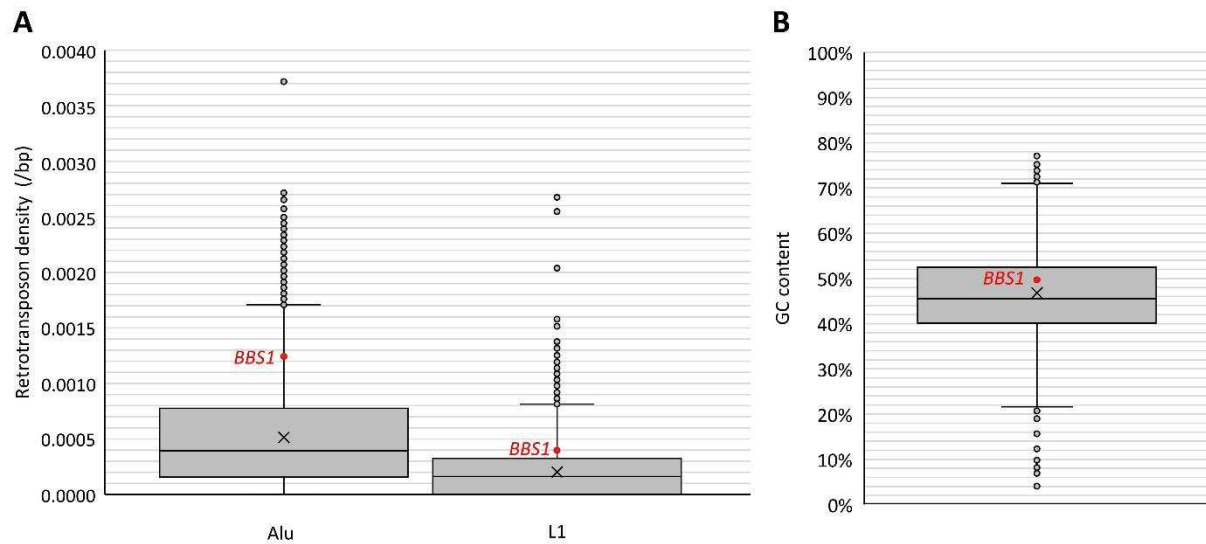
**Supplementary Figure 7.** Comparison of next generation sequencing data surrounding the *BBS1* locus from patient A.II-3, B.II-1 and C.II-1. All harbour the heterozygous c.1169T>G (p.M390R) in exon 12. One can observe the absence of the miscalled variant in exon 13 in B.II-1 and C.II-1 compared to A.II-3 as described in supplementary Figure 2. Depth of coverage profiles among exon 13 are slightly different between the 3 examples but a small dropout is visible upon SVA F insertion site.



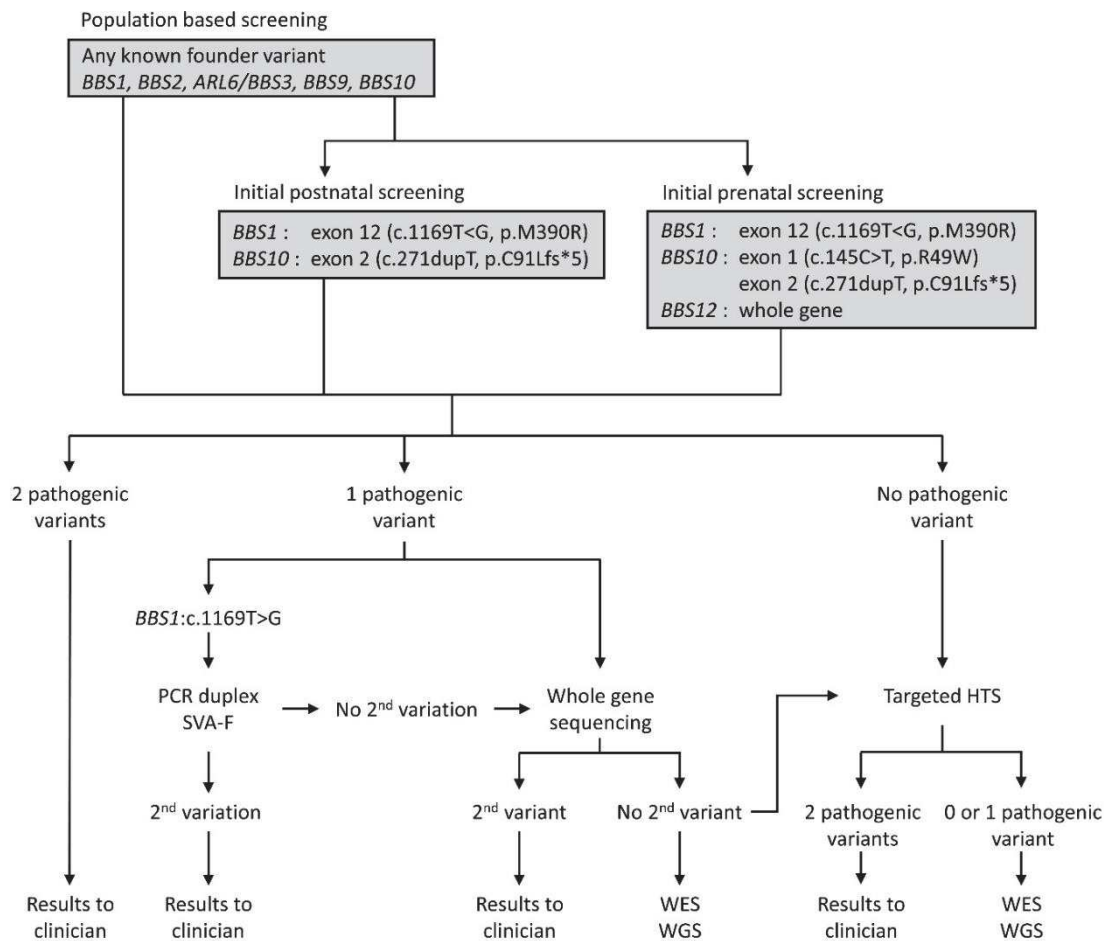
**Supplementary Figure 8.** Principal Component Analysis performed on the genotype data from common SNPs (MAF  $\geq 0.05$ ) for the four SVA-F carriers compared against FranceGenRef (FGR) and 1000 Genomes (The 1000 Genomes Project Consortium et al. 2015) European Non Finnish populations. A total of 118,577 SNVs were included in the analysis. CEU: Utah Residents (CEPH) with Northern and Western European Ancestry, GBR: British in England and Scotland, IBS: Iberian Population in Spain, TSI: Toscani in Italy.



**Supplementary Figure 9.** Linked-read analysis of the genome of patient A.II-3. (A) Loupe visualization of the *BBS1* locus, linear (top) and matrix (bottom) type. No specific signal is detected for exon 13 of *BBS1*. (B) IGV visualization of the insertion point in exon 13 of *BBS1* shows that there is a difference in depth of coverage.



**Supplementary Figure 10. Genomic context analysis of *BBS1*.** In the boxplots, the lower and upper boundaries indicate the 2<sup>nd</sup> and 3<sup>rd</sup> quartile respectively. The median is represented as a black line and the mean as an “X” within the box. Whiskers above and below the box indicate the 1<sup>st</sup> and 4<sup>th</sup> quartiles and points outside represents outliers. (A) Boxplot distribution of the *Alu* and L1 elements within the 19,941 human genes (RefSeq). (B) GC content distribution of the human gene. *BBS1* is highlighted in red.



**Supplementary Figure 11.** Updated BBS diagnostic strategy including SVA F fast detection in exon 13 using a duplex PCR.

## SUPPLEMENTARY REFERENCES

Akman, Hasan O., Guido Davidzon, Kurenai Tanji, Emma J. MacDermott, Louann Larsen, Mercy M. Davidson, Ronald G. Haller, et al. 2010a. « Neutral Lipid Storage Disease with Subclinical Myopathy Due to a Retrotransposal Insertion in the PNPLA2 Gene ». *Neuromuscular Disorders* 20 (6): 397-402. <https://doi.org/10.1016/j.nmd.2010.04.004>.

Auton, Adam, Gonçalo R. Abecasis, David M. Altshuler, Richard M. Durbin, Gonçalo R. Abecasis, David R. Bentley, Aravinda Chakravarti, et al. 2015. « A global reference for human genetic variation ». *Nature* 526 (7571): 68-74. <https://doi.org/10.1038/nature15393>.

Backenroth, Daniel, Jason Homsy, Laura R. Murillo, Joe Glessner, Edwin Lin, Martina Brueckner, Richard Lifton, Elizabeth Goldmuntz, Wendy K. Chung, et Yufeng Shen. 2014. « CANOES: Detecting Rare Copy Number Variants from Whole Exome Sequencing Data ». *Nucleic Acids Research* 42 (12): e97. <https://doi.org/10.1093/nar/gku345>.

Badano, Jose L., Jun Chul Kim, Bethan E. Hoskins, Richard Alan Lewis, Stephen J. Ansley, David J. Cutler, Claudio Castellan, Philip L. Beales, Michel R. Leroux, et Nicholas Katsanis. 2003. « Heterozygous Mutations in BBS1, BBS2 and BBS6 Have a Potential Epistatic Effect on Bardet–Biedl Patients with Two Mutations at a Second BBS Locus ». *Human Molecular Genetics* 12 (14): 1651-59. <https://doi.org/10.1093/hmg/ddg188>.



Beales, P. L., N. Elcioglu, A. S. Woolf, D. Parker, et F. A. Flinter. 1999. « New Criteria for Improved Diagnosis of Bardet-Biedl Syndrome: Results of a Population Survey ». *Journal of Medical Genetics* 36 (6): 437-46.

Beales, Philip L., Jose L. Badano, Alison J. Ross, Stephen J. Ansley, Bethan E. Hoskins, Brigitta Kirsten, Charles A. Mein, et al. 2003. « Genetic Interaction of BBS1 Mutations with Alleles at Other BBS Loci Can Result in Non-Mendelian Bardet-Biedl Syndrome ». *The American Journal of Human Genetics* 72 (5): 1187-99. <https://doi.org/10.1086/375178>.

Bragg, D. Cristopher, Kotchaphorn Mangkalaphiban, Christine A. Vaine, Nichita J. Kulkarni, David Shin, Rachita Yadav, Jyotsna Dhakal, et al. 2017. « Disease Onset in X-Linked Dystonia-Parkinsonism Correlates with Expansion of a Hexameric Repeat within an SVA Retrotransposon in *TAF1* ». *Proceedings of the National Academy of Sciences* 114 (51): E11020-28. <https://doi.org/10.1073/pnas.1712526114>.

Brett, Maggie, George Korovesis, Angeline H. M. Lai, Eileen C. P. Lim, et Ene-Choo Tan. 2017. « Intragenic Multi-Exon Deletion in the *FBN1* Gene in a Child with Mildly Dilated Aortic Sinus: A Retrotransposal Event ». *Journal of Human Genetics* 62 (7): 711-15. <https://doi.org/10.1038/jhg.2017.32>.

Buckley, Reuben M., R. Daniel Kortschak, Joy M. Raison, et David L. Adelson. 2017. « Similar Evolutionary Trajectories for Retrotransposon Accumulation in Mammals ». *Genome Biology and Evolution* 9 (9): 2336-53. <https://doi.org/10.1093/gbe/evx179>.

Burns, Kathleen H. 2020. « Our Conflict with Transposable Elements and Its Implications for Human Disease ». *Annual Review of Pathology: Mechanisms of Disease* 15 (1): 51-70. <https://doi.org/10.1146/annurev-pathmechdis-012419-032633>.

Byeon, Boseon, et Igor Kovalchuk. 2016. « Non-Coding RNAs Match the Deleted Genomic Regions in Humans ». *Scientific Reports* 6 (1). <https://doi.org/10.1038/srep37452>.

Chang, Christopher C, Carson C Chow, Laurent CAM Tellier, Shashaank Vattikuti, Shaun M Purcell, et James J Lee. 2015. « Second-generation PLINK: rising to the challenge of larger and richer datasets ». *GigaScience* 4 (s13742-015-0047-8). <https://doi.org/10.1186/s13742-015-0047-8>.

Chatron, Nicolas, Kevin Cassinari, Olivier Quenez, Stéphanie Baert-Desurmont, Claire Bardel, Marie-Pierre Buisine, Eduardo Calpena, et al. 2019. « Identification of mobile retrocopies during genetic testing: Consequences for routine diagnosis ». *Human Mutation* 40 (11): 1993-2000. <https://doi.org/10.1002/humu.23845>.

Chen, Jian-Min, Emmanuelle Masson, Milan Macek, Odile Raguénès, Tereza Piskackova, Brigitte Fercot, Libor Fila, David N. Cooper, Marie-Pierre Audrézet, et Claude Férec. 2008. « Detection of Two Alu Insertions in the *CFTR* Gene ». *Journal of Cystic Fibrosis* 7 (1): 37-43. <https://doi.org/10.1016/j.jcf.2007.04.001>.

Collins, Ryan L., Harrison Brand, Konrad J. Karczewski, Xuefang Zhao, Jessica Alföldi, Laurent C. Francioli, Amit V. Khera, et al. 2020. « A structural variation reference for medical and population genetics ». *Nature* 581 (7809): 444-51. <https://doi.org/10.1038/s41586-020-2287-8>.

Conley, Mary Ellen, Julie D. Partain, Shannon M. Norland, Sheila A. Shurtleff, et Haig H. Kazazian. 2005. « Two Independent Retrotransposon Insertions at the Same Site within the Coding Region of *BTK* ». *Human Mutation* 25 (3): 324-25. <https://doi.org/10.1002/humu.9321>.



Crivelli, Louise, Virginie Bubien, Natalie Jones, Jennifer Chiron, Françoise Bonnet, Emmanuelle Barouk-Simonet, Patrice Couzigou, Nicolas Sevenet, Frédéric Caux, et Michel Longy. 2017. « Insertion of Alu Elements at a PTEN Hotspot in Cowden Syndrome ». *European Journal of Human Genetics : EJHG* 25 (9): 1087-91. <https://doi.org/10.1038/ejhg.2017.81>.

Cummings, Beryl B., Jamie L. Marshall, Taru Tukiainen, Monkol Lek, Sandra Donkervoort, A. Reghan Foley, Veronique Bolduc, et al. 2017. « Improving genetic diagnosis in Mendelian disease with transcriptome sequencing ». *Science Translational Medicine* 9 (386): eaal5209. <https://doi.org/10.1126/scitranslmed.aal5209>.

DePristo, Mark A, Eric Banks, Ryan Poplin, Kiran V Garimella, Jared R Maguire, Christopher Hartl, Anthony A Philippakis, et al. 2011. « A framework for variation discovery and genotyping using next-generation DNA sequencing data ». *Nature Genetics* 43 (5): 491-98. <https://doi.org/10.1038/ng.806>.

Estrada-Cuzcano, Alejandro, Christelle Etard, Clarisse Delvallée, Corinne Stoetzel, Elise Schaefer, Sophie Scheidecker, Véronique Geoffroy, et al. 2020. « Novel IQCE variations confirm its role in postaxial polydactyly and cause ciliary defect phenotype in zebrafish ». *Human Mutation* 41 (1): 240-54. <https://doi.org/10.1002/humu.23924>.

Frag, Talaat I., et Ahmad S. Teebi. 1989. « High incidence of Bardet Biedl syndrome among the Bedouin ». *Clinical Genetics* 36 (6): 463-64. <https://doi.org/10.1111/j.1399-0004.1989.tb03378.x>.

Firth, Helen V., Shola M. Richards, A. Paul Bevan, Stephen Clayton, Manuel Corpas, Diana Rajan, Steven Van Vooren, Yves Moreau, Roger M. Pettett, et Nigel P. Carter. 2009. « DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources ». *The American Journal of Human Genetics* 84 (4): 524-33. <https://doi.org/10.1016/j.ajhg.2009.03.010>.

Gardner, Eugene J., Elena Prigmore, Giuseppe Gallone, Petr Danecek, Kaitlin E. Samocha, Juliet Handsaker, Sebastian S. Gerety, et al. 2019. « Contribution of retrotransposition to developmental disorders ». *Nature Communications* 10 (1): 4630. <https://doi.org/10.1038/s41467-019-12520-y>.

Gasior, Stephen L., Graeme Preston, Dale J. Hedges, Nicolas Gilbert, John V. Moran, et Prescott L. Deisinger. 2007. « Characterization of Pre-Insertion Loci of de Novo L1 Insertions ». *Gene* 390 (1-2): 190-98. <https://doi.org/10.1016/j.gene.2006.08.024>.

Genin, E, A Tullio-Pelet, F Begeot, S Lyonnet, et L Abel. 2004. « Estimating the age of rare disease mutations: the example of Triple-A syndrome ». *Journal of Medical Genetics* 41 (6): 445. <https://doi.org/10.1136/jmg.2003.017962>.

Geoffroy, Véronique, Yvan Herenger, Arnaud Kress, Corinne Stoetzel, Amélie Piton, Hélène Dollfus, et Jean Muller. 2018. « AnnotSV: An Integrated Tool for Structural Variations Annotation ». *Bioinformatics (Oxford, England)* 34 (20): 3572-74. <https://doi.org/10.1093/bioinformatics/bty304>.

Geoffroy, Véronique, Cécile Pizot, Claire Redin, Amélie Piton, Nasim Vasli, Corinne Stoetzel, André Blavier, Jocelyn Laporte, et Jean Muller. 2015. « VaRank: A Simple and Powerful Tool for Ranking Genetic Variants ». *PeerJ* 3: e796. <https://doi.org/10.7717/peerj.796>.

Geoffroy, Véronique, Corinne Stoetzel, Sophie Scheidecker, Elise Schaefer, Isabelle Perrault, Séverine Bär, Ariane Kröll, et al. 2018. « Whole-genome sequencing in patients with ciliopathies uncovers a novel recurrent tandem duplication in IFT140 ». *Human Mutation* 39 (7): 983-92. <https://doi.org/10.1002/humu.23539>.

- Goerner-Potvin, Patricia, et Guillaume Bourque. 2018. « Computational tools to unmask transposable elements ». *Nature Reviews Genetics* 19 (11): 688-704. <https://doi.org/10.1038/s41576-018-0050-x>.
- Gouronc, Aurélie, Vincent Zilliox, Marie-Line Jacquemont, Françoise Darcel, Anne-Sophie Leuvrey, Elsa Nourisson, Manuela Antin, et al. 2020. « High prevalence of Bardet-Biedl syndrome in La Réunion Island is due to a founder variant in ARL6/BBS3 ». *Clinical Genetics* n/a (n/a). <https://doi.org/10.1111/cge.13768>.
- Ha, Hongseok, Jui Wan Loh, et Jinchuan Xing. 2016. « Identification of Polymorphic SVA Retrotransposons Using a Mobile Element Scanning Method for SVA (ME-Scan-SVA) ». *Mobile DNA* 7 (1). <https://doi.org/10.1186/s13100-016-0072-x>.
- Hancks, Dustin C., et Haig H. Kazazian. 2010. « SVA Retrotransposons: Evolution and Genetic Instability ». *Seminars in Cancer Biology* 20 (4): 234-45. <https://doi.org/10.1016/j.semcancer.2010.04.001>.
- Hancks, Dustin C, et Haig H Kazazian. 2012. « Active Human Retrotransposons: Variation and Disease ». *Current Opinion in Genetics & Development* 22 (3): 191-203. <https://doi.org/10.1016/j.gde.2012.02.006>.
- Hancks, Dustin C., et Haig H. Kazazian. 2016. « Roles for retrotransposon insertions in human disease ». *Mobile DNA* 7 (mai). <https://doi.org/10.1186/s13100-016-0065-9>.
- Hassoun, H., T. L. Coetzer, J. N. Vassiliadis, K. E. Sahr, G. J. Maalouf, S. T. Saad, L. Catanzariti, et J. Palek. 1994. « A Novel Mobile Element Inserted in the Alpha Spectrin Gene: Spectrin Dayton. A Truncated Alpha Spectrin Associated with Hereditary Elliptocytosis ». *The Journal of Clinical Investigation* 94 (2): 643-48. <https://doi.org/10.1172/JCI117380>.
- Hayano, Takahide, Shiro Yamada, Kazuyoshi Hosomichi, Hirofumi Nakaoka, Kosuke Yoshihara, Sosuke Adachi, Katsunori Kashima, Kenichi Tanaka, Takayuki Enomoto, et Ituro Inoue. 2015. « Identification of Novel Exonic Mobile Element Insertions in Epithelial Ovarian Cancers ». *Human Genome Variation* 2 (octobre): 15030-15030. <https://doi.org/10.1038/hgv.2015.30>.
- Karaca, Ender, Jennifer E. Posey, Zeynep Coban Akdemir, Davut Pehlivan, Tamar Harel, Shalini N. Jhangiani, Yavuz Bayram, et al. 2018. « Phenotypic expansion illuminates multilocus pathogenic variation ». *Genetics in Medicine* 20 (12): 1528-37. <https://doi.org/10.1038/gim.2018.33>.
- Katsanis, Nicholas, Stephen J. Ansley, Jose L. Badano, Erica R. Eichers, Richard Alan Lewis, Bethan E. Hoskins, Peter J. Scambler, William S. Davidson, Philip L. Beales, et James R. Lupski. 2001. « Triallelic Inheritance in Bardet-Biedl Syndrome, a Mendelian Recessive Disorder ». *Science* 293 (5538): 2256. <https://doi.org/10.1126/science.1063525>.
- Kazazian, Haig H., Corinne Wong, Hagop Youssoufian, Alan F. Scott, Deborah G. Phillips, et Stylianos E. Antonarakis. 1988a. « Haemophilia A Resulting from de Novo Insertion of L1 Sequences Represents a Novel Mechanism for Mutation in Man ». *Nature* 332 (6160): 164-66. <https://doi.org/10.1038/332164a0>.
- Kent, W James. 2002. « BLAT--the BLAST-like Alignment Tool ». *Genome Research* 12 (4): 656-64. <https://doi.org/10.1101/gr.229202>.
- Khan, Sher Alam, Noor Muhammad, Muzammil Ahmad Khan, Anwar Kamal, Zia Ur Rehman, et Saadullah Khan. 2016. « Genetics of Human Bardet-Biedl Syndrome, an Updates ». *Clinical Genetics*, janvier. <https://doi.org/10.1111/cge.12737>.

- Kherraf, Zine-Eddine, Amir Amiri-Yekta, Denis Dacheux, Thomas Karaouzène, Charles Coutton, Marie Christou-Kent, Guillaume Martinez, et al. 2018. « A Homozygous Ancestral SVA-Insertion-Mediated Deletion in WDR66 Induces Multiple Morphological Abnormalities of the Sperm Flagellum and Male Infertility ». *The American Journal of Human Genetics* 103 (3): 400-412. <https://doi.org/10.1016/j.ajhg.2018.07.014>.
- Klift, Heleen M. van der, Carli M. Tops, Frederik J. Hes, Peter Devilee, et Juul T. Wijnen. 2012. « Insertion of an SVA Element, a Nonautonomous Retrotransposon, in *PMS2* Intron 7 as a Novel Cause of Lynch Syndrome ». *Human Mutation* 33 (7): 1051-55. <https://doi.org/10.1002/humu.22092>.
- Kobayashi, K., Y. Nakahori, M. Miyake, K. Matsumura, E. Kondo-Iida, Y. Nomura, M. Segawa, et al. 1998. « An Ancient Retrotransposal Insertion Causes Fukuyama-Type Congenital Muscular Dystrophy ». *Nature* 394 (6691): 388-92. <https://doi.org/10.1038/28653>.
- Kremer, Laura S., Daniel M. Bader, Christian Mertes, Robert Kopajtich, Garwin Pichler, Arcangela Iuso, Tobias B. Haack, et al. 2017. « Genetic diagnosis of Mendelian disorders via RNA sequencing ». *Nature Communications* 8 (1): 15824. <https://doi.org/10.1038/ncomms15824>.
- Kwon, Yun-Jeong, Yuri Choi, Jungwoo Eo, Yu-Na Noh, Jeong-An Gim, Yi-Deun Jung, Ja-Rang Lee, et Heui-Soo Kim. 2013a. « Structure and Expression Analyses of SVA Elements in Relation to Functional Genes ». *Genomics & Informatics* 11 (3): 142. <https://doi.org/10.5808/GI.2013.11.3.142>.
- Lee, Christopher M, Galt P Barber, Jonathan Casper, Hiram Clawson, Mark Diekhans, Jairo Navarro Gonzalez, Angie S Hinrichs, et al. 2019. « UCSC Genome Browser enters 20th year ». *Nucleic Acids Research* 48 (D1): D756-61. <https://doi.org/10.1093/nar/gkz1012>.
- Li, Heng. 2011. « A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data ». *Bioinformatics* 27 (21): 2987-93. <https://doi.org/10.1093/bioinformatics/btr509>.
- Li, Heng, et Richard Durbin. 2009. « Fast and accurate short read alignment with Burrows–Wheeler transform ». *Bioinformatics* 25 (14): 1754-60. <https://doi.org/10.1093/bioinformatics/btp324>.
- Li, Quan, et Kai Wang. 2017. « InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines ». *The American Journal of Human Genetics* 100 (2): 267-80. <https://doi.org/10.1016/j.ajhg.2017.01.004>.
- Lindstrand, Anna, Erica E. Davis, Claudia M. B. Carvalho, Davut Pehlivan, Jason R. Willer, I.-Chun Tsai, Subhadra Ramanathan, et al. 2014. « Recurrent CNVs and SNVs at the NPHP1 Locus Contribute Pathogenic Alleles to Bardet-Biedl Syndrome. » *American Journal of Human Genetics* 94 (5): 745-54. <https://doi.org/10.1016/j.ajhg.2014.03.017>.
- Lindstrand, Anna, Stephan Frangakis, Claudia M.B. Carvalho, Ellen B. Richardson, Kelsey A. McFadden, Jason R. Willer, Davut Pehlivan, et al. 2016. « Copy-Number Variation Contributes to the Mutational Load of Bardet-Biedl Syndrome ». *The American Journal of Human Genetics* 99 (2): 318-36. <https://doi.org/10.1016/j.ajhg.2015.04.023>.
- Martin, Sandra L. 2018. « On the move ». *eLife* 7 (février): e34901. <https://doi.org/10.7554/eLife.34901>.

- Mary, Laura, Kirsley Chennen, Corinne Stoetzel, Manuela Antin, Anne Leuvrey, Elsa Nourisson, Elisabeth Alanio-Detton, et al. 2019. « Bardet-Biedl syndrome: Antenatal presentation of forty-five fetuses with biallelic pathogenic variants in known Bardet-Biedl syndrome genes ». *Clinical Genetics* 95 (3): 384-97. <https://doi.org/10.1111/cge.13500>.
- Mathieson, Iain, et Gil McVean. 2014. « Demography and the Age of Rare Variants ». *PLOS Genetics* 10 (8): e1004528. <https://doi.org/10.1371/journal.pgen.1004528>.
- Medina, Julie, et Hervé Perron. 2017. « Séquences provenant d'éléments génétiques mobiles, face cachée du génome humain ». *médecine/sciences* 33 (2): 151-58. <https://doi.org/10.1051/medsci/20173302010>.
- Medstrand, Patrik, Louie N. van de Lagemaat, et Dixie L. Mager. 2002. « Retroelement Distributions in the Human Genome: Variations Associated With Age and Proximity to Genes ». *Genome Research* 12 (10): 1483-95. <https://doi.org/10.1101/gr.388902>.
- Miki, Yoshio, Toyomasa Katagiri, Fujio Kasumi, Takamasa Yoshimoto, et Yusuke Nakamura. 1996. « Mutation Analysis in the BRCA2 Gene in Primary Breast Cancers ». *Nature Genetics* 13 (2): 245-47. <https://doi.org/10.1038/ng0696-245>.
- Mills, Ryan E., E. Andrew Bennett, Rebecca C. Iskow, et Scott E. Devine. 2007. « Which Transposable Elements Are Active in the Human Genome? » *Trends in Genetics* 23 (4): 183-91. <https://doi.org/10.1016/j.tig.2007.02.006>.
- Mitchison, Hannah M, et Enza Maria Valente. 2017. « Motile and non-motile cilia in human pathology: from function to phenotypes ». *The Journal of Pathology* 241 (2): 294-309. <https://doi.org/10.1002/path.4843>.
- Muller, Jean, C. Stoetzel, M. C. Vincent, C. C. Leitch, V. Laurier, J. M. Danse, S. Hellé, et al. 2010. « Identification of 28 novel mutations in the Bardet-Biedl syndrome genes: the burden of private mutations in an extensively heterogeneous disease ». *Human genetics* 127 (5): 583-93. <https://doi.org/10.1007/s00439-010-0804-9>.
- Mykytyn, Kirk, Darryl Y. Nishimura, Charles C. Searby, Gretel Beck, Kevin Bugge, Heidi L. Haines, Alberto S. Cornier, et al. 2003. « Evaluation of Complex Inheritance Involving the Most Common Bardet-Biedl Syndrome Locus (BBS1) ». *The American Journal of Human Genetics* 72 (2): 429-37. <https://doi.org/10.1086/346172>.
- Mykytyn, Kirk, Darryl Y. Nishimura, Charles C. Searby, Mythreyi Shastri, Hsan-jan Yen, John S. Beck, Terry Braun, et al. 2002. « Identification of the Gene (BBS1) Most Commonly Involved in Bardet-Biedl Syndrome, a Complex Human Obesity Syndrome ». *Nature Genetics* 31 (4): 435-38. <https://doi.org/10.1038/ng935>.
- Nakamura, Yuki, Moe Murata, Yuki Takagi, Toshihiro Kozuka, Yukiko Nakata, Ryo Hasebe, Akira Takagi, Jun-ichi Kitazawa, Midori Shima, et Tetsuhito Kojima. 2015. « SVA Retrotransposition in Exon 6 of the Coagulation Factor IX Gene Causing Severe Hemophilia B ». *International Journal of Hematology* 102 (1): 134-39. <https://doi.org/10.1007/s12185-015-1765-5>.
- Niederlova, Veronika, Martin Modrak, Oksana Tsyklauri, Martina Huranova, et Ondrej Stepanek. 2019. « Meta-analysis of genotype-phenotype associations in Bardet-Biedl Syndrome uncovers differences among causative genes ». *Human Mutation* 0 (ja). <https://doi.org/10.1002/humu.23862>.

Okonechnikov, Konstantin, Ana Conesa, et Fernando García-Alcalde. 2015. « Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data ». *Bioinformatics* 32 (2): 292-94. <https://doi.org/10.1093/bioinformatics/btv566>.

O'Leary, Nuala A., Mathew W. Wright, J. Rodney Brister, Stacy Ciuffo, Diana Haddad, Rich McVeigh, Bhanu Rajput, et al. 2015. « Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation ». *Nucleic Acids Research* 44 (D1): D733-45. <https://doi.org/10.1093/nar/gkv1189>.

Ostertag, Eric M., John L. Goodier, Yue Zhang, et Haig H. Kazazian. 2003a. « SVA Elements Are Nonautonomous Retrotransposons That Cause Disease in Humans ». *The American Journal of Human Genetics* 73 (6): 1444-51. <https://doi.org/10.1086/380207>.

Raiz, Julija, Annette Damert, Sergiu Chira, Ulrike Held, Sabine Klawitter, Matthias Hamdorf, Johannes Löwer, Wolf H. Strätling, Roswitha Löwer, et Gerald G. Schumann. 2012a. « The Non-Autonomous Retrotransposon SVA Is Trans -Mobilized by the Human LINE-1 Protein Machinery ». *Nucleic Acids Research* 40 (4): 1666-83. <https://doi.org/10.1093/nar/gkr863>.

Redin, Claire, Stéphanie Le Gras, Oussema Mhamdi, Véronique Geoffroy, Corinne Stoetzel, Marie-Claire Vincent, Pietro Chiurazzi, et al. 2012a. « Targeted high-throughput sequencing for diagnosis of genetically heterogeneous diseases: efficient mutation detection in Bardet-Biedl and Alström Syndromes ». *Journal of Medical Genetics* 49 (8): 502-12. <https://doi.org/10.1136/jmedgenet-2012-100875>.

Sayers, Eric W., Jeff Beck, J. Rodney Brister, Evan E. Bolton, Kathi Canese, Donald C. Comeau, Kathryn Funk, et al. 2019. « Database Resources of the National Center for Biotechnology Information. » *Nucleic Acids Research*, octobre. <https://doi.org/10.1093/nar/gkz899>.

Scheidecker, Sophie, Christelle Etard, Nathan W. Pierce, Véronique Geoffroy, Elise Schaefer, Jean Muller, Kirsley Chennen, et al. 2014. « Exome Sequencing of Bardet-Biedl Syndrome Patient Identifies a Null Mutation in the BBSome Subunit BBIP1 (BBS18) ». *Journal of Medical Genetics* 51 (2): 132-36. <https://doi.org/10.1136/jmedgenet-2013-101785>.

Shaheen, Ranad, Katarzyna Szymanska, Basudha Basu, Nisha Patel, Nour Ewida, Eissa Faqeih, Amal Al Hashem, et al. 2016. « Characterizing the morbid genome of ciliopathies ». *Genome Biology* 17 (1): 242. <https://doi.org/10.1186/s13059-016-1099-5>.

Shamseldin, Hanan E., Ranad Shaheen, Nour Ewida, Dalal K. Bubshait, Hisham Alkuraya, Elham Almardawi, Ali Howaidi, et al. 2020. « The morbid genome of ciliopathies: an update ». *Genetics in Medicine*, février. <https://doi.org/10.1038/s41436-020-0761-1>.

Smit, AFA, R Hubley, et P Green. 2013. « RepeatMasker Open-4.0 », 2015 2013. <http://www.repeatmasker.org>.

Stemers, Frank J, Weihua Chang, Grace Lee, David L Barker, Richard Shen, et Kevin L Gunderson. 2006. « Whole-genome genotyping with the single-base extension assay ». *Nature Methods* 3 (1): 31-33. <https://doi.org/10.1038/nmeth842>.

Stewart, Chip, Deniz Kural, Michael P Strömberg, Jerilyn A Walker, Miriam K Konkel, Adrian M Stütz, Alexander E Urban, et al. 2011. « A Comprehensive Map of Mobile Element Insertion Polymorphisms in Humans ». *PLoS Genetics* 7 (8): e1002236-e1002236. <https://doi.org/10.1371/journal.pgen.1002236>.



- Stoetzel, Corinne, Virginie Laurier, Erica E. Davis, Jean Muller, Suzanne Rix, José L. Badano, Carmen C. Leitch, et al. 2006. « BBS10 Encodes a Vertebrate-Specific Chaperonin-like Protein and Is a Major BBS Locus ». *Nature Genetics* 38 (5): 521-24. <https://doi.org/10.1038/ng1771>.
- Sudmant, Peter H., Tobias Rausch, Eugene J. Gardner, Robert E. Handsaker, Alexej Abyzov, John Huddleston, Yan Zhang, et al. 2015. « An integrated map of structural variation in 2,504 human genomes ». *Nature* 526 (7571): 75-81. <https://doi.org/10.1038/nature15394>.
- Takasu, M., R. Hayashi, E. Maruya, M. Ota, K. Imura, K. Kougo, C. Kobayashi, et al. 2007. « Deletion of Entire HLA-A Gene Accompanied by an Insertion of a Retrotransposon ». *Tissue Antigens* 70 (2): 144-50. <https://doi.org/10.1111/j.1399-0039.2007.00870.x>.
- Taniguchi-Ikeda, Mariko, Kazuhiro Kobayashi, Motoi Kanagawa, Chih-chieh Yu, Kouhei Mori, Tetsuya Oda, Atsushi Kuga, et al. 2011. « Pathogenic Exon-Trapping by SVA Retrotransposon and Rescue in Fukuyama Muscular Dystrophy ». *Nature* 478 (7367): 127-31. <https://doi.org/10.1038/nature10456>.
- Taşkesen, Mustafa, Gayle B Collin, Alexei V Evsikov, Ayşegül Güzel, R Köksal Özgül, Jan D Marshall, et Jürgen K Naggert. 2012. « Novel Alu Retrotransposon Insertion Leading to Alström Syndrome ». *Human Genetics* 131 (3): 407-13. <https://doi.org/10.1007/s00439-011-1083-9>.
- Tavares, Erika, Chen Yu Tang, Anjali Vig, Shuning Li, Gail Billingsley, Wilson Sung, Ajoy Vincent, Bhooma Thiruvahindrapuram, et Elise Héon. 2019a. « Retrotransposon Insertion as a Novel Mutational Event in Bardet-Biedl Syndrome ». *Molecular Genetics & Genomic Medicine* 7 (2): e00521. <https://doi.org/10.1002/mgg3.521>.
- The 1000 Genomes Project Consortium, Richard A. Gibbs, Eric Boerwinkle, Harsha Doddapaneni, Yi Han, Viktoriya Korchina, Christie Kovar, et al. 2015. « A Global Reference for Human Genetic Variation ». *Nature* 526 (7571): 68-74. <https://doi.org/10.1038/nature15393>.
- Thorvaldsdottir, H., J. T. Robinson, et J. P. Mesirov. 2013. « Integrative Genomics Viewer (IGV): High-Performance Genomics Data Visualization and Exploration ». *Briefings in Bioinformatics* 14 (2): 178-92. <https://doi.org/10.1093/bib/bbs017>.
- Thung, Djie Tjwan, Joep de Ligt, Lisenka EM Vissers, Marloes Steehouwer, Mark Kroon, Petra de Vries, Eline P. Slagboom, Kai Ye, Joris A. Veltman, et Jayne Y. Hehir-Kwa. 2014. « Mobster: accurate detection of mobile element insertions in next generation sequencing data ». *Genome Biology* 15 (10): 488. <https://doi.org/10.1186/s13059-014-0488-x>.
- Torene, Rebecca I., Kevin Galens, Shuxi Liu, Kevin Arvai, Carlos Borroto, Julie Scuffins, Zhancheng Zhang, et al. 2020. « Mobile element insertion detection in 89,874 clinical exomes ». *Genetics in Medicine*, janvier. <https://doi.org/10.1038/s41436-020-0749-x>.
- Uguen, Kévin, Claire Jubin, Yannis Duffourd, Claire Bardel, Valérie Malan, Jean-Michel Dupont, Laila El Khattabi, et al. 2020. « Genome Sequencing in Cytogenetics: Comparison of Short-Read and Linked-Read Approaches for Germline Structural Variant Detection and Characterization ». *Molecular Genetics & Genomic Medicine* 8 (3): e1114-e1114. <https://doi.org/10.1002/mgg3.1114>.
- Vogt, Julia, Kathrin Bengesser, Kathleen BM Claes, Katharina Wimmer, Victor-Felix Mautner, Rick van Minkelen, Eric Legius, et al. 2014. « SVA retrotransposon insertion-associated deletion represents a novel mutational mechanism underlying large genomic copy number changes with non-recurrent breakpoints ». *Genome Biology* 15 (6): R80. <https://doi.org/10.1186/gb-2014-15-6-r80>.

- Wang, Hui, Jinchuan Xing, Deepak Grover, Dale J. Hedges, Kyudong Han, Jerilyn A. Walker, et Mark A. Batzer. 2005. « SVA Elements: A Hominid-Specific Retroposon Family ». *Journal of Molecular Biology* 354 (4): 994-1007. <https://doi.org/10.1016/j.jmb.2005.09.085>.
- Wang, Jianxin, Lei Song, Deepak Grover, Sami Azrak, Mark A Batzer, et Ping Liang. 2006. « DbRIP: A Highly Integrated Database of Retrotransposon Insertion Polymorphisms in Humans ». *Human Mutation* 27 (4): 323-29. <https://doi.org/10.1002/humu.20307>.
- Waters, Aoife M., et Philip L. Beales. 2011. « Ciliopathies: An Expanding Disease Spectrum ». *Pediatric Nephrology (Berlin, Germany)* 26 (7): 1039-56. <https://doi.org/10.1007/s00467-010-1731-7>.
- Williams, Amy L., Nick Patterson, Joseph Glessner, Hakon Hakonarson, et David Reich. 2012. « Phasing of Many Thousands of Genotyped Samples ». *The American Journal of Human Genetics* 91 (2): 238-51. <https://doi.org/10.1016/j.ajhg.2012.06.013>.
- Wilund, Kenneth R., Ming Yi, Filomena Campagna, Marcello Arca, Giovanni Zuliani, Renato Fellin, Yiu-Kee Ho, J. Victor Garcia, Helen H. Hobbs, et Jonathan C. Cohen. 2002. « Molecular Mechanisms of Autosomal Recessive Hypercholesterolemia ». *Human Molecular Genetics* 11 (24): 3019-30. <https://doi.org/10.1093/hmg/11.24.3019>.
- Wormser, Ohad, Libe Gradstein, Yuval Yogev, Yonatan Perez, Rotem Kadir, Inna Goliand, Yair Sadka, et al. 2019. « SCAPER localizes to primary cilia and its mutation affects cilia length, causing Bardet-Biedl syndrome ». *European Journal of Human Genetics* 27 (6): 928-40. <https://doi.org/10.1038/s41431-019-0347-z>.
- Yıldız Bölükbaşı, Esra, Sara Mumtaz, Muhammad Afzal, Ute Woehlbier, Sajid Malik, et Aslihan Tolun. 2018. « Homozygous mutation in *CEP19*, a gene mutated in morbid obesity, in Bardet-Biedl syndrome with predominant postaxial polydactyly ». *Journal of Medical Genetics* 55 (3): 189. <https://doi.org/10.1136/jmedgenet-2017-104758>.
- Zaghloul, Norann A., et Nicholas Katsanis. 2010. « Functional modules, mutational load and human genetic disease ». *Trends in Genetics* 26 (4): 168-76. <https://doi.org/10.1016/j.tig.2010.01.006>.
- Zerbino, Daniel R., et Ewan Birney. 2008. « Velvet: Algorithms for de Novo Short Read Assembly Using de Bruijn Graphs. » *Genome Research* 18 (5): 821-29. <https://doi.org/10.1101/gr.074492.107>.



Le projet III

portant sur *IQCE*

## I. Le synopsis.

Ce projet publié en 2019 dans *Human mutation* rapporte un travail collaboratif portant sur 3 familles pour lequel la validation fonctionnelle a été réalisée au laboratoire. Des poisson-zèbres *IQCE* ont également été générés en recourant à des morpholinos et étudiés grâce à une collaboration établie avec l'équipe de C. Etard et U. Strahle du KIT (Karlsruhe). *IQCE* est un gène rapporté chez une seule famille consanguine comme responsable de polydactylie post-axiale (PAP) chez l'homme (Umair et al. 2017). Cette information était inconnue au commencement de ce projet. 7 autres gènes sont également connus pour être responsables de polydactylie : *FAM92A1*, *GLI1*, *GLI3*, *KIAA0825*, *MIPOL1*, *PITX1* et *ZNF141*.

2 des 3 familles résolues dans ce projet sont consanguines et les patients présentent des phénotypes proches de ceux des ciliopathies syndromiques tel que le BBS. Voilà pourquoi ils ont été adressés à notre laboratoire. Ils présentent tous une polydactylie et une obésité. Une rétinopathie pigmentaire (RP) et un hypogonadisme viennent compléter le phénotype du patient de la famille 1. Une atteinte rénale et une surdité viennent compléter les phénotypes des patients de la famille 3. La première recherche de mutations a d'abord été limitée aux gènes *BBS* et s'est avérée sans succès. Ensuite, un WES a été réalisé et a permis d'identifier à l'état homozygote, des mutations tronquantes dans le gène *IQCE* et des mutations dans d'autres gènes dont la pathogénicité a déjà été rapportée dans la littérature. Pour la famille 1, la mutation homozygote dans *TULP1* est responsable du phénotype de RP (Hanein et al. 2004). Pour la famille 3, la mutation homozygote dans le gène *ATP6V1B1* explique l'atteinte rénale et la surdité (Vargas-Poussou et al. 2006). Au début du projet, *IQCE* était considéré comme un gène candidat responsable de ciliopathie syndromique : nous avons envisagé un chevauchement des phénotypes.

Nous avons effectué les expériences de caractérisation moléculaire sur le matériel biologique extrait des cellules du patient de la famille 1. Dans son cas, la délétion mène à l'apparition d'un codon stop prématuré. La protéine tronquée résultante présente une séquence réduite de moitié et n'est sûrement pas produite. Cette information n'a pas pu être confirmée à cause de l'aspécificité de l'anticorps anti-*IQCE* disponible sur le marché. *TULP1* est un gène dont l'expression est limitée à la rétine. La mutation de ce gène ne risquait donc pas de venir impacter nos expériences sur les fibroblastes. Nous avons, tout de même, confirmé l'absence d'expression du gène dans ces cellules. Nous avons suivi le logigramme d'expériences présenté en Figure 42 afin de préciser l'impact d'une perte d'expression d'*IQCE* sur les fibroblastes des patients.

Pour rappel, la protéine *IQCE* interagit et permet la concentration des protéines *EVC*, *EVC2* et *EFCAB7* au niveau de l'*EVC* zone. Nous avons fait l'hypothèse que l'absence de protéine *IQCE* allait entraîner la dispersion du complexe *EVC-EVC2*. Pour confirmer cette hypothèse, nous avons marqué la protéine

EVC2 par IF sur des cellules en conditions ciliées où le cil a été marqué avec un anticorps dirigé contre l'alpha tubuline acétylée. Nous avons compté le nombre de fois où nous pouvions visualiser un point jaune à la base du cil correspondant à une colocalisation entre les 2 protéines. Nous avons également mesuré la longueur du cil pour déterminer si la protéine IQCE est impliquée dans la formation du cil primaire.

Une étude transcriptomique par RNAseq a été mise en œuvre pour déterminer les variations du profil d'expression chez ces patients. Les différences d'expressions considérées comme significatives ont été validées par qRT-PCR. Elles concernaient des gènes connus pour être impliqués dans la voie SHH et la formation des membres : des informations obtenues à l'issue d'analyses d'enrichissement. Les variations d'expressions validées ont confirmé l'impact de la perte d'IQCE sur l'expression de gènes codants pour des protéines actrices de la voie SHH. Nous avons donc voulu explorer davantage l'impact de la perte d'IQCE sur l'activation de cette voie en réalisant une qRT-PCR et un WB à partir de culots cellulaires obtenus dans des conditions -SVF+SAG. Nous avons poursuivi la validation fonctionnelle en utilisant le modèle poisson-zèbre afin de caractériser l'impact *in vivo* de la perte d'IQCE et déterminer si le phénotype résultant concordait avec celui des patients.



Collaboration avec  
C. Etard et U. Strahle  
(Karlsruhe)

Création de poisson zèbre *IQCE*  
(par injection de morpholino)

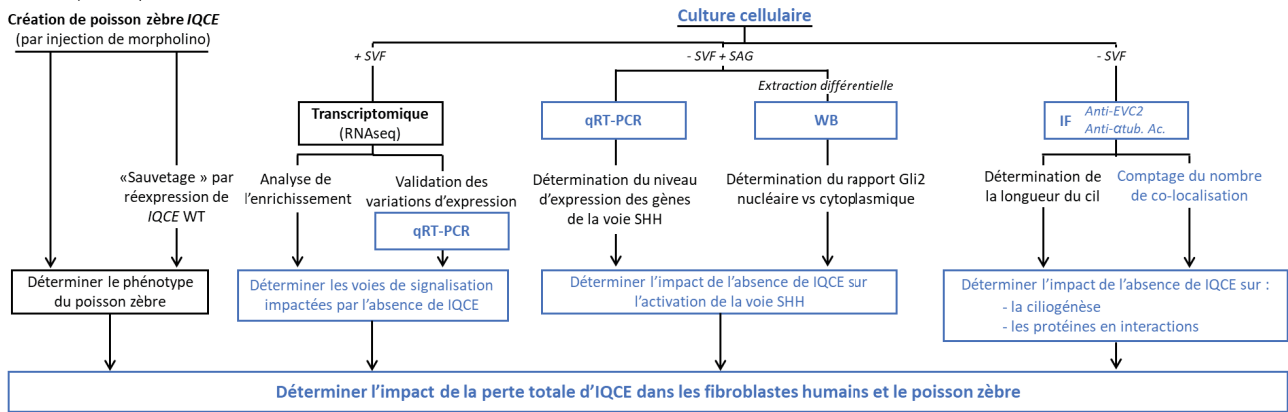
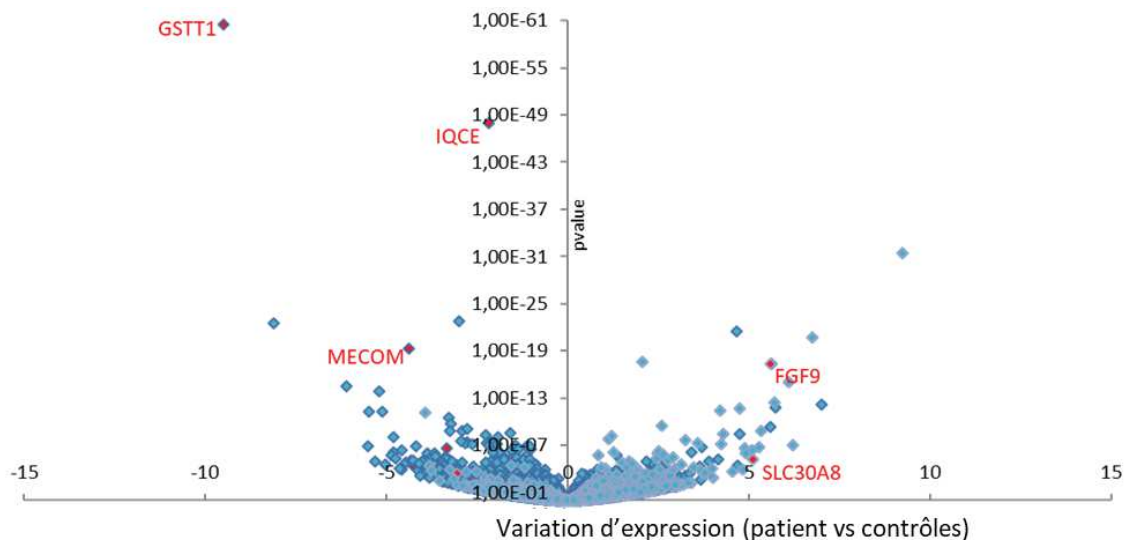


Figure 42 : Logigramme. Résumé des expériences mises en place pour répondre aux différents objectifs. Mon travail est visualisé en bleu.

## II. Les données supplémentaires.

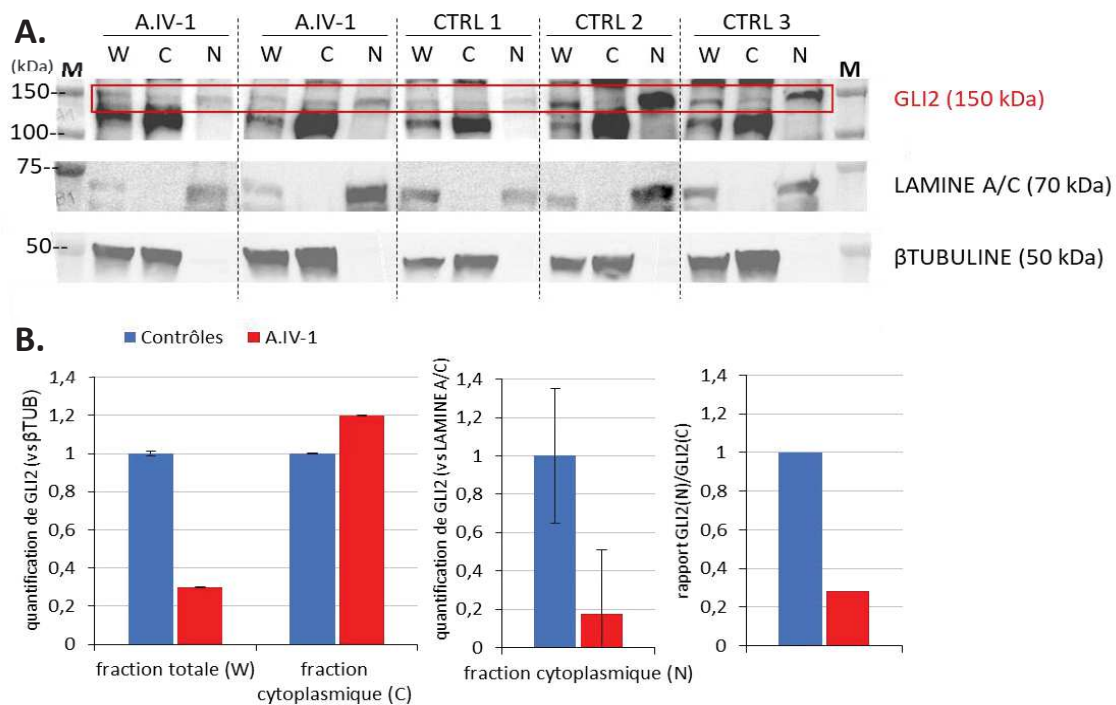
Les résultats de l'analyse transcriptomique peuvent être présentés sous la forme d'un volcano plot afin de visualiser la pvalue en fonction des variations d'expressions géniques. Ce type de représentation rend l'information plus visuelle et permet d'identifier facilement les gènes pour lesquels la variation d'expression chez le patient est significativement (partie haute) diminuée (gauche) ou augmentée (droite) par rapport aux données des contrôles. Les gènes soulignés en rouge sont ceux dont les variations d'expression ont été validées par qRTPCR.



**Figure 43 : Volcano plot.** Représentation des résultats du RNAseq.

Afin de déterminer l'impact de la perte d'IQCE sur la voie SHH, nous avons réalisé des extractions protéiques différentielles à partir de culots cellulaires cultivés en condition -SVF+SAG. Un WB a été réalisé avec les différentes fractions. Nous avons utilisé plusieurs anticorps dont un dirigé contre la protéine GLI2. Ce facteur de transcription activateur subit une translocation depuis le cytoplasme vers le noyau lorsque la voie SHH est activée. Nous avons obtenu une multitude de bandes aspécifiques pour GLI2. Pour la quantification, nous avons pris en considération la bande de 150kDa (le PM attendu) présente dans les 3 fractions. En effet, la condition -SVF+SAG permet d'activer la voie SHH : dans cet état, une bande est attendue dans la fraction nucléaire synonyme de translocation nucléaire. Les autres anticorps étaient dirigés contre des protéines marqueurs du cytoplasme (la  $\beta$ -TUBULINE) et du noyau (la LAMINE A/C). Comme attendu, pour la  $\beta$ -TUBULINE, une bande apparaît dans les fractions W et C. Pour la LAMINE A/C, les bandes apparaissent dans les fractions W et N (Figure 44). Nous avons déterminé le rapport entre les quantités de GLI2 cytoplasmiques et nucléaires pour estimer le niveau d'activation de la voie entre le patient et les contrôles. Pour le patient, le rapport est fortement diminué par rapport aux contrôles. Il semble que dans cette condition, la translocation nucléaire de GLI2 soit moindre chez le patient. Ces résultats soutiennent l'hypothèse selon laquelle la voie SHH est

sous-activée en l'absence d'IQCE. Il faudrait reproduire et optimiser cette expérience pour éliminer les marquages aspécifiques.



**Figure 44 : Expériences menées sur la voie SHH.** Résultats du WB réalisé sur les extractions protéiques différentielles issues des culots cellulaires cultivés en condition -SVF+SAG. **A.** Photos obtenues après le marquage des protéines GLI2, LAMINE A/C (référence nucléaire) et βTUBULINE (référence cytoplasmique). **B.** Résultats de la quantification de GLI2 par rapport à la βTUBULINE, la LAMINE A/C et le rapport GLI2N/GLI2C. M : marqueur, A.IV-1 : le patient, CTRL : contrôle, W : whole/protéines totales, C : cytoplasme, N : noyau.

### III. La conclusion et les perspectives.

L'ensemble du travail mis en œuvre dans ce projet a permis de montrer que la perte d'IQCE provoque la perte de localisation d'EVC2 à la base du cil primaire et mène à une sous-activation de la voie SHH dans des conditions où elle est déjà activée. L'identification de mutations dans les gènes *TULP1* et *ATP6V1B1* explique certains des traits cliniques des patients. Leur polydactylie est expliquée par des mutations identifiées dans le gène *IQCE*.

Les expériences réalisées sur les fibroblastes humains nous permettent de confirmer l'implication d'*IQCE* dans le fonctionnement du cil primaire et de confirmer qu'il s'agit effectivement d'un gène de ciliopathie. Les poisson-zèbres pour lesquels l'expression d'*IQCE* a été fortement réduite (par l'usage de morpholinos) présentent un phénotype typique de ciliopathie syndromique avec une courbature de l'axe antéro-postérieur, une désorganisation des couches de la rétine, un défaut d'asymétrie droite-gauche et une désorganisation des cils dans la lumière du pronéphros. Ces résultats montrent qu'un défaut d'*IQCE* peut conduire à un phénotype syndromique dans cette espèce. Pour rappel, le modèle

murin KO *Iqce* présente une atteinte rétinienne avec une diminution de l'épaisseur totale de la rétine et des troubles du comportement.

Il est envisageable en se basant sur les phénotypes des modèles animaux d'identifier des patients avec une atteinte rétinienne causée par des mutations dans le gène *IQCE*. Le nombre de familles rapporté à ce jour (n=4 au total) ne permet pas de savoir si cette hypothèse est réaliste ou non. Afin de déterminer si des phénotypes autres que la polydactylie pourraient être causés par des mutations tronquantes dans ce gène chez l'homme, nous pourrions générer des organoïdes de mini-rétine. Des cellules humaines pour lesquelles l'expression d'*IQCE* seraient abolie ou les cellules des patients pourraient être utilisée comme matériel biologique de base. La capacité de réponse des photorécepteurs de ces mini-rétines aux stimuli lumineux peut être testée. Il s'agit du même principe que celui d'un ERG mais appliqué à des cellules en culture (Hughes et al. 2016).

Il n'a pas été possible de confirmer l'absence de protéine IQCE chez le patient car les anticorps dirigés contre cette protéine disponible sur le marché ne fournissent pas de résultats concluants. Pour contourner ce problème, nous pourrions envisager de mettre en place plusieurs expériences. Nous pourrions faire exprimer le gène *IQCE* porteur d'une des mutations tronquantes dans des cellules contrôles où l'expression de *IQCE* (WT) serait abolie avec des si-RNA par exemple. Dans ce cas, il faudrait ajouter un « tag » (la GFP par exemple) aux protéines IQCE mutée et WT afin les localiser.

## IV. L'article





## RESEARCH ARTICLE

# Novel *IQCE* variations confirm its role in postaxial polydactyly and cause ciliary defect phenotype in zebrafish

Alejandro Estrada-Cuzcano<sup>1\*</sup>  | Christelle Etard<sup>2\*</sup> | Clarisse Delvallée<sup>1\*</sup> |  
Corinne Stoetzel<sup>1</sup> | Elise Schaefer<sup>1,3</sup> | Sophie Scheidecker<sup>1,4</sup> | Véronique Geoffroy<sup>1</sup> |  
Aline Schneider<sup>1</sup> | Fouzia Studer<sup>5</sup> | Francesca Mattioli<sup>6,7,8,9</sup> | Kirsley Chennen<sup>1,10</sup> |  
Sabine Sigaudy<sup>11</sup> | Damien Plassard<sup>12</sup> | Olivier Poch<sup>10</sup> | Amélie Piton<sup>4,6,7,8,9</sup>  |  
Uwe Strahle<sup>2</sup> | Jean Muller<sup>1,4\*</sup>  | Hélène Dollfus<sup>1,3,5\*</sup>

<sup>1</sup>Laboratoire de Génétique médicale, UMR\_S INSERM U1112, IGMA, Faculté de Médecine, FMTS, Université de Strasbourg, Strasbourg, France

<sup>2</sup>Institute of Toxicology and Genetics (ITG), Karlsruhe Institute of Technology (KIT), Karlsruhe, Eggenstein-Leopoldshafen, Germany

<sup>3</sup>Service de Génétique Médicale, Institut de Génétique Médicale d'Alsace, Hôpitaux Universitaires de Strasbourg, Strasbourg, France

<sup>4</sup>Laboratoires de Diagnostic Génétique, Hôpitaux Universitaires de Strasbourg, Strasbourg, France

<sup>5</sup>Centre de Référence pour les affections rares en génétique ophtalmologique, CARGO, Filière SENSGENE, Hôpitaux Universitaires de Strasbourg, Strasbourg, France

<sup>6</sup>Institut de Génétique et de Biologie Moléculaire et Cellulaire, Illkirch-Graffenstaden, France

<sup>7</sup>Institut de Génétique et de Biologie Moléculaire et Cellulaire, INSERM U1258, Illkirch-Graffenstaden, France

<sup>8</sup>Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS UMR 7104, Illkirch-Graffenstaden, France

<sup>9</sup>Institut de Génétique et de Biologie Moléculaire et Cellulaire, Université de Strasbourg, Illkirch, France

<sup>10</sup>Complex Systems and Translational Bioinformatics, ICube UMR 7357, Fédération de Médecine Translationnelle, Université de Strasbourg, Strasbourg, France

<sup>11</sup>Département de Génétique Médicale, Hôpital de la Timone, Marseille, France

<sup>12</sup>Plateforme GenomEast, IGBMC, Illkirch-Graffenstaden, France

## Correspondence

Jean Muller, Hélène Dollfus, Laboratoire de Génétique médicale, UMR\_S INSERM U1112, IGMA, Faculté de Médecine, FMTS, Université de Strasbourg, Strasbourg, France.  
Email: jeanmuller@unistra.fr (J. M.) and dollfus@unistra.fr (H. D.)

## Funding information

Fondation Jérôme Lejeune; Hôpitaux Universitaires de Strasbourg, Grant/Award Number: API 2013–2014 HUS N° 5885; Fondation Recherche Médicale, Grant/Award Number: ECO20170637509; Fondation Maladie Rare; Fondation JED-Belgique; Retina France, Grant/Award Number: 100 exomes Program; HGF Biointerfaces Programme

## Abstract

Polydactyly is one of the most frequent inherited defects of the limbs characterized by supernumerary digits and high-genetic heterogeneity. Among the many genes involved, either in isolated or syndromic forms, eight have been implicated in postaxial polydactyly (PAP). Among those, *IQCE* has been recently identified in a single consanguineous family. Using whole-exome sequencing in patients with uncharacterized ciliopathies, including PAP, we identified three families with biallelic pathogenic variations in *IQCE*. Interestingly, the c.895\_904del (p.Val301Serfs\*8) was found in all families without sharing a common haplotype, suggesting a recurrent mechanism. Moreover, in two families, the systemic phenotype could be explained by additional pathogenic variants in known genes (*TULP1*, *ATP6V1B1*). RNA expression analysis on patients' fibroblasts confirms that the dysfunction of *IQCE* leads to the dysregulation of genes associated with the hedgehog-signaling pathway, and zebrafish experiments demonstrate a full spectrum of phenotypes linked to defective cilia: Body curvature, kidney cysts, left-right asymmetry, misdirected cilia in the pronephric duct, and retinal defects. In conclusion, we identified three additional

\*Alejandro Estrada-Cuzcano and Christelle Etard, Clarisse Delvallée, Jean Muller, and Hélène Dollfus contributed equally to this study.

families confirming *IQCE* as a nonsyndromic PAP gene. Our data emphasize the importance of taking into account the complete set of variations of each individual, as each clinical presentation could finally be explained by multiple genes.

**KEYWORDS**

cilia, hedgehog signaling, *IQCE*, polydactyly, RNA-seq, zebrafish

## 1 | INTRODUCTION

Polydactyly is one of the most frequently inherited defects of the limb and has a prevalence of 0.3–3.6/1,000 in live births (Castilla et al., 1973; Malik, Ullah, Afzal, Lal, & Haque, 2014). It is characterized by supernumerary digits, ranging from extra-additive soft tissue without bone structure to an integral and fully developed digit (Umair, Ahmad, Bilal, Ahmad, & Alfadhel, 2018). According to the position of the extra digit, polydactyly has been classified into postaxial polydactyly (PAP), mesoaxial, and preaxial (Malik, 2014). This defect is also associated with at least 221 syndromes according to the London dysmorphology database (Phadke & Sankar, 2010), covering 146 genes (Umair et al., 2018) and can be a hallmark for a group of syndromes such as for ciliopathies. For this group of diseases, in addition to skeletal malformations such as PAP, the phenotypic spectrum can associate retinal degeneration, obesity, kidney dysfunction, and sometimes intellectual disability as observed in the classical Bardet–Biedl syndrome (BBS; MIM# 209900) for which 22 genes have been identified. In such heterogeneous diseases, high-throughput genetic testing has made possible the diagnosis for many individuals (Shaheen et al., 2016) but also revealed more complex cases. Indeed, multiloci-bearing pathogenic variations have been identified in several known diseases related to loci in single patients with apparent syndromic presentation (Karaca et al., 2018).

To date, eight genes *FAM92A1*, *GLI1*, *GLI3*, *IQCE*, *KIAA0825*, *MIPOL1*, *PITX1*, and *ZNF141* (Ullah et al., 2019; Umair et al., 2018) have been implicated in nonsyndromic polydactyly accounting for 286 disease-causing variations according to human gene mutation database (2018.1). Among those genes, IQ domain-containing protein E (*IQCE*) encodes a ciliary protein located at the base of the primary cilia and is linked to the hedgehog (Hh) signaling pathway (Pusapati et al., 2014). Hh signaling is one of many major pathways that controls key steps of embryonic development (Ingham, Nakano, & Seger, 2011), tissue homeostasis, and regeneration (Petrova & Joyner, 2014) and related to human cancers (Barakat, Humke, & Scott, 2010). Human hand and feet development is based on a highly conserved pentadactyl pattern (Abbasi, 2011), the conserved Hh signaling pathway regulates the precise digital shape and pattern in the limbs. *IQCE* together with *EFCAB7* has an important regulatory role in the EvC complex (*EVCEVCE2*), a positive tissue regulator of Hh signaling. Recently, a homozygous splice acceptor variation in *IQCE* has been described in a single consanguineous family with PAP (Umair et al., 2017).

In this study, we report three families with biallelic pathogenic variations in *IQCE* identified by whole-exome sequencing (WES).

Interestingly, these families with PAP were initially recruited as syndromic ciliopathies and two have additional pathogenic variations in other genes explaining their apparent syndromic phenotype. Functional studies based on the patient's cells or zebrafish (*Danio rerio*) assays confirm the ciliary role of *IQCE*.

## 2 | MATERIALS AND METHODS

### 2.1 | Subjects

The patients were recruited by the Strasbourg University Hospital Medical Genetics Department as well as by the Center for Rare Genetic Ophthalmologic Diseases (Hôpitaux Universitaires de Strasbourg, Strasbourg, France) because they were suspected to have a ciliopathy. Study protocols used in each cohort have been approved by the corresponding Institutional Review Board or equivalent committees (in Strasbourg, "Comité de Protection des Personnes" EST IV, N°DC-20142222), and written informed consent was given by each participant or parents. Our research complies with the Declaration of Helsinki. Written informed consent for open-access publication was provided by the participants or their parents.

### 2.2 | Whole-exome sequencing

WES was performed by IntegraGen (Evry, France) for Families A, B, and C in 2011, 2013, and 2018, respectively. After shearing of the genomic DNA, library preparation was performed using either the Agilent Human All Exon Kits v2 (Families A and B) or the Twist Human Core Exome Enrichment System (Twist Bioscience) and IntegraGen Custom additions (Family C). Sequencing was performed on an Illumina HiSeq 2000/4000 to generate 75-base pair (bp) paired-end reads following the manufacturer's protocols. Image analysis and base calling were performed using Illumina Real-Time Analysis (1.14/2.7.7) with default parameters.

### 2.3 | Bioinformatics analysis

The reads were mapped to the reference human genome (GRCh37/hg19) using BWA v0.7.12 (Etard, Roostalu, & Strahle, 2010). GATK UG v3.4-46 was used to call single-nucleotide variant (SNV) and indel variations (DePristo et al., 2011). Annotation and ranking of SNV and indel were performed by VaRank 1.4.3 (Geoffroy et al., 2015) in combination with the Alamut Batch software (1.11; Interactive Biosoftware, Rouen, France). Very stringent filtering criteria were

applied to filter out nonpathogenic variants (Table S1): (a) Variants represented with an allele frequency >1% in public variation databases including the 1000Genomes (Auton et al., 2015), the genome aggregation database (gnomAD; Lek et al., 2016), the database of genomic variants (MacDonald, Ziman, Yuen, Feuk, & Scherer, 2014), or our internal exome database, and (b) variants in 5' and 3' untranslated region, downstream, upstream, intronic, and synonymous locations without pathogenic prediction of local splice effect. Structural variants (SVs) were predicted using CANOES (Backenroth et al., 2014) and annotated by AnnotSV 2.0 (Geoffroy, Herenger et al., 2018). Our analysis considered all inheritance modes but given that two families were consanguineous, we focused on compound heterozygous and homozygous variants consistent with a recessive transmission. Variations have been classified according to the American College of Medical Genetics and Genomics guidelines (Richards et al., 2015) and only Class 3 (likely pathogenic) and Class 4 (pathogenic) have been considered.

## 2.4 | Homozygosity mapping

Based on the WES data, homozygosity mapping has been performed using PLINK 1.9 (Chang et al., 2015). We followed the parameters described by Kancheva et al. (2016). The results are available in Table S2.

## 2.5 | Cell culture

Skin fibroblasts from patient A.IV-1 and age-matched healthy control fibroblasts cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum (FCS) and 1% penicillin-streptomycin-glutamine (PSG). To induce primary cilium formation, the cells were deprived of serum by growth for 24 hr in DMEM with 1% PSG but only 0.1% FCS. To activate the Hh pathway, the cells were incubated in a DMEM medium with 1% PSG without serum (-FCS) and with smoothed agonist (SAG; 100 nM). SAG has been resuspended in dimethyl sulfoxide (DMSO) at a stock concentration of 4 mM.

## 2.6 | RNA extraction, complementary DNA synthesis, quantitative polymerase chain reaction, and TaqMan

For RNA-seq validation, cells have been cultured in normal condition DMEM containing 1% PSG and 10% FCS. Total RNA was prepared from patient A.IV-1 fibroblast pellets (technical triplicate) and three controls pellets (biological triplicate) using QIAshredder Kit (Qiagen Ref 79654) and RNeasy Kit (Qiagen Ref 74104) followed by a DNase treatment with the TURBO DNA-free™ Kit. Experiments have been repeated three times. For the Hh genes expression analysis (real-time quantitative polymerase chain reaction [qRT-PCR]), cells have been cultured in DMEM with 1% PSG (-FCS) + SAG (100 nM). Total RNA was prepared as described above and experiments have been repeated twice with two different control pellets (biological duplicate) and one patient pellet (technical triplicate). RNA integrity

was assessed by gel electrophoresis and RNA concentration by Eppendorf BioPhotometer Plus with the Hellma® TrayCell. Reverse transcription of 1 µg total RNA to complementary DNA (cDNA) was performed using the Bio-Rad iScript™ cDNA Synthesis Kit. qRT-PCR amplification was performed in a Bio-Rad CFX96™ Real-Time System using the iQ™ SYBR® Green SuperMix and primer sets (Table S3) optimized for tested targets for SYBR Green-based real-time PCR for the RT PCR. The normalized fold expression of the target gene was calculated using the comparative cycle threshold ( $C_t$ ) method by normalizing target messenger RNA (mRNA)  $C_t$  to those for both GAPDH and HPRT reference gene using the CFX Manager Software Version 1.5 and excel calculation. Statistical analysis has been done using the "GraphPad prism" software.

## 2.7 | Immunocytochemistry and fluorescence microscopy

A.IV-1 and control-ciliated fibroblasts were seeded on 12 mm diameter cover slips. Ciliogenesis has been induced by 24 hr incubation in a DMEM with 1% PSG medium, without serum (-FCS). Then, cells have been fixed with 4% paraformaldehyde-phosphate-buffered saline (PFA-PBS). Subsequently, the cells were stained with the ciliary markers antiacetylated  $\alpha$ -tubulin (Lys40) mouse monoclonal primary antibody (1:200, #32-2700; Invitrogen, Carlsbad, CA) and anti-EVC2 rabbit polyclonal primary antibody (1:250, #ab198930; Abcam). As secondary antibody, donkey anti-mouse Alexa 594 (1:1000; Invitrogen) and donkey anti-rabbit DyLight 488 (1:1000, #611-741-127; Rockland, Tebu-bio) were used. Ciliated cells count between A.IV-1 and control were performed manually using ~100 cells per genotype (exact Fisher test was applied). Counts of EVC2 colocalization at the base of the cilia were performed manually using also ~100 cells per genotype in triplicate (Student's *t* test has been applied). Pictures were taken with a Zeiss Axio Imager Z2 and the analysis with the ZEN2012 software (Carl Zeiss Inc., Oberkochen, Germany).

## 2.8 | Transcriptome analysis

RNA samples were extracted from two fibroblast replicates of A.IV-1 (A.IV-1.1 and A.IV-1.2), and replicates from six individuals affected by other ciliopathies or intellectual disabilities of known molecular origin (ARN1.1, ARN1.2, ARN2.1, ARN2.2, ARN3.1, ARN3.2, ARN4.1, ARN4.2, ARN5.1, ARN6.1, and ARN6.2) using TRI Reagent® (Molecular Research Center) or using the RNeasy Mini Kit Qiagen®. Both protocols included an additional step of DNase I recombinant treatment (Sigma-Aldrich®). The integrity of the RNA was visualized on a 1% bleach agarose gel by electrophoresis. Quantification and further quality analysis were performed using the NanoDrop®. Samples should have a 260/280 ratios ~2 and a 260/230 ratio >1.7. The integrity and quality of the RNA were also evaluated by running samples on an RNA 6000 Nano Chip on the Bioanalyzer (Agilent Technologies) and samples should have an RNA integrity number  $\geq 8$ . Library preparation was performed at the GenomEast platform at the Institute of Genetics and Molecular and Cellular

Biology (Strasbourg), using the TruSeq® stranded mRNA Sample Preparation Kit (Illumina) starting from 1 µg of extracted total RNA. Libraries were then 100 + 75 bp paired-end sequenced, with two samples per lane on an Illumina HiSeq 4000 sequencer. The corresponding bioinformatics pipeline is described in detail in the Supporting Information.

### 2.8.1 | Zebrafish stocks

Fish were bred and raised at 28.5°C as described previously (Westerfield, 1993). The AB wild-type line (European Zebrafish Resource Centre, Karlsruhe, Germany) was used for all the experiments. Zebrafish husbandry and experimental procedures were performed in accordance with the German animal protection regulations (AZ35-9185.81/G-137/10; Regierungspräsidium, Karlsruhe). Fish were crossed pairwise. For experiments, fertilized eggs were raised in 1X Instant Ocean salt solution (Aquarium Systems, Inc.) supplemented with 200 µM 1-phenyl 2-thiourea to suppress melanogenesis. For visualization of situs inversus, we used a transgenic line expressing GFP under the control of the skeletal and cardiac muscle-specific promoter *unc45b* (Roostalu & Strahle, 2012).

### 2.8.2 | In situ hybridization

Whole-mount in situ hybridization was performed as previously described (Costa, Escalera, Rodrigues, Manasfi, & Mermelstein, 2002; Oxtoby & Jowett, 1993). *Ptc1* probe was obtained from Concordet et al. (1996). *wt1a* Probe was obtained from Armant et al. (2013).

### 2.8.3 | Morpholinos and microinjections

A search of the zebrafish genome (GRCz11) revealed a single orthologue of the human *IQCE* gene (RefSeq: NM\_001287204.1) encoding a protein sharing 58% amino acid similarity with the human *IQCE* protein. Injections were performed as described previously (F. Muller et al., 1999). In brief, zebrafish eggs were collected shortly after being laid. Cleaned eggs were transferred to a Petri dish with a minimal amount of water. Embryos were injected (FemtoJet; Eppendorf) through the chorion into the yolk at the one-cell stage. Injection needles were pulled from borosilicate glass capillary tubes with filament (Warner Instruments) using a micropipette puller (Sutter Instrument Co.). Morpholinos were designed against the start codon of *iqce* (*iqce-mo*). Morpholino controls consist of the 5 bp mismatch *iqce-mo* (*iqce-mocont*). Morpholinos (Gene Tools, LLC) were injected at the following concentrations: *iqce-Mo*: CAAGTTCTCCA GCTACCACAGACAT (0.6 µM); *iqce-mocont*: CAACTTgTcGAGCTACC AgA cACAT (0.6 µM). All dilutions were made in distilled water. Phenol red was added to the samples before injection (0.1% final concentration). *iqce-gfp* plasmid was injected at the final concentration of 40 ng/µl. When injected at 0.6 µM, neither the anti *iqce* nor the *iqce-mocont* (control) morpholino elicited necrosis, an unspecific effect of morpholino injections occasionally observed with some

morpholinos (Robu et al., 2007). The efficiency of *iqce-mo* to block translation was assessed by coinjection of *iqce-mo* with a plasmid containing the *iqce* cDNA in-frame with *gfp* (*iqce-gfp*). As expected, the embryos were depleted from GFP expression, indicating that the morpholinos indeed target efficiently *iqce*. In contrast, the five-mismatch *iqce-mocont* was unable to abrogate the translation of the fusion gene (Figure S3a).

### 2.9 | PCR amplification

Total RNA was isolated from 24 to 72 hour post fertilization (hpf) embryos using TRI Reagent (Invitrogen). RT-PCR was carried out following standard protocols. For *iqce* in situ hybridization probe, a fragment of 580 bp was amplified by PCR with the following primers: *iqce*5': GCTGCATCATGTAGGACGATG, *iqce*3': CACTCTGAAGCTG ACTCAAG, and cloned into pGEM-T Easy vector (Promega). *iqce* full-length sequence was amplified from zebrafish cDNA with the following primers: *iqce*5': tcgGAATTCATGTCTGTGGTAGCTGGAGAAC; *iqce*3': agactcgaggTAAATCAAGTCA TCGGCACTG and cloned into *pcs2* + GFP with *EcoRI-XhoI*. For the rescue experiment *iqce5' modif*: tcgGAATTC ATGTGGTGTAGCGGGCGAAGTACTAG was used in combination with *iqce3'*, and cloned into *pcs2* + GFP morpholino-resistant *iqce* mRNA was synthesized with the SP6kit (Promega).

### 2.10 | Immunohistochemistry

Fish embryos were fixed in 4% PFA/1X PBS 1 hr at room temperature. Embryos were permeabilized 7 min at -20°C with cold acetone, rinsed five times, 5 min with 1X PBS and blocked 4 hr with binding protein dependent (1% DMSO, 5% BSA, and 1X PBS). Incubation with primary antibody occurred overnight at 4°C, followed by thorough washing with PTw (PBSTween: 1X PBS, 0.1% Tween 20). Embryos were incubated with secondary antibody 2 hr at room temperature, and washed with PTw. Antibodies dilutions: Acetylated-tubulin (Sigma-Aldrich): 1/500; engrailed (*eng*; Developmental Studies Hybridoma Bank, The University of Iowa): 1/50; and goat anti-rabbit Alexa 488 (Invitrogen): 1/1000; and goat anti-mouse Alexa 488 (Invitrogen): 1/1000. For nucleus staining, we used Draq5 at 1/1000 dilution. Staining was analyzed by confocal microscopy (TCS SP2; Leica). For sectioning of the eyes, embryos were fixed in 4% PFA/PBS, progressively dehydrated in ethanol and stepwise transferred into Epon resin. Five-micrometer sections were cut with a Leica microtome and stained with toluidine blue.

## 3 | RESULTS

### 3.1 | Clinical features

Three families with ciliopathy like phenotypes were referred to our laboratory as they remained unsolved following standard genetic testing. They all shared PAP as a common feature but each presented other symptoms as described below (Figure 1a and Table 1). The affected patient of Family A (A.IV-1) was born after an unremarkable

pregnancy and is the only child of related Algerian parents. Bilateral polydactyly and brachydactyly of hands and feet were noted at birth associated with syndactyly of the second and third digits of the feet. At the age of 6 years, he was diagnosed with a decreased visual acuity and his ophthalmologic evaluation revealed retinal degeneration. He had hypogonadism and overweight was noted during his last clinical evaluation (15 years old). Family B is composed of three children presenting with PAP of hands and feet (bilateral for B.II-1 and unilateral for B.II-2 and B.II-3). The older one and the twin sister presented overweight, respectively, at the age of 7 and 5 years in the context of obese parents. The ophthalmologic evaluation, including electroretinography, was normal for all the siblings. In Family C, two affected children were reported. Individual C.IV-3 is the last child of related Moroccan parents and presented with PAP of both hands and one foot, learning disabilities, and overweight. Her two sisters and parents share also the two last features. Her cousin, patient C.IV-4, born also from related parents is more severe. He had also PAP of the four limbs associated with renal anomalies, deafness, and speech delay. The observed kidney disease consisted of tubular nephropathy with nephrocalcinosis diagnosed at 1 year old evocative of a distal renal tubular acidosis. Deafness was detected at the age of 2 years. His development was in the normal range, except for a language delay, in the context of hearing loss, which improved with speech therapy (Table 1).

The association of PAP and the other clinical features of each family member was assumed to be compatible with a syndromic ciliopathy (Wolff et al., 2004) and, in particular, to the BBS (MIM# 209900). In line with the described phenotypes, the three families have been screened without success for the known BBS genes at that time (*BBS1–BBS18*) using either Sanger or targeted exome sequencing (J. Muller et al., 2010; Redin et al., 2012).

### 3.2 | Identification of *IQCE* pathogenic variations

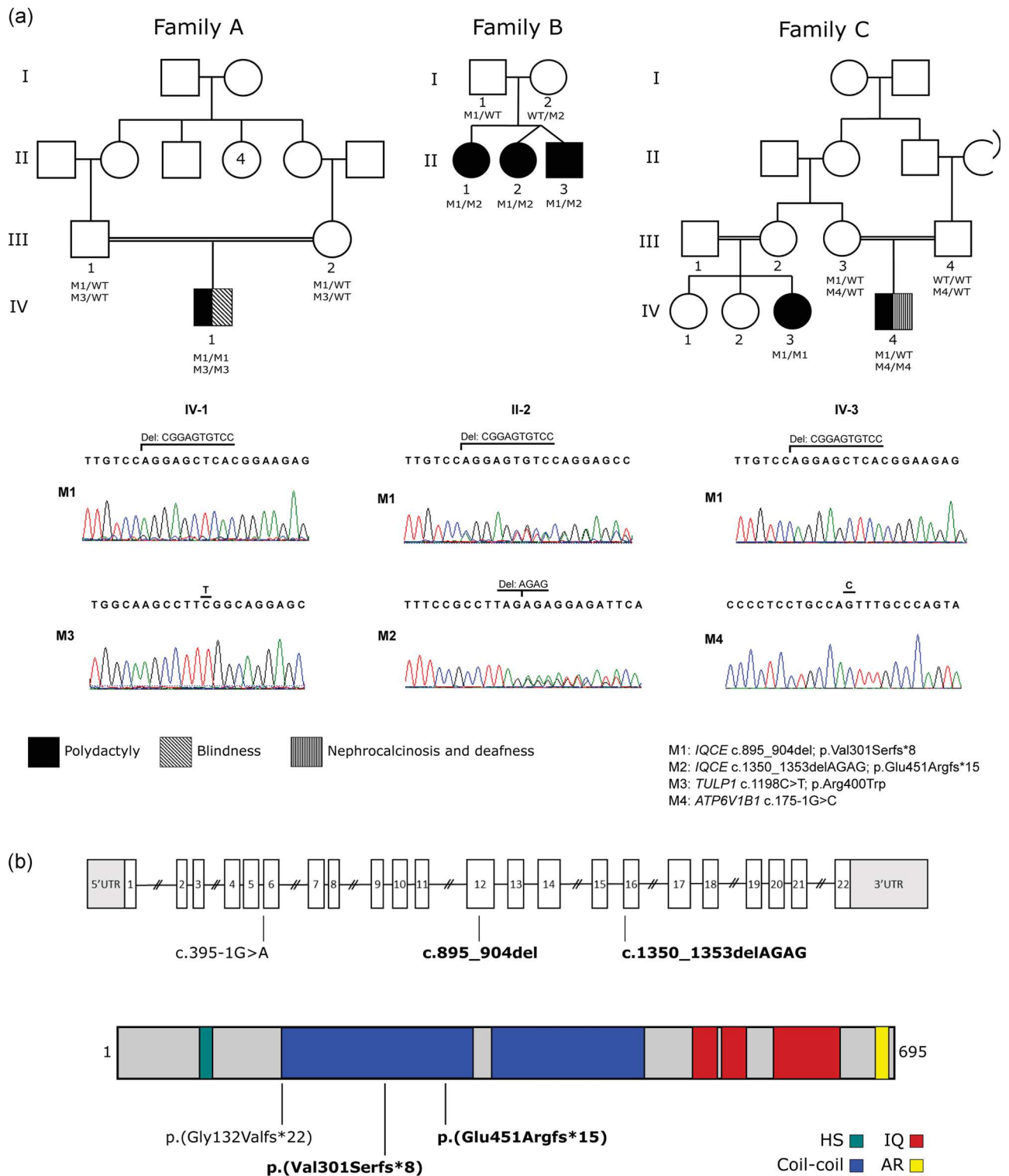
WES was applied to all three families, generating on average a total of 81,000 SNVs, 11,400 insertions–deletions (indels), and 21 SVs (Table S1). In Family A, as the proband (A.IV-1) was supposedly born from consanguineous parents, we identified 39 sizeable homozygous regions encompassing 238 Mb (Table S2); supporting the assumption of consanguinity. In this case, autosomal recessive inheritance is more likely, and we, thus, focused on homozygous variants, reducing the number of variants to 7 in six genes (*DNAAF1*, *IQCE*, *MADD*, *SHANK2*, *TULP1*, and *ZDHHC2*; Table S1). A.IV-1 is affected by polydactyly, retinal dystrophy, and hypogonadism, which made us concentrate on two genes. The first one is *IQCE* (NM\_152558.5), a gene that was just described as associated to PAP (Umair et al., 2017), with a homozygous 10 bp deletion in exon 12:c.895\_904del (p.Val301Serfs\*8; Figure 1b) located in a conspicuous homozygous region of 2.4 Mb. This variant is present in gnomAD at a very low combined allele frequency (0.0009534) with two homozygous carriers, which is compatible with the frequency of the disease. Subsequently, we analyzed *IQCE* exons 12 in 96 matched controls of Algerian ethnicity and could not find this variant (Supporting

Information Methods). Interestingly, the second gene is *TULP1* (NM\_001289395.1), a gene reported in Leber congenital amaurosis or retinitis pigmentosa (RP; Hagstrom, North, Nishina, Berson, & Dryja, 1998; Hanein et al., 2004) for which we identified a known homozygous pathogenic (Class 5) variation, c.1198C>T (p.Arg400Trp), also located in a homozygous region of 8.5 Mb. This variant does not appear in public databases (gnomAD, 1000Genomes). The variants in both genes (*IQCE* and *TULP1*) segregated with the phenotypes within the family (Figure 1b) and in combination explained most of the clinical features of this individual. In the nonconsanguineous Family B (Figure 1a), WES revealed only two Class-4 variations in *IQCE* at compound heterozygous state: c.895\_904del (p.Val301Serfs\*8) in exon 12 and c.1350\_1353delA-GAG (p.Glu451Argfs\*15) in exon 16 (Figure 1b). Both variations segregated in all affected members (B.II-1, B.II-2, and B.II-3). The second variation (c.1350\_1353delA-GAG) was found in gnomAD with a frequency of 0.0002381 with 66 heterozygous carriers. In Family C, WES analysis of the most severe case, C.IV-4 (Figure 1a), revealed a known homozygous variant in *ATP6V1B1* (NM\_001692.3), a gene that is associated with distal renal tubular acidosis with deafness (Vargas-Poussou et al., 2006; distal renal tubular acidosis [dRTA]; MIM# 267300). This Class-4 variant, c.175-1G>C, is located in a canonical splice site and is predicted to cause the skipping of exon 3, this variation was not found in public databases (gnomAD, 1000Genomes). Interestingly, the c.895\_904del (p.Val301Serfs\*8) variation in *IQCE* that has been already observed in Families A and B (respectively, at the homozygous and compound heterozygous state) was only found in C.IV-4 at the heterozygous state (whereas no second pathogenic allele could be identified by cDNA sequencing and quantification, data not shown). The WES did not reveal any other variations in known PAP genes for this individual. Given that patient C.IV-3 had only PAP and given the C.IV-4 variation in *IQCE*, we sequenced the *IQCE* gene in C.IV-3 and surprisingly found the c.895\_904del (p.Val301Serfs\*8) at the homozygous state (Figure 1).

Given the recurrence of the c.895\_904del (p.Val301Serfs\*8) variant in our cohort, we investigated a possible founder effect. Nevertheless, no common haplotype could be observed in the corresponding region in Families A and B (Table S4). Analysis of the gnomAD record of this variation revealed its presence in different ethnical groups at a very low allele frequency: European (221/126,532 alleles), Ashkenazi (12/10,134 alleles), Latino (17/34,392 alleles), African (6/23,966 alleles), Finish (2/25,788 alleles), and South Asian (1/30,782 alleles). The diverse origin of individuals with this *IQCE* variant and the differences in the haplotype suggested a recurrent mutational mechanism.

To better understand the effect of *IQCE* variants, we investigated the expression and subcellular localization of *IQCE* interacting partners in patient cells. *IQCE* is ubiquitously expressed in human adult tissues with higher levels of mRNA expression in lung, brain, and hippocampus (Figure S1a). Endogenous *IQCE* is located at the base of the cilia and forms a complex with EFCAB7 to anchor the EVC–EVC2 subcomplex at the base of the cilia (EvC zone) for the propagation of Hh signaling (Pusapati et al., 2014). We investigated





**FIGURE 1** Clinical description and molecular genetic characterization of *IQCE* variations. (a) Graphic representation of three autosomal recessive families with complicated polydactyly in which *IQCE*, *TULP1*, and *ATP6V1B1* variations were identified. Sanger sequencing traces confirming the presence of the pathogenic variant and segregation in available family members are demonstrated. The “WT” symbol represents wild-type and “M” symbol represents pathogenic variant. Blackened symbols indicate affected individuals. (b) Genomic structure (NC\_000007.14) and schematic representation of the protein layout (NP\_689771.3) of *IQCE* and location of a currently known disease-associated variant (in bold, novel pathogenic variants identified in this study). Topology of *IQCE* represents one hydrophobic segment (HS), coil-coil conserved domain regions, three IQ calmodulin-binding motifs (IQ), and one acid region (AR). UTR, untranslated region

**TABLE 1** Clinical characteristic of individuals with variants in *IQCE*

	Family A		Family B			Family C	
	A.IV-1	B.II-1	B.II-2	B.II-3	C.IV-3	C.IV-4	
<i>IQCE</i>	p.[Val301Serfs*8]; [Val301Serfs*8]	p.[Val301Serfs*8]; [Glu451fs*15]	p.[Val301Serfs*8]; [Glu451fs*15]	p.[Val301Serfs*8]; [Glu451fs*15]	p.[Val301Serfs*8]; [Val301Serfs*8]	p.[Val301Serfs*8];[=] ATP6V1B1 c.[175-1G>C]; [175-1G>C]	
Other gene	<i>TULP1</i> p.[Arg400Trp];[Arg400Trp]						
Clinical features							
Hands and feet abnormalities							
Postaxial polydactyly	Y	Y	Y	Y	Y	Y	
Hands	Y	Y	Y*	Y*	Y	Y	
Feet	Y	Y	Y*	Y*	Y	Y	
Brachydactyly	Y	Y	Y	Y	N	N	
Syndactyly	Y	N	N	Y	N	N	
Neurosensory disease							
Retinal dystrophy	Y	N	N	N	N	N	
Deafness	N	N	N	N	N	Y	
Pondostatural evaluation							
Overweight	Y	Y	Y	N	Y	N	
Development psychomotor							
Intellectual abnormalities	N	N	N	N	N	N	
Learning disabilities	N	N	N	N	Y	N	
Speech delay	Y	N	N	N	Y	Y	
Ataxia	N	N	N	N	N	N	
Kidney function							
Renal abnormality	N	N	N	N	N	Tubulopathy-nephrocalcinosis	
Genital abnormalities							
Hypogonadism	Micropenis	N			N	N	

Abbreviations: Y, yes; N, not present; \*, unilateral.



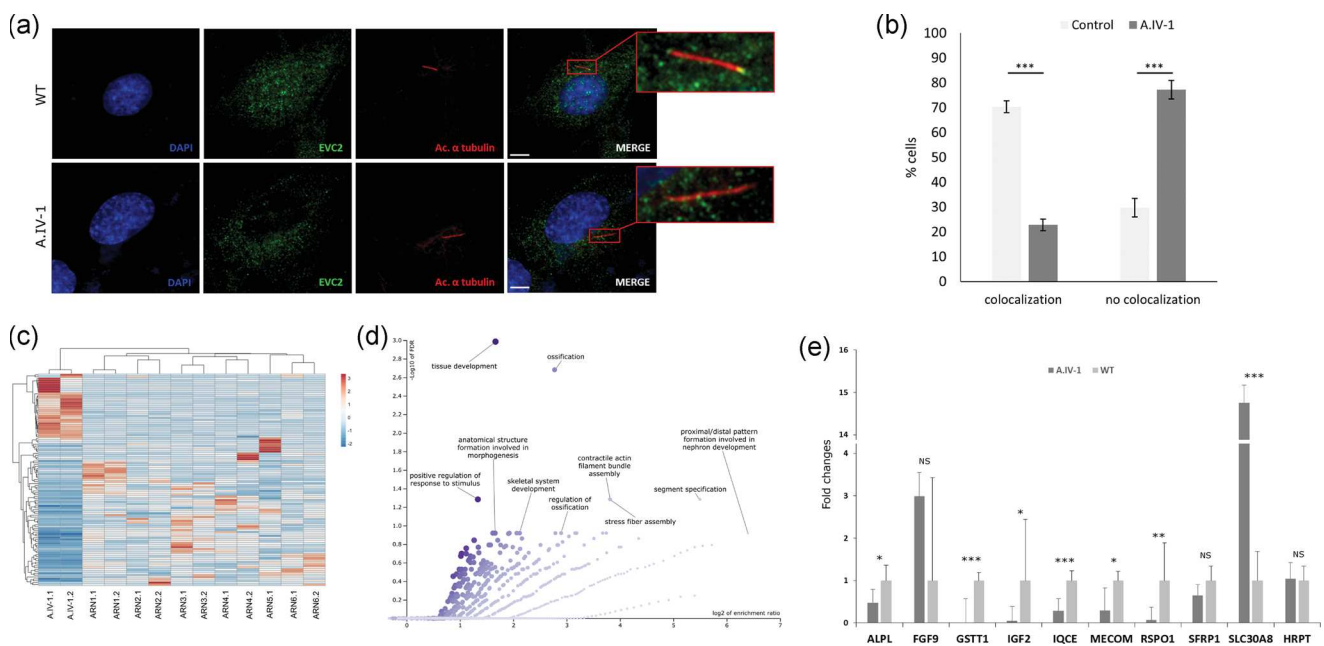
the impact of IQCE deficiency in the formation of the primary cilia and localization of EVC–EVC2 subcomplex. Immunohistochemical studies on ciliated fibroblasts from individual A.IV-1 showed no differences in the number of ciliated cells (Figure S1b). Nevertheless, for 77% of patient's cells, we did not observe colocalization of EVC2 at the base of the cilia (Figure 2a,b) that presumably results in mislocalization of the EVC–EVC2 subcomplex and may compromise the Hh signaling.

### 3.3 | Transcriptome analysis and Hh signaling effect

To get insight into the mechanistic impact of IQCE, we explored the transcriptome variation of individual A.IV-1 by RNA sequencing. We considered the possible impact of *TULP1* variation. However, under physiological conditions, *TULP1* is expressed exclusively in the retina, where it is involved in protein trafficking which is essential for the transport of rhodopsin (North, Naggert, Yan, Noben-Trauth, & Nishina, 1997). This was confirmed as no *TULP1* expression could be observed in the patient's fibroblasts. Given the *IQCE* genotype (frameshift) for A.IV-1, no *IQCE* expression could be detected. RNA differential expression analysis resulted in 98 significantly differentially expressed genes (Figure 2c and Table S5). None of the known *IQCE* interactors—*EVC*, *EVC2*, *EFCAB7*, and *TCC23* (Breslow et al., 2018; Pusapati et al., 2018)—showed

significant variations in their expression levels neither did the downstream effectors of the Hh signaling pathway *GLI1*, *GLI2*, and *GLI3*. We identified among the most significantly up and downregulated genes *SLC35D3*, *SLC30A8*, and *GSTT1*, respectively. *SLC35D3* has been shown to be implicated in metabolic syndrome and obesity (Z. Zhang et al., 2014), and *SLC30A8* is associated with a protective role against Type 2 diabetes (Flannick et al., 2014).

Regarding *GSTT1*, the protein catalyzes the conjugation of glutathione to a wide range of potential toxins as the first step in detoxification and is identified as a risk factor for open-angle glaucoma or ventricular diastolic dysfunction (Singh, Kumar, Tewari, & Agarwal, 2018; Stamenkovic et al., 2018). To obtain functional clues about the *IQCE* protein modules, we applied functional enrichment analysis (WebGestalt 2019; Wang, Vasaiakar, Shi, Greer, & Zhang, 2017); the overrepresentation analysis identified 10 significant clusters involving tissue development, ossification, and skeletal system development among others (Figure 2d). A number of genes associated with Hh signaling and limb formation are included in these protein clusters (*ALPL*, *ALDH1A1*, *FGF9*, *HOXC10*, *IGF2*, *IGFBP1*, *MECOM*, *MLLT3*, *NDP*, *SFRP1*, and *SMOC1*; Table S6). Selected genes with significant  $|\log_2$  fold change| >0.5 were tested compared with controls by RT-qPCR, and we confirmed the expression change of *ALPL*, *GSTT1*, *IGF2*, *IQCE*, *MECOM*, *RSPO1*, and *SLC30A8* but we did not detect significant differences in *FGF9* and *SFRP1* (Figure 2e).



**FIGURE 2** (a) Immunostaining of A.IV-1 and control-ciliated fibroblasts, antibodies against human EVC2 (in green), and the ciliary marker acetylated alpha tubulin (in red) revealed that in absence of endogenous *IQCE*, EVC2 fails to localize at the base of the cilia. (b) Count of mislocalized EVC2 at the base of the cilia in fibroblasts under ciliated conditions (–FCS) between A.IV-1 and control (three times ~100 cells per condition). Statistical significance was determined using Student's *t* test,  $p = .0004$  (\*\*\*). (c) Heat map generated using ClustVis (Metsalu & Vilo, 2015) from the transcriptome sequencing data reflecting gene counts of the significantly differentially expressed genes across patient and control samples. (d) Volcano plot of overrepresentation enrichment analysis of significantly misregulated transcripts in A.IV-1 fibroblast compared with wild-type. These results were obtained using WebGestalt 2019 (Table S6). (e) Real-time quantitative polymerase chain reaction analysis from three technical replicates of A.IV-1 and three controls fibroblasts showed relative expression levels (normalized to GAPDH) of nine significantly differentially expressed genes, data include  $\pm$  SD. Scale bar = 5  $\mu$ m. \* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$  (alpha risk, 5%). ns, not significant, SD, standard deviation

To assess whether the Hh signaling is affected, we determined the expression levels of key genes in the patient's cells both under ciliary conditions and Hh activation (–FCS + SAG treatment). qRT-PCR revealed a significant underexpression of genes involved in the EVC-zone complex including *EFCAB7*, *EVC*, *EVC2* and also *GLI2* and *GLI3* (Figure S2).

### 3.4 | Lack of *iqce* causes ciliary defects in zebrafish embryos

Considering the ciliary localization of the human IQCE and encouraged by the previous studies, where genetic screens in zebrafish identified numerous defects of subpopulations of cilia (Zhao & Malicki, 2007), we investigated expression and function of *iqce* in zebrafish (*Danio rerio*). We first assessed the *iqce* expression in the developing embryo. In situ hybridization of *iqce* mRNA showed ubiquitous expression at the 10 somites and 20 hpf stages with higher expression in the pronephric ducts at 20 hpf. At 48 hpf, higher levels of *iqce* mRNA were found in the pronephric ducts and in the mid- and hindbrain (Figure S3b).

A morpholino directed against the *iqce* start codon (*iqce-mo*) and a morpholino with a five mismatch control sequence (*iqce-mocont*) were used to elucidate the function of *iqce* in zebrafish embryos. Embryos injected with *iqce-mo* (77%,  $n = 126$ ) but not with the *iqce-mocont* control (2%,  $n = 221$ ) showed ventral curvature of the body axis (Figure 3a,b) which is often found in cilia mutants (Figure S3c). Occasionally we observed hydrocephalus (Figure 3b). To prove that the observed phenotype is due to the specific knockdown of *iqce*, we coinjected *iqce-mo* together with *iqce* mRNA resistant to *iqce-mo*. Fifty-four percent of embryos exhibited a characteristic curved phenotype compared with 77% for *iqce-mo* alone (Figure 3d,  $n = 193$ ). We conclude from this experiment that the *iqce-mo* effectively abrogate the *iqce* expression.

Bilaterally dilated kidney cysts were present in 37% of *iqce-mo* morphants, but neither in uninjected nor in *iqce-mocont* injected embryos (Figure 3c and Figure S3d). These cysts are localized within the glomerulus as indicated by *wt1a* staining (Figure S3e).

We next examined the cilia of the pronephric duct. Embryos at 48 hpf were stained with acetylated-tubulin antibody. In *iqce-mocont* and uninjected embryos, cilia are aligned along with the duct forming a thin line (Figure 3e,g). In contrast, the cilia organization in *iqce* morphants was severely perturbed showing cilia oriented in all directions (Figure 3f). The misorientation and shortening of pronephric duct cilia can be due to dilation of the tubules and suggest that pronephric fluid flow may be affected (Zhao & Malicki, 2007). We next tested whether other cilia are also affected, the dense ring of cilia on the epithelium surrounding the olfactory placode of 48-hpf wild-type larvae was strongly reduced in morphants (Figure 3h–j). In addition, the photoreceptor outer segment connects to the cell body via a narrow constriction that forms around a ciliary axoneme, and is known as the connecting cilium. Defects of photoreceptor cilia are associated with outer segment abnormalities and degeneration (Zhao & Malicki, 2007). Therefore, we analyzed eye sections of 4

day old morphants. In noninjected embryos or control morphants, all the layers are properly formed (Figure 3k). In contrast in *iqce* morphants, the lamination is disrupted and the retinal layer altered with no visible photoreceptor outer segment (Figure 3l).

The perturbation of motile cilia often results in a problem of left–right asymmetry. We, thus, examined the percentage of situs inversus in *iqce* morphants. *iqce-mo* or *iqce-mocont* were injected into embryos carrying a transgene that expresses GFP in the heart (–1.8 *unc45b*-TFP; Roostalu & Strahle, 2012). This allowed us to observe the looping of the heart. Forty-four percent of morphant embryos exhibited situs inversus (with the ventricle on the right and atrium on the left) or heterotaxia demonstrating that cilia from the Kupffer's vesicle, the fish equivalent of the mouse node, were affected by the *iqce* knockdown. In contrast, only 3% and 6% of uninjected and *iqce-mocont* injected embryos, respectively, showed reverse heart jogging (Figure 3m).

In mutants with defective cilia (*iguana*) or ciliary motility (*lrrc50*), otoliths are frequently found ectopically, untethered, or fused (Stooke-Vaughan, Huang, Hammond, Schier, & Whitfield, 2012). We, thus, examined the number of otoliths in *iqce* morphants at 2 dpf. At 48 dpf, each ear should contain two otoliths and it is the case for 86% of the morphant. However, for 14% of them, one of the two ears presented either one extra otolith (three instead of two) or only one (one instead of two; data not shown,  $n = 262$  embryos).

Cilia are necessary for normal activation as well as repression of the Hh signaling pathway (Goetz & Anderson, 2010; Huang & Schier, 2009; Wolff et al., 2004). It was shown that in the absence of the ciliary protein BBS1 and under stimulation of Hh, the ciliary localization of smoothed (*Smo*) was further enhanced compared with wild-type cells (Nozaki, Katoh, Kobayashi, & Nakayama, 2018). We, thus, examined the expression of the Hh receptor *Patch1* (*Ptc1*) and the downstream gene *eng* in the developing somites in *iqce* morphants. At 10 somites stages, transcription of *ptc1* was expanded to the myotome, whereas it is exclusively restricted to the adaxial cells in control or uninjected embryos (Figure 3n–p and Figure S3f). Similarly, the domain of *eng* expression was significantly expanded among the fast muscle fibers by 24 hpf, indicative of an increased number of Hh-dependent muscle fast fibers (Figure 3q–s). These results indicate that the Hh pathway is perturbed by the loss of *iqce* function.

## 4 | DISCUSSION

An IQCE homozygous variant affecting the splicing of the gene was recently identified by WES in a single Pakistani family presenting with PAP Type A limited to lower limb (Umair et al., 2017). Here, we report three additional families with pathogenic variations in IQCE referred initially as ciliopathies with PAP involving the upper and lower limbs. Interestingly, the three families shared the same c.895\_904del (p.Val301Serfs\*8) variation in IQCE exon 12 in a homozygous or in a heterozygous states with another truncating variant. These two IQCE variations have already been observed in the

largest public databases. The c.1350\_1353del (p.Glu451Argfs\*15) was described in exome variant server (Tennessen et al., 2012) with an average allele frequency >1.5% with 90 homozygous calls. Initially, this overestimated allele frequency prevented its identification as candidate pathogenic variation. The higher false-positive rate for indel calls was a known issue solved, while integrating into ExAC/gnomAD (0.024% and no homozygous calls).

The index case A.IV-1 had a suspected ciliopathy-like phenotype, due to the presence of RP, hypogonadism, polydactyly, and speech delay. Although the proband was assumed as a patient with one syndrome, the occurrence of pathogenic variations in both *IQCE* and *TULP1* can explain independently the polydactyly and the RP. In Family C, the phenotypes of the patients were different; C.IV-3 has only PAP and overweight compared with her cousin

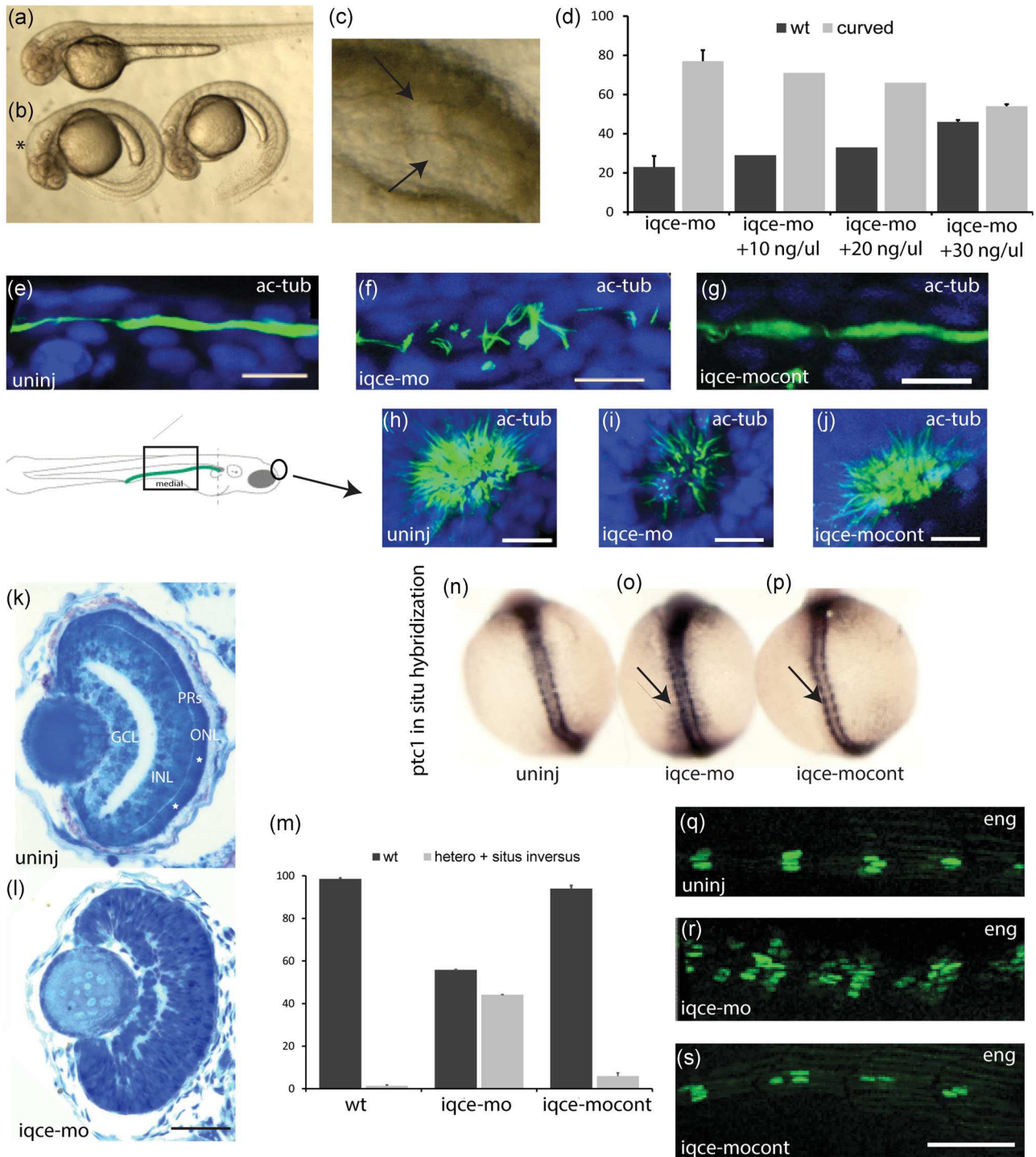


FIGURE 3 Continued.



who has PAP, deafness, and renal anomalies. Phenotypic variations in a single family are not unusual in ciliopathies even with a single gene involved (Abdelhamed et al., 2013; Gonzalez-Del Pozo et al., 2014; Katsanis et al., 2001; Ware, Aygun, & Hildebrandt, 2011; W. Zhang et al., 2018). The variations in *IQCE* can clearly explain the isolated PAP in C.IV-3. However, in the C.IV-4 individual, the association of deafness and nephropathy could finally be explained by the identification of variations in the *ATP6V1B1* gene (dRTA; MIM# 267300), known to be responsible for renal tubular acidosis associated with deafness. Nevertheless, the identification of only a single pathogenic allele in *IQCE* and the absence of any other variation in known polydactyly genes do not provide any explanation to date for the PAP for this patient. Although gene dosage and RNA expression have been performed for this individual revealing still 50% of the *IQCE* expression level (data not shown), one cannot rule out missed variation(s) either in *IQCE* or in another gene. Whole-genome sequencing might reveal intronic, intragenic, or structural variations missed by the WES (Geoffroy, Stoetzel et al., 2018). These two examples illustrate well how the incidental co-occurrence of two different conditions can mimic the clinical features of a known Mendelian disease (DeLuca, Weed, Haas, Halder, & Stone, 2015; Karaca et al., 2018). This point out also the importance of searching for multiple hits in WES or whole-genome sequencing data to explain some apparent syndromic phenotype, such as ciliopathy-like phenotype. The occurrence of the other clinical manifestations, overweight, or learning disability in these families seemed to be due to other factors than the presence of neurosensory impairment (RP or deafness). However, we should take into consideration the appreciable variations of systemic features in individuals affected by pathogenic variants in different ciliopathy genes (Carmi, Elbedour, Stone, & Sheffield, 1995; Estrada-Cuzcano et al., 2012; Waters & Beales, 2011).

Vertebrate limb patterning is controlled by two signaling centers, the zone of polarizing activity (ZPA) and the distal posterior mesenchyme, and the apical ectodermal ridge (AER) at the distal limb bud. The ZPA and AER produce signaling molecules—Hh and FGFs, respectively—that control the formation of anteroposterior and proximodistal limb axes in

a cooperative manner (Yamamoto et al., 2012). Additionally, Hh can induce a positive-feedback insulin-like growth factor (Igf) signaling during osteoblast differentiation (mostly Igf2), regulating in this way the long bone growth (Fisher, Meyer, Garber, & Dealy, 2005; Shi, Chen, Karner, & Long, 2015). Until recently, the combination of temporal and spatial exposure to long-range Hh signaling was accepted as the main molecular mechanism underlying digit patterning (Ahn & Joyner, 2004; Harfe et al., 2004). However, current investigations in chicken limb buds show that mesenchymal cells in the posterior limb bud produce actin-based filopodia that accumulate particles of Hh, allowing the Hh particles to act at a distance without extracellular diffusion (Sanders, Llagostera, & Barna, 2013). This study suggested that the proper activation of the Hh target genes is more critical for digit patterning than the diffusible long-range Hh morphogen threshold. The polydactyly phenotype of other carriers of *IQCE* variations supports the involvement of *IQCE* in the formation of limbs, a process that is related to Hh signaling. At the base of the cilia, the EvC complex (constituted by EVC-EVC2 and EFCA-B7-IQCE) is not essential for Hh target gene induction. However, it can reduce the Hh-stimulated induction of *GLI1* or *PTCH1* (Pusapati et al., 2014). Interestingly, gene ontology analysis of misregulated transcripts revealed enrichment for skeletal system development. Notably, *IGF2*—a critical protein for bone formation (Fisher et al., 2005; Shi et al., 2015)—is misregulated in the absence of *IQCE*. Defects in *IGF2* in humans cause Silver-Russell syndrome, ectrodactyly, under masculinized genitalia, developmental delay, and placental hypoplasia (Begemann et al., 2015; Yamoto et al., 2017). Additionally, we observed the upregulation of *IGFBP1* and *IGFBP3*, these two proteins bind *IGF2* and their excess may limit *IGF2* function (Chao & D'Amore, 2008).

Knockdown of *iqce* in zebrafish showed the full spectrum of phenotypes associated with defective cilia: Body curvature, kidney cysts, left-right asymmetry, misdirected cilia in pronephric duct (Shi, Su, Lipschutz, & Lobo, 2017), as well as retinal defects resembling very much those observed in *bbs9* morphants (Veleri et al., 2012). In addition, full *iqce* knockout (KO) mouse showed short tibias, thin retinas, metabolic and behavioral alterations, as well as polydactyly. Hh expression during limb development is pivotal for defining the identity and number of digits. When its expression is activated at an ectopic site in the limb bud, this defines a pathogenic mechanism that leads to preaxial polydactyly

**FIGURE 3** *iqce* Morphants present ciliopathy phenotypes. (a, b) *iqce* Morphants exhibit a curved body axis. (a) Uninjected and (b) *iqce-mo* morphants, \* indicates hydrocephalus. (c) Cysts in the paired pronephric glomeruli in *iqce* morphants (arrows). (d) Rescue of *iqce-mo* (6  $\mu$ M) injected embryos with different concentrations of *iqce* mRNA; 10 ng/ $\mu$ l ( $n = 121$ ), 20 ng/ $\mu$ l ( $n = 163$ ), and 30 ng/ $\mu$ l ( $n = 193$ ). The rescue was evaluated by the loss of curved body, data include  $\pm$  SD only when needed (WT and the best condition [i.e., 30 ng/ $\mu$ l]). (e–g). Cilia in the pronephric duct of *iqce* morphants are misdirected, shortened, and malformed (f), whereas in (e) uninjected or (g) control morpholinos, they are elongated. Scale bar = 11.6  $\mu$ m. (h–j) Cilia in nasal placode are reduced in number and size in (i) morphants ( $n = 5$ ) compared to (h) uninjected ( $n = 5$ ) and (j) morpholino control ( $n = 5$ ). Scale bar = 11.6  $\mu$ m. Scheme: Square shows the area of observation for pronephric cilia, and circle for nasal placode cilia. (k, l) 5  $\mu$ m sections through eyes of uninjected or *iqce-mo* injected embryos stained with toluidine blue (4 dpf). The uninjected embryos show normal lamination, whereas morphants lack a properly formed retina. \* indicates the outer segments of photoreceptors. Scale bar = 0.2 cm. (m) Chart indicating the percentage of normal and situs inversus or heterotaxia in uninjected ( $n = 70$ ), *iqce-mo* ( $n = 250$ ), or *iqce-mocont* ( $n = 160$ ) injected embryos, data include  $\pm$  SD. (n–p) Patched1 (*ptc1*) in situ hybridization revealed expanded expression in (o) *iqce* morphant ( $n = 81$ ) but not in (n) uninjected ( $n = 88$ ) or (p) control morpholino ( $n = 30$ ) injected embryos. (q–s) In situ immunohistochemistry showing expanded engrailed expression in (r) *iqce* morphants ( $n = 10$ ) compared with (q) uninjected ( $n = 10$ ) and (s) *iqce-mocont* ( $n = 10$ ) injected embryos. Scale bar = 47.2  $\mu$ m. (a–j) 48 hpf; (k–m) 72 hpf; (n–p) 10 somites; and (q–s) 24 hpf. GCL, ganglion cell layer; INL, inner nuclear layer; mRNA, messenger RNA; ONL, outer nuclear layer; PRs, photoreceptors; SD, standard deviation; WT, wild-type

(Lettice et al., 2012). The *iqce* zebrafish model exhibits modified Hh signaling. Indeed, we observed a broadening of the Hh *ptc1* receptor outside of the adaxial cell, as well as extended *eng* expression. It is well known that the primary cilium acts as a Hh signal transduction machine (Goetz, Ocbina, & Anderson, 2009). Recently, it was shown that in a *bbs1* KO cell line, the entry of Smo into cilia upon Hh stimulation was increased compared with the wild-type cell line (Nozaki et al., 2018). This, in turn, activated ectopic *eng* expression as observed in *iqce* zebrafish morphants. This result was unexpected since Pusapati et al. (2014) provided evidence by analyzing osteoblasts and the NIH3T3 cell line that the EFCAB7-IQCE module anchors the EVC-EVC2 complex in a signaling microdomain at the base of the cilia. Moreover, they showed that EFCAB7 depletion triggered the mislocalization of EVC-EVC2 within cilia and impaired the activation of the transcription factor GLI2. These data suggest an impairment of Hh signaling rather than the gain of Hh signalling activity in the *iqce* mutant background that we observed in the zebrafish. However, in zebrafish, no *Evc* or *Evc2* orthologues have been identified but only a paralog absent in human (Pusapati et al., 2014). Transcriptome analysis of the patient's cells under normal condition did not provide evidence of GLI genes modified expression compared to controls. However, under Hh activation, the patient's cells show reduced expression of most of the key Hh pathway genes (*GLI2*, *GLI3*, *EFCAB7*, *EVC*, *EVC2*; Figure S2). At the protein level, we observed a significant amount of EVC2 mislocalization at the base of the cilia (Figure 2b) as well as a reduced GLI2 protein expression (data not shown).

Blindness is not present in patients with only *iqce* defect. However, it has been previously observed in other cases that KO of genes in zebrafish can give additional phenotypes not observed in patients. For example, the zebrafish *cc2d2* adult mutants exhibit a scoliosis that is not observed in Joubert syndrome (JBTS; MIM# 213300; Bachmann-Gagescu et al., 2011). The defective retina of *iqce* morphants is very similar to the *bbs9* morphant model showing a lack of lamination (Veleri et al., 2012). This observation fits with the implication of *iqce* in ciliopathies.

In conclusion, our findings confirm the role of IQCE in nonsyndromic PAP, and its negative impact on the Hh signaling pathway. The differences between *iqce* zebrafish and mice models compared with humans raise the question of additional genetic factors. This study illustrates how detailed and accurate clinical data are needed to interpret exome-scale genetic results and translate those results in correct counseling and therapeutic intervention.

## ACKNOWLEDGMENTS

We would like to thank the patient's family for their participation and the patient's associations FormiCoeur, Bardet-Biedl France, and Unadev, for their constant and strong support. They would like to thank the members of the diagnostic laboratories (molecular and cytogenetic units, in particular, Manuela Antin, Anne-Sophie Leuvrey, and Elsa Nourisson) at the University Hospital in Strasbourg, France, for technical assistance. They would like to thank Nadine Borel and the zebrafish house team. They would also like to thank the collaborators for screening their negative cohort (Kinga Bujakowska, Eric Pierce, Phil Beales, and

Friedhelm Hildebrandt). Finally, they would like to thank the GenomEast platform (Strasbourg), especially Serge Vicaire for library preparation, Céline Keime for supervising the transcriptome analysis, as well as Bernard Jost and Christelle Thibault. The computing resources for this study were provided by the BICS and BISTRO bioinformatics platforms in Strasbourg. Whole-exome sequencing was funded by Retina France (100 Exomes Program), by Hôpitaux Universitaires de Strasbourg (API 20132013-2014-2014 HUS N° 58855885). RNA-Seq analysis performed in patients with IQCE pathogenic variant and other patients were funded by Fondation Maladie Rare and Fondation Jérôme Lejeune. The work performed by C. E. and U. S. was supported by the HGF Biointerfaces Programme; the work performed by A. E.-C. work was supported by the Fondation JED-Belgique; and the work performed by C. D. was supported by a FRM Grant (ECO20170637509).

## CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

## DATA AVAILABILITY STATEMENT

Data generated or analyzed during this study are included in the published article and the corresponding supporting information. The raw sequencing data generated in the course of this study are not publicly available due to the protocol and the corresponding consents used that did not include such information. All variants have been submitted to ClinVar using the following range of accessions numbers SCV000929968–SCV000929972 (<https://www.ncbi.nlm.nih.gov/clinvar/>). Anonymized NGS data and genomic variant data files will be made available upon request from qualified investigators studying the molecular basis of genomic disorders. Datasets can be obtained via the corresponding author on reasonable request.

## ORCID

Alejandro Estrada-Cuzcano  <http://orcid.org/0000-0001-9868-2465>

Amélie Piton  <http://orcid.org/0000-0003-0408-7468>

Jean Muller  <http://orcid.org/0000-0002-7682-559X>

## REFERENCES

- Abbasi, A. A. (2011). Evolution of vertebrate appendicular structures: Insight from genetic and palaeontological data. *Developmental Dynamics*, 240(5), 1005–1016. <https://doi.org/10.1002/dvdy.22572>
- Abdelhamed, Z. A., Wheway, G., Szymanska, K., Natarajan, S., Toomes, C., Inglehearn, C., & Johnson, C. A. (2013). Variable expressivity of ciliopathy neurological phenotypes that encompass Meckel-Gruber syndrome and Joubert syndrome is caused by complex de-regulated ciliogenesis, Shh and Wnt signalling defects. *Human Molecular Genetics*, 22(7), 1358–1372. <https://doi.org/10.1093/hmg/dds546>
- Ahn, S., & Joyner, A. L. (2004). Dynamic changes in the response of cells to positive hedgehog signaling during mouse limb patterning. *Cell*, 118(4), 505–516. <https://doi.org/10.1016/j.cell.2004.07.023>

- Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., Korbel, J. O., ... Abecasis, G. R., 1000 Genomes Project Consortium (2015). A global reference for human genetic variation. *Nature*, 526(7571), 68–74. <https://doi.org/10.1038/nature15393>
- Armant, O., März, M., Schmidt, R., Ferg, M., Diotel, N., Ertzer, R., ... Rastegar, S. (2013). Genome-wide, whole mount in situ analysis of transcriptional regulators in zebrafish embryos. *Developmental Biology*, 380(2), 351–362. <https://doi.org/10.1016/j.ydbio.2013.05.006>
- Bachmann-Gagescu, R., Phelps, I. G., Stearns, G., Link, B. A., Brockerhoff, S. E., Moens, C. B., & Doherty, D. (2011). The ciliopathy gene *cc2d2a* controls zebrafish photoreceptor outer segment development through a role in Rab8-dependent vesicle trafficking. *Human Molecular Genetics*, 20(20), 4041–4055. <https://doi.org/10.1093/hmg/ddr332>
- Backenroth, D., Homsy, J., Murillo, L. R., Glessner, J., Lin, E., Brueckner, M., ... Shen, Y. (2014). CANOES: Detecting rare copy number variants from whole exome sequencing data. *Nucleic Acids Research*, 42(12), e97. <https://doi.org/10.1093/nar/gku345>
- Barakat, M. T., Humke, E. W., & Scott, M. P. (2010). Learning from Jekyll to control Hyde: Hedgehog signaling in development and cancer. *Trends in Molecular Medicine*, 16(8), 337–348. <https://doi.org/10.1016/j.molmed.2010.05.003>
- Begemann, M., Zirn, B., Santen, G., Wirthgen, E., Soellner, L., Büttel, H. M., ... Eggemann, T. (2015). Paternally inherited IGF2 mutation and growth restriction. *New England Journal of Medicine*, 373(4), 349–356. <https://doi.org/10.1056/NEJMoa1415227>
- Breslow, D. K., Hoogendoorn, S., Kopp, A. R., Morgens, D. W., Vu, B. K., Kennedy, M. C., ... Nachury, M. V. (2018). A CRISPR-based screen for hedgehog signaling provides insights into ciliary function and ciliopathies. *Nature Genetics*, 50, 460–471. <https://doi.org/10.1038/s41588-018-0054-7>
- Carmi, R., Elbedour, K., Stone, E. M., & Sheffield, V. C. (1995). Phenotypic differences among patients with Bardet-Biedl syndrome linked to three different chromosome loci. *American Journal of Medical Genetics*, 59(2), 199–203. <https://doi.org/10.1002/ajmg.1320590216>
- Castilla, E., Paz, J., Mutchinick, O., Muñoz, E., Giorgiutti, E., & Gelman, Z. (1973). Polydactyly: A genetic study in South America. *American Journal of Human Genetics*, 25(4), 405–412. <http://www.ncbi.nlm.nih.gov/pubmed/4716659>
- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, 4, 7. <https://doi.org/10.1186/s13742-015-0047-8>
- Chao, W., & D'Amore, P. A. (2008). IGF2: Epigenetic regulation and role in development and disease. *Cytokine & Growth Factor Reviews*, 19(2), 111–120. <https://doi.org/10.1016/j.cytogr.2008.01.005>
- Concordet, J. P., Lewis, K. E., Moore, J. W., Goodrich, L. V., Johnson, R. L., Scott, M. P., & Ingham, P. W. (1996). Spatial regulation of a zebrafish patched homologue reflects the roles of sonic hedgehog and protein kinase A in neural tube and somite patterning. *Development*, 122(9), 2835–2846. <http://www.ncbi.nlm.nih.gov/pubmed/8787757>
- Costa, M. L., Escaleira, R. C., Rodrigues, V. B., Manasfi, M., & Mermelstein, C. S. (2002). Some distinctive features of zebrafish myogenesis based on unexpected distributions of the muscle cytoskeletal proteins actin, myosin, desmin,  $\alpha$ -actinin, troponin and titin. *Mechanisms of Development*, 116(1–2), 95–104. [https://doi.org/10.1016/S0925-4773\(02\)00149-1](https://doi.org/10.1016/S0925-4773(02)00149-1)
- DeLuca, A. P., Weed, M. C., Haas, C. M., Halder, J. A., & Stone, E. M. (2015). Apparent Usher syndrome caused by the combination of BBS1-associated retinitis pigmentosa and SLC26A4-associated deafness. *JAMA Ophthalmology*, 133(8), 967–968. <https://doi.org/10.1001/jamaophthol.2015.1463>
- DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., ... Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, 43(5), 491–498. <https://doi.org/10.1038/ng806>
- Estrada-Cuzcano, A., Koenekoop, R. K., Senechal, A., De Baere, E. B. W., de Ravel, T., Banfi, S., ... Klevering, B. J. (2012). BBS1 mutations in a wide spectrum of phenotypes ranging from nonsyndromic retinitis pigmentosa to Bardet-Biedl syndrome. *Archives of Ophthalmology*, 130(11), 1425–1432. <https://doi.org/10.1001/archophthol.2012.2434>
- Etard, C., Roostalu, U., & Strähle, U. (2010). Lack of Apobec2-related proteins causes a dystrophic muscle phenotype in zebrafish embryos. *The Journal of Cell Biology*, 189(3), 527–539. <https://doi.org/10.1083/jcb.200912125>
- Fisher, M. C., Meyer, C., Garber, G., & Dealy, C. N. (2005). Role of IGFBP2, IGF-I and IGF-II in regulating long bone growth. *Bone*, 37(6), 741–750. <https://doi.org/10.1016/j.bone.2005.07.024>
- Flannick, J., Thorleifsson, G., Beer, N. L., Jacobs, S. B. R., Grarup, N., Burtt, N. P., ... Altshuler, D. (2014). Loss-of-function mutations in SLC30A8 protect against type 2 diabetes. *Nature Genetics*, 46(4), 357–363. <https://doi.org/10.1038/ng.2915>
- Geoffroy, V., Herenger, Y., Kress, A., Stoetzel, C., Piton, A., Dollfus, H., & Muller, J. (2018). AnnotSV: An integrated tool for structural variations annotation. *Bioinformatics*, 34, 3572–3574. <https://doi.org/10.1093/bioinformatics/bty304>
- Geoffroy, V., Pizot, C., Redin, C., Piton, A., Vasli, N., Stoetzel, C., ... Muller, J. (2015). VaRank: A simple and powerful tool for ranking genetic variants. *PeerJ*, 3, e796. <https://doi.org/10.7717/peerj.796>
- Geoffroy, V., Stoetzel, C., Scheidecker, S., Schaefer, E., Perrault, I., Bär, S., ... Muller, J. (2018). Whole-genome sequencing in patients with ciliopathies uncovers a novel recurrent tandem duplication in IFT140. *Human Mutation*, 39(7), 983–992. <https://doi.org/10.1002/humu.23539>
- Goetz, S. C., & Anderson, K. V. (2010). The primary cilium: A signalling centre during vertebrate development. *Nature Reviews Genetics*, 11(5), 331–344. <https://doi.org/10.1038/nrg2774>
- Goetz, S. C., Ocbina, P. J. R., & Anderson, K. V. (2009). The primary cilium as a hedgehog signal transduction machine. *Methods in Cell Biology*, 94, 199–222. [https://doi.org/10.1016/S0091-679x\(08\)94010-3](https://doi.org/10.1016/S0091-679x(08)94010-3)
- González-Del Pozo, M., Méndez-Vidal, C., Santoyo-Lopez, J., Vela-Boza, A., Bravo-Gil, N., Rueda, A., ... Antiñolo, G. (2014). Deciphering intrafamilial phenotypic variability by exome sequencing in a Bardet-Biedl family. *Molecular Genetics & Genomic Medicine*, 2(2), 124–133. <https://doi.org/10.1002/mgg3.50>
- Hagstrom, S. A., North, M. A., Nishina, P. M., Berson, E. L., & Dryja, T. P. (1998). Recessive mutations in the gene encoding the tubby-like protein TULP1 in patients with retinitis pigmentosa. *Nature Genetics*, 18(2), 174–176. <https://doi.org/10.1038/ng0298-174>
- Hanein, S., Perrault, I., Gerber, S., Tanguy, G., Barbet, F., Ducroq, D., ... Kaplan, J. (2004). Leber congenital amaurosis: Comprehensive survey of the genetic heterogeneity, refinement of the clinical definition, and genotype-phenotype correlations as a strategy for molecular diagnosis. *Human Mutation*, 23(4), 306–317. <https://doi.org/10.1002/humu.20010>
- Harfe, B. D., Scherz, P. J., Nissim, S., Tian, H., McMahon, A. P., & Tabin, C. J. (2004). Evidence for an expansion-based temporal Shh gradient in specifying vertebrate digit identities. *Cell*, 118(4), 517–528. <https://doi.org/10.1016/j.cell.2004.07.024>
- Huang, P., & Schier, A. F. (2009). Dampened hedgehog signaling but normal Wnt signaling in zebrafish without cilia. *Development*, 136(18), 3089–3098. <https://doi.org/10.1242/dev.041343>
- Ingham, P. W., Nakano, Y., & Seger, C. (2011). Mechanisms and functions of hedgehog signalling across the metazoa. *Nature Reviews Genetics*, 12(6), 393–406. <https://doi.org/10.1038/nrg2984>
- Kancheva, D., Atkinson, D., De Rijk, P., Zimon, M., Chamova, T., Mitev, V., ... Jordanova, A. (2016). Novel mutations in genes causing hereditary spastic paraplegia and Charcot-Marie-Tooth neuropathy identified by an optimized protocol for homozygosity mapping based on whole-exome sequencing. *Genetics in Medicine*, 18(6), 600–607. <https://doi.org/10.1038/gim.2015.139>
- Karaca, E., Posey, J. E., Coban Akdemir, Z., Pehlivan, D., Harel, T., Jhangiani, S. N., ... Lupski, J. R. (2018). Phenotypic expansion illuminates multilocus



- pathogenic variation. *Genetics in Medicine*, 20, 1528–1537. <https://doi.org/10.1038/gim.2018.33>
- Katsanis, N., Ansley, S. J., Badano, J. L., Eichers, E. R., Lewis, R. A., Hoskins, B. E., & Lupski, J. R. (2001). Triallelic inheritance in Bardet-Biedl syndrome, a Mendelian recessive disorder. *Science*, 293(5538), 2256–2259. <https://doi.org/10.1126/science.1063525>
- Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., ... MacArthur, D. G. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, 536(7616), 285–291. <https://doi.org/10.1038/nature19057>
- Lettice, L. A., Williamson, I., Wiltshire, J. H., Peluso, S., Devenney, P. S., Hill, A. E., ... Hill, R. E. (2012). Opposing functions of the ETS factor family define Shh spatial expression in limb buds and underlie polydactyly. *Developmental Cell*, 22(2), 459–467. <https://doi.org/10.1016/j.devcel.2011.12.010>
- MacDonald, J. R., Ziman, R., Yuen, R. K. C., Feuk, L., & Scherer, S. W. (2014). The database of genomic variants: A curated collection of structural variation in the human genome. *Nucleic Acids Research*, 42(D1), D986–D992. <https://doi.org/10.1093/nar/gkt958>
- Malik, S. (2014). Polydactyly: Phenotypes, genetics and classification. *Clinical Genetics*, 85(3), 203–212. <https://doi.org/10.1111/cge.12276>
- Malik, S., Ullah, S., Afzal, M., Lal, K., & Haque, S. (2014). Clinical and descriptive genetic study of polydactyly: A Pakistani experience of 313 cases. *Clinical Genetics*, 85(5), 482–486. <https://doi.org/10.1111/cge.12217>
- Metsalu, T., & Vilo, J. (2015). ClustVis: A web tool for visualizing clustering of multivariate data using principal component analysis and heatmap. *Nucleic Acids Research*, 43(W1), W566–W570. <https://doi.org/10.1093/nar/gkv468>
- Muller, F., Blader, P., Rastegar, S., Fischer, N., Knöchel, W., & Strähle, U. (1999). Characterization of zebrafish smad1, smad2 and smad5: The amino-terminus of Smad1 and Smad5 is required for specific function in the embryo. *Mechanisms of Development*, 88(1), 73–88. [https://doi.org/10.1016/S0925-4773\(99\)00173-2](https://doi.org/10.1016/S0925-4773(99)00173-2)
- Muller, J., Stoetzel, C., Vincent, M. C., Leitch, C. C., Laurier, V., Danse, J. M., ... Dollfus, H. (2010). Identification of 28 novel mutations in the Bardet-Biedl syndrome genes: The burden of private mutations in an extensively heterogeneous disease. *Human Genetics*, 127(5), 583–593. <https://doi.org/10.1007/s00439-010-0804-9>
- North, M. A., Naggert, J. K., Yan, Y., Noben-Trauth, K., & Nishina, P. M. (1997). Molecular characterization of TUB, TULP1, and TULP2, members of the novel tubby gene family and their possible relation to ocular diseases. *Proceedings of the National Academy of Sciences of the United States of America*, 94(7), 3128–3133. <http://www.ncbi.nlm.nih.gov/pubmed/9096357>
- Nozaki, S., Katoh, Y., Kobayashi, T., & Nakayama, K. (2018). BBS1 is involved in retrograde trafficking of ciliary GPCRs in the context of the BBSome complex. *PLoS One*, 13(3), e0195005. <https://doi.org/10.1371/journal.pone.0195005>
- Oxtoby, E., & Jowett, T. (1993). Cloning of the zebrafish krox-20 gene (Krx-20) and its expression during hindbrain development. *Nucleic Acids Research*, 21(5), 1087–1095. <https://doi.org/10.1093/nar/21.5.1087>
- Petrova, R., & Joyner, A. L. (2014). Roles for hedgehog signaling in adult organ homeostasis and repair. *Development*, 141(18), 3445–3457. <https://doi.org/10.1242/dev.083691>
- Phadke, S. R., & Sankar, V. H. (2010). Polydactyly and genes. *Indian Journal of Pediatrics*, 77(3), 277–281. <https://doi.org/10.1007/s12098-010-0033-1>
- Pusapati, G. V., Hughes, C. E., Dorn, K. V., Zhang, D., Sugianto, P., Aravind, L., & Rohatgi, R. (2014). EFCAB7 and IQCE regulate hedgehog signaling by tethering the EVC-EVC2 complex to the base of primary cilia. *Developmental Cell*, 28(5), 483–496. <https://doi.org/10.1016/j.devcel.2014.01.021>
- Pusapati, G. V., Kong, J. H., Patel, B. B., Krishnan, A., Sagner, A., Kinnebrew, M., ... Rohatgi, R. (2018). CRISPR screens uncover genes that regulate target cell sensitivity to the morphogen sonic hedgehog. *Developmental Cell*, 44(2), 271. <https://doi.org/10.1016/j.devcel.2018.01.002>
- Redin, C., Le Gras, S., Mhamdi, O., Geoffroy, V., Stoetzel, C., Vincent, M. C., ... Muller, J. (2012). Targeted high-throughput sequencing for diagnosis of genetically heterogeneous diseases: Efficient mutation detection in Bardet-Biedl and Alström Syndromes. *Journal of Medical Genetics*, 49(8), 502–512. <https://doi.org/10.1136/jmedgenet-2012-100875>
- Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., ... Rehm, H. L. (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine*, 17(5), 405–423. <https://doi.org/10.1038/gim.2015.30>
- Robu, M. E., Larson, J. D., Nasevicius, A., Beiraghi, S., Brenner, C., Farber, S. A., & Ekker, S. C. (2007). p53 activation by knockdown technologies. *PLoS Genetics*, 3(5), e78. <https://doi.org/10.1371/journal.pgen.0030078>
- Roostalu, U., & Strähle, U. (2012). In vivo imaging of molecular interactions at damaged sarcolemma. *Developmental Cell*, 22(3), 515–529. <https://doi.org/10.1016/j.devcel.2011.12.008>
- Sanders, T. A., Llagostera, E., & Barna, M. (2013). Specialized filopodia direct long-range transport of SHH during vertebrate tissue patterning. *Nature*, 497(7451), 628–632. <https://doi.org/10.1038/nature12157>
- Shaheen, R., Szymanska, K., Basu, B., Patel, N., Ewida, N., Faqeh, E., ... Alkuraya, F. S. (2016). Characterizing the morbid genome of ciliopathies. *Genome Biology*, 17(1), 242. <https://doi.org/10.1186/s13059-016-1099-5>
- Shi, Y., Chen, J., Karner, C. M., & Long, F. (2015). Hedgehog signaling activates a positive feedback mechanism involving insulin-like growth factors to induce osteoblast differentiation. *Proceedings of the National Academy of Sciences of the United States of America*, 112(15), 4678–4683. <https://doi.org/10.1073/pnas.1502301112>
- Shi, Y., Su, Y., Lipschutz, J. H., & Lobo, G. P. (2017). Zebrafish as models to study ciliopathies of the eye and kidney. *Clinical Nephrology and Research*, 1(1), 6–9. <http://www.ncbi.nlm.nih.gov/pubmed/29553143>
- Singh, M. M., Kumar, R., Tewari, S., & Agarwal, S. (2018). Association of GSTT1/GSTM1 and ApoE variants with left ventricular diastolic dysfunction in thalassaemia major patients. *Hematology*, 24, 20–25. <https://doi.org/10.1080/10245332.2018.1502397>
- Stamenkovic, M., Lukic, V., Suvakov, S., Simic, T., Sencanic, I., Pljesa-Ercegovac, M., & Djukic, T. (2018). GSTM1-null and GSTT1-active genotypes as risk determinants of primary open angle glaucoma among smokers. *International Journal of Ophthalmology*, 11(9), 1514–1520. <https://doi.org/10.18240/ijo.2018.09.14>
- Stooke-Vaughan, G. A., Huang, P., Hammond, K. L., Schier, A. F., & Whitfield, T. T. (2012). The role of hair cells, cilia and ciliary motility in otolith formation in the zebrafish otic vesicle. *Development*, 139(10), 1777–1787. <https://doi.org/10.1242/dev.079947>
- Tennessen, J. A., Bigham, A. W., O'Connor, T. D., Fu, W., Kenny, E. E., Gravel, S., ... Akey, J. M. (2012). Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science*, 337(6090), 64–69. <https://doi.org/10.1126/science.1219240>
- Ullah, I., Kakar, N., Schrauwen, I., Hussain, S., Chakchouk, I., Liaqat, K., ... Leal, S. M. (2019). Variants in KIAA0825 underlie autosomal recessive postaxial polydactyly. *Human Genetics*, 138, 593–600. <https://doi.org/10.1007/s00439-019-02000-0>
- Umair, M., Ahmad, F., Bilal, M., Ahmad, W., & Alfadhel, M. (2018). Clinical genetics of polydactyly: An updated review. *Frontiers in Genetics*, 9, 447. <https://doi.org/10.3389/fgene.2018.00447>
- Umair, M., Shah, K., Alhaddad, B., Haack, T. B., Graf, E., Strom, T. M., ... Ahmad, W. (2017). Exome sequencing revealed a splice site variant in the IQCE gene underlying post-axial polydactyly type A restricted to lower limb. *European Journal of Human Genetics*, 25(8), 960–965. <https://doi.org/10.1038/ejhg.2017.83>
- Vargas-Poussou, R., Houillier, P., Le Pottier, N., Strompf, L., Loirat, C., Baudouin, V., ... Blanchard, A. (2006). Genetic investigation of autosomal recessive distal renal tubular acidosis: Evidence for early sensorineural hearing loss associated with mutations in the ATP6-VOA4 gene. *Journal of the American Society of Nephrology*, 17(5), 1437–1443. <https://doi.org/10.1681/ASN.2005121305>



- Veleri, S., Bishop, K., Dalle Nogare, D. E., English, M. A., Fokkett, T. J., Chitnis, A., ... Swaroop, A. (2012). Knockdown of Bardet-Biedl syndrome gene BBS9/PTHB1 leads to cilia defects. *PLoS One*, 7(3), e34389.
- Wang, J., Vasaiakar, S., Shi, Z., Greer, M., & Zhang, B. (2017). WebGestalt 2017: A more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Research*, 45(W1), W130–W137. <https://doi.org/10.1093/nar/gkx356>
- Ware, S. M., Aygun, M. G., & Hildebrandt, F. (2011). Spectrum of clinical diseases caused by disorders of primary cilia. *Proceedings of the American Thoracic Society*, 8(5), 444–450. <https://doi.org/10.1513/pats.201103-025SD>
- Waters, A. M., & Beales, P. L. (2011). Ciliopathies: An expanding disease spectrum. *Pediatric Nephrology*, 26(7), 1039–1056. <https://doi.org/10.1007/s00467-010-1731-7>
- Westerfield, M. (1993). *The zebrafish book: A guide for the laboratory use of zebrafish (Brachydanio rerio)* (p. 300). Eugene, OR: University of Oregon Press.
- Wolff, C., Roy, S., Lewis, K. E., Schauerte, H., Joerg-Rauch, G., Kirn, A., ... Ingham, P. W. (2004). Iguana encodes a novel zinc-finger protein with coiled-coil domains essential for hedgehog signal transduction in the zebrafish embryo. *Genes & Development*, 18(13), 1565–1576. <https://doi.org/10.1101/gad.296004>
- Yamamoto, M., Matsuzaki, T., Takahashi, R., Adachi, E., Maeda, Y., Yamaguchi, S., ... Noda, M. (2012). The transformation suppressor gene Reck is required for postaxial patterning in mouse forelimbs. *Biology Open*, 1(5), 458–466. <https://doi.org/10.1242/bio.2012638>
- Yamoto, K., Saitsu, H., Nakagawa, N., Nakajima, H., Hasegawa, T., Fujisawa, Y., ... Ogata, T. (2017). De novo IGF2 mutation on the paternal allele in a patient with Silver-Russell syndrome and ectrodactyly. *Human Mutation*, 38(8), 953–958. <https://doi.org/10.1002/humu.23253>
- Zhang, Z., Hao, C. J., Li, C. G., Zang, D. J., Zhao, J., Li, X. N., ... Li, W. (2014). Mutation of SLC35D3 causes metabolic syndrome by impairing dopamine signaling in striatal D1 neurons. *PLoS Genetics*, 10(2), e1004124. <https://doi.org/10.1371/journal.pgen.1004124>
- Zhang, W., Taylor, S. P., Ennis, H. A., Forlenza, K. N., Duran, I., Li, B., ... Cohn, D. H. (2018). Expanding the genetic architecture and phenotypic spectrum in the skeletal ciliopathies. *Human Mutation*, 39(1), 152–166. <https://doi.org/10.1002/humu.23362>
- Zhao, C., & Malicki, J. (2007). Genetic defects of pronephric cilia in zebrafish. *Mechanisms of Development*, 124(7-8), 605–616. <https://doi.org/10.1016/j.mod.2007.04.004>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Estrada-Cuzzano A, Etard C, Delvallée C, et al. Novel *IQCE* variations confirm its role in postaxial polydactyly and cause ciliary defect phenotype in zebrafish. *Human Mutation*. 2020;41:240–254. <https://doi.org/10.1002/humu.23924>

## **SUPPLEMENTARY INFORMATION**

### **Sanger Sequencing**

Sanger sequencing was performed by way of PCR amplification with 50ng of genomic DNA template. The primers were designed with Primer 3 (<http://frodo.wi.mit.edu/primer3>). Bidirectional sequencing of the purified PCR products was performed by Eurofins Genomics Sequencing Facilities (<https://www.eurofinsgenomics.eu/en/custom-dna-sequencing>).

### **Control screening**

*IQCE* exon 12 was tested in 93 control DNA of matched origin (Algerian) and 93 multi ethnic. High resolution melting (HRM) assay was performed in a 20 µl reaction volume containing 50 ng of template DNA, 15 µl Precision Melt Supermix (Bio-Rad) and 10 µmol of each primer of the primer set. PCR was performed using a CFX96 (Bio-Rad) thermal cycler and consisted of predenaturation (98°C for 2 min) followed by 35 cycles of amplification involving denaturation (98°C for 5 s), annealing (60°C for 10 s) and a melting curve involving denaturation (95°C for 30 s), annealing (70°C for 30 s) and a melt curve (72–95°C with an increment of 0.2°C per 10 s). Data generated were analyzed using Precision Melt Analysis software (Bio-Rad). In HRM analysis, differences in  $T_m$  and normalized curve shape were used together to discriminate even the most difficult-to-detect sequence variation.

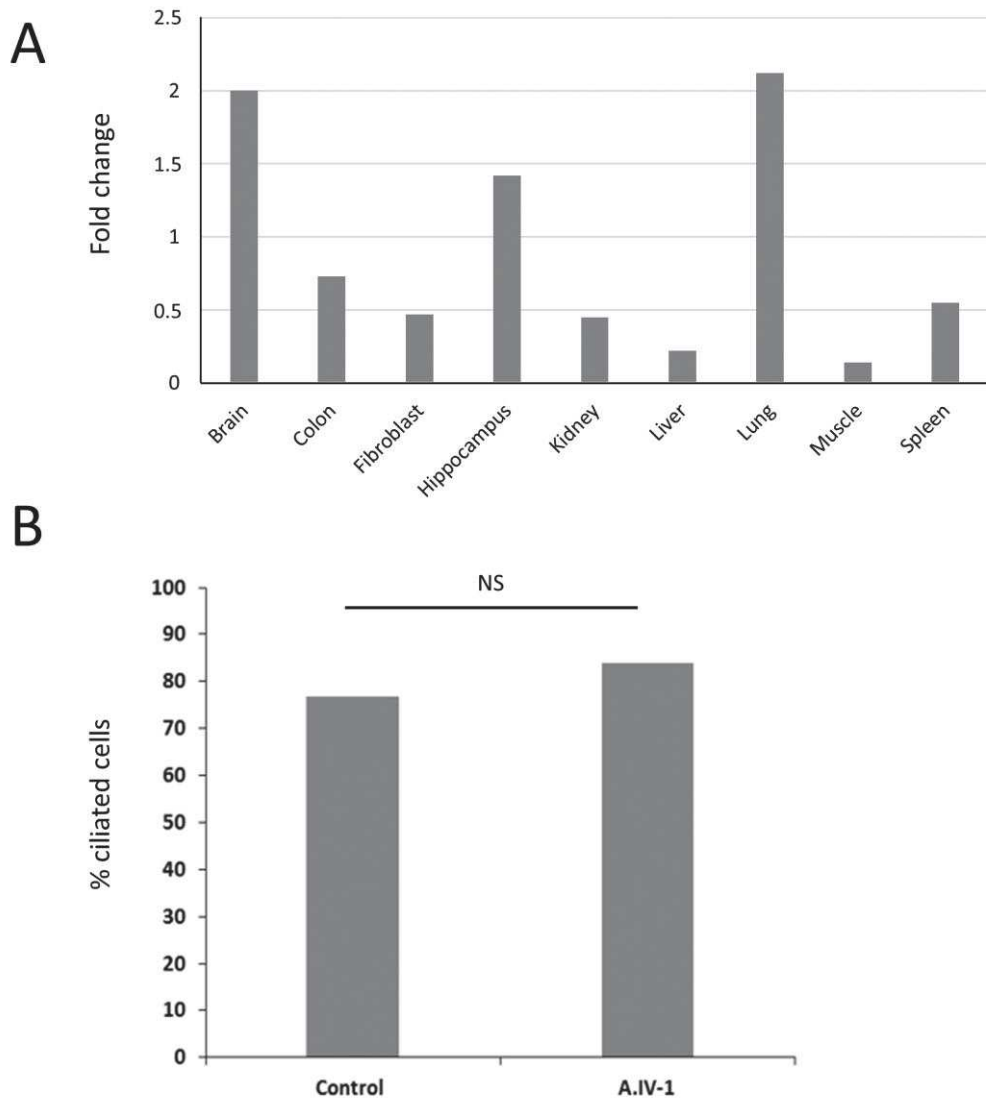
### **Bioinformatic pipeline RNA seq**

Image analysis and base calling were performed using RTA v.2.7.3 and bcl2fastq v.2.17.1.14 (Illumina). Sequence reads were mapped onto hg19/GRCh37 assembly of human genome using Tophat 2.0.14 (Kim et al., 2013) and bowtie version 2-2.1.0 (Langmead, Trapnell, Pop, & Salzberg, 2009), A summary of the mapping results is available in Supplementary Table 6. Gene expression was quantified from uniquely aligned reads using HTSeq-0.6.1 with gene annotations from Ensembl release 75 and intersection non-empty mode.

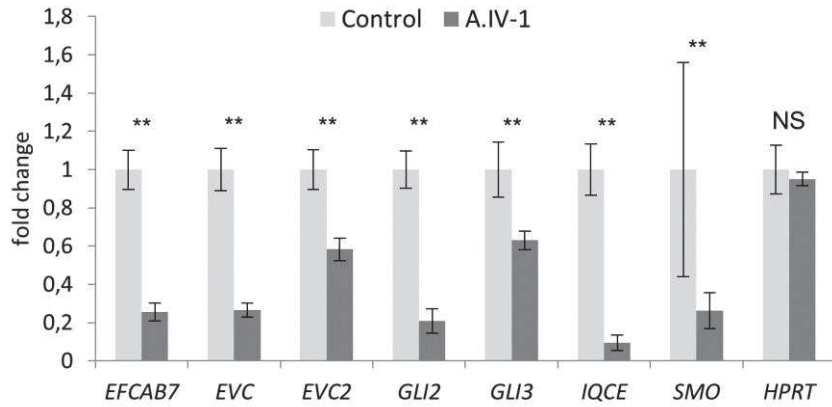
### **Differential expression RNA-seq**

In order to identify variation in mRNA expression due to the absence of *IQCE* in the patient cells, comparisons of read counts were performed using R 3.2.5 with the statistical method proposed by Anders and Huber implemented in the DESeq2 v.1.10.1 Bioconductor package

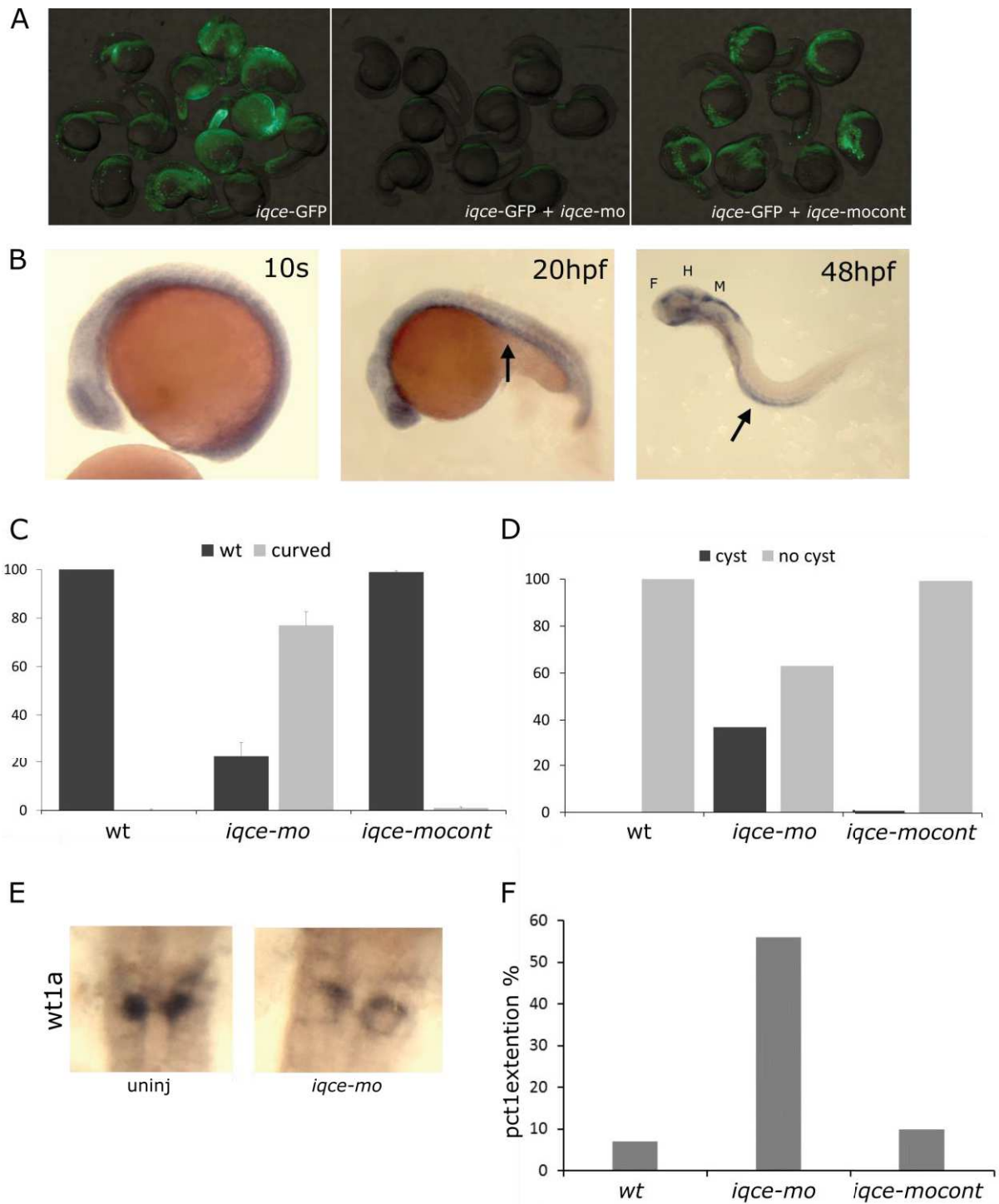
(Love, et al., 2014). The generated data were explored and visualized to assess and check data quality and to eventually remove bad quality data. Especially, PCA was built after stabilizing variance using the regularized log transformation method or RSeqQC was used to do some quality controls such as gene body coverage. Only genes with  $|\log_2 \text{fold-change}| > 0,5$  were considered, as we expected to identify a large amount of change between transcriptomic results from one patient compared to other individuals (indirect consequences of the pathogenic mutation or unrelated events due to variability between humans). To avoid as much as possible false positive variants, we filtered out all genes with a Benjamini and Hochberg adjusted p-value above 0.05.



**Supplementary Figure S1. A.** Highest mRNA expression of *IQCE* in adult human brain/hippocampus and lung. **B.** Count of ciliated fibroblasts between A.IV-1 and control (~100 cells per condition). Percentages are shown in histogram, N.S.: non-significant, Fisher exact test 0.224,  $P < 0.05$ ).



**Supplementary Figure S2. Expression levels of EVC-zone and GLI genes upon ciliogenesis and Hh activation.** Patient's (A.IV-1) and control cells upon treatment (–FCS + SAG 100nM) were compared (n=4 controls, n=6 patient). Statistical significance has been performed using the non-parametric Mann-Whitney test,  $p=0.0095$  (\*\*), NS : non-significant.



**Supplementary Figure S3. Efficacy of *iqce* morpholinos.** **A.** Morpholinos directed against the start codon of *iqce* blocks the expression of a construct containing *iqce* cDNA in frame with *gfp* (central panel). In contrast, 5 mismatch morpholinos did not reduce expression of the chimeric protein (left panel) compared to *iqce-gfp* alone (right panel). 24 hpf. **B.** *in situ* hybridization showing *iqce* expression pattern in 10 somites, 20 and 48 hpf embryos. Arrow indicates the pronephric duct. F: forebrain ventricule; H: hindbrain ventricule; M: midbrain ventricule. **C.** Chart showing the percentage of curved embryos in uninjected (n=200), *iqce-*

*mo* (n=120); and *iqce-mocont* (n=221) injected embryos. 48 hpf, data includes  $\pm$  SD and statistical significance was determined using the Student's t-test, all comparisons were significant with a p-value<0.05. **D.** Chart showing the percentage of cysts in uninjected (n=101), *iqce-mo* (n=70) and *iqce-mocont* (n=90) injected embryos. Statistical significance was determined using the Student's t-test, all comparisons were significant with a p-value<0.05. **E.** *wt1a in situ* hybridization showing expression within glomeruli in non-injected embryos and at the periphery of the cyst in morphants (48 hpf). **F.** Chart showing the percentage of embryos with *ptc1* extension in 10 somites embryos (n=191).



**Supplementary Table S1.** Summary of the whole exome sequencing results. SNV: Single nucleotide variation, indel: gain or loss of up to 50 nucleotides at a single locus, SV: Structural Variation. Exclusion of SV with a DGV frequency > 1% is done only with studies of more than 1000 individuals

	A.IV-1			B.II-2			B.II-3			C.IV-4		
	SNV	indel	SV	SNV	indel	SV	SNV	Indel	SV	SNV	Indel	SV
Total number of variants	83129	11664	31	81972	11319	17	82116	11195	16	78906	11727	31
After exclusion of variants with an allele frequency >1% (gnomAD, 1000G, internal exome database, DGV)	9771	1125	16	7069	993	6	6856	943	9	5323	982	17
After exclusion of SNV/indel in 5'UTR, 3'UTR, downstream, upstream, intron and synonymous locations without local splice effect prediction	788	76	/	553	59	/	560	65	/	753	113	/
After exclusion of missense with no deleterious effect according to SIFT, PPH2 or PhastCons	669	76	/	458	59	/	465	65	/	641	113	/
After exclusion of variants not in a ciliary genes list (Nevers et al., 2017) and not in non-syndromic polydactyly known genes	52	6	1	31	2	2	40	4	2	78	15	1
After selection of variants consistent with recessive transmission (compound heterozygous, homozygous variants)	7 hom.: <i>DNAAF1</i> , <i>IQCE</i> , <i>MADD</i> , <i>SHANK2</i> , <i>TULP1</i> , <i>ZDHHC2</i> 3 comp. het.: <i>ARMC3</i> , <i>DNAH11</i> , <i>MGA</i>			0 homozygous variant 1 comp. het. In <i>IQCE</i>						3 hom.: <i>ZMAT1</i> , <i>AGTPBP1</i> , <i>KDM4C</i> 11 comp. het: <i>CTDSPL2</i> , <i>DDX1</i> , <i>DNAH7</i> , <i>DYNLL1</i> , <i>EP400</i> , <i>IFT20</i> , <i>LRRCC1</i> , <i>LRRC46</i> , <i>MRPL10</i> , <i>PRKDC</i> , <i>SMARCA5</i> , <i>TMEM97</i> , <i>TPTE</i>		

**Supplementary Table S2.** Homozygous regions identified in individual A.IV-1.

Chr	SNP1	SNP2	POS1	POS2	MB	#SNP	DENSITY	
1	rs4147830	rs11431894	94544276	100680594	6.14	33	185.949	
1	rs4915053	rs3761934	108113526	114189067	6.08	136	44.673	
1	1_148741720_T_C	rs11428742	148741720	150550020	1.81	22	82.196	
1	rs4636	rs10927011	234744413	243579112	8.83	125	70.678	
2	rs7577088	rs3764906	39233500	42588195	3.35	20	167.735	
2	rs522893	2_188250301_T_A	169690729	188250301	18.56	224	82.855	
2	rs2304704	rs12477076	190430177	191848312	1.42	23	61.658	
3	rs901854	rs113574694	4856234	10089550	5.23	58	90.230	
3	rs9814557	rs2246945	135720540	137843476	2.12	21	101.092	
4	rs13146318	rs1051447	41615690	49063872	7.45	68	109.532	
4	rs4864727	rs12621	54011424	57897570	3.89	61	63.707	
4	rs10434219	rs344122	62598689	77637502	15.04	175	85.936	
6	rs2076506	6_32485853_C_T	22570064	32485853	9.92	646	15.350	
6	rs1059547	rs12189627	32557529	41059458	8.50	309	27.514	<i>TULP1</i>
7	rs4916935	rs798479	197267	2691201	2.49	41	60.828	<i>IQCE</i>
7	rs7794563	rs80024169	5347295	14775824	9.43	78	120.879	
7	rs4718101	rs6460315	64152353	66289040	2.14	38	56.229	
8	rs3989699	rs13271468	7218621	12236202	5.02	132	38.012	
8	8_12273335_G_A	rs66480340	12273335	24211885	11.94	202	59.102	
8	rs10086200	rs59331088	61193588	74005131	12.81	72	177.938	
8	rs11777189	rs12545587	99057150	136560952	37.50	204	183.842	
9	rs4977555	rs11515	19550308	21968199	2.42	30	80.596	
9	rs7863560	rs3217100	119495697	123734288	4.24	30	141.286	
11	11_32452000_GG_-	rs2076623	32452000	33564123	1.11	25	44.485	
11	rs12274095	rs7101792	69063767	71249386	2.19	28	78.058	
11	rs3740912	rs4937391	125891269	128786294	2.90	32	90.470	
12	rs11054683	rs1600	12247616	14656768	2.41	31	77.715	
12	rs3088008	rs3210837	49724955	51685831	1.96	55	35.652	
12	rs74651927	rs2242497	52695754	56992630	4.30	185	23.226	
12	rs1663564	rs7488309	105546172	118520170	12.97	152	85.355	
12	rs1794956	rs11059841	124848137	129190021	4.34	37	117.348	
14	rs34800262	rs2249922	51370852	52906081	1.54	34	45.154	
14	rs2297113	rs863091	58036494	59112475	1.08	23	46.782	
14	rs35533709	rs2057482	60433392	62213848	1.78	30	59.349	
15	rs4923929	rs555001	42439376	43545728	1.11	37	29.901	
15	rs11351249	rs2623989	72105929	73545732	1.44	23	62.600	
15	rs752270	rs4777755	87099537	93510603	6.41	171	37.492	
16	rs6564764	rs931713	80581631	85023855	4.44	121	36.713	
17	rs8071836	rs67749219	47238028	49711042	2.47	40	61.825	

**Supplementary Table S3. List and characteristics of primers used in this study**

Application	Gene	Exon	Forward (5'-3')	Reverse (5'-3')	Size (bp)
DNA	<i>IQCE</i>	12	CTCTGAGCTCACAACCAACAG	GTTGCCCAAGGGTTCACA	399
		16	TGGTTTCTGGTGTTCAGG	TTTTCAGTGGCTGGTCAGAA	382
	<i>ATP6V1B1</i>	3	AGCCGTGGGAAGTAGGTGTT	ATGTCCCATTGACCCCACT	385
	<i>TULP1</i>	12/13	GGGATGTAGGATCCCCTCA	CATGCCAGGAATGATGACG	481
cDNA	<i>IQCE</i>	3-4/5	ACATCGCCAAAGTCACCTTATC	AGGGCCTGGGTCAGACTT	158
			<i>GSTT1</i>	2-3/3	ACCTTGACGGAGAGTGTGG
	<i>SLC30A8</i>	9-10/10	CCATCTTACTCATGGAAGGTGTG	GCTGTAGCAACATGAGCTGAGA	154
	<i>RSPO1</i>	4/5	TGGAGAGGAACGACATCCG	CCTTACACTTGGTGCAGAAGTTA	156
	<i>FGF9</i>	2/3	GGCCTGGTCAGCATTTCGAG	GTATCGCCTTCCAGTGCCAC	180
	<i>SFRP1</i>	1/2	GCTTCTACTGGCCCGAGATG	TGGCCTCAGATTTCAACTCGT	180
	<i>MECOM</i>	4/5	AAACTCGAAAGCGAGAATGATCT	TGGTGGCGAATTAATTGGACTT	185
	<i>ALPL</i>	6-7/7	AACATCAGGGACATTGACGTG	GTATCTCGGTTTGAAGCTCTTCC	159
	<i>IGF2</i>	2-3/3-4	TTCTACTTCAGCAGGCCCG	GAAGTTGTCCGGAAGCACG	168
	<i>GAPDH</i>	4-5/6	GGAGCGAGATCCCTCCAAAAT	GGCTGTTGCATACTTCTCATGG	197
	<i>HPRT1</i>	1-2/2-3	CCTGGCGTCGTGATTAGTGAT	AGACGTTCAGTCCTGTCCATAA	131
	<i>EFCAB7</i>	6/6-7	GTCAGCAACCAGGAAGTTCA	CATGTGTTGCCAGTCCTTTATT	115
	<i>EVC2</i>	6-7/8	CTCGTCACGGAACAGAACAC	CAGGTCAGCACAAGGGAGAG	135
	<i>EVC</i>	2/4	TCGAAGGACAAGGAAGCTGT	TTTCATGCAGAGACGGGTTG	148
	<i>GLI2</i>	2/3	AAGCAAGAAGCCAAAAGTGG	TGGTACCTTCTTCTGGTG	188
	<i>GLI3</i>	2/3	GGCATTTTTGGTCAAGAGA	GGACATTCTGTGGCTGCATA	238
	<i>SMO</i>	2/3	TGCCCAAGTGTGAGAATGAC	TACCAGCTCTTGGGGTTGTC	229

**Supplementary Table S4.** Comparison of haplotypes of individuals A.IV-1 and B.II-2.

ID	Gene	Chr	Position	Allele_1_A.IV-1	Allele_2_A.IV-1	Allele_1_II-2_Family2	Allele_1_II-2_Family2
7_2255974_C_T	<i>MAD1L1</i>	7	2255974	C	C	C	T
7_2256017_C_A	<i>MAD1L1</i>	7	2256017	C	C	C	A
7_2257612_G_A	<i>MAD1L1</i>	7	2257612	A	A	A	G
7_2259134_C_T	<i>MAD1L1</i>	7	2259134	C	C	C	T
7_2260715_G_T	<i>MAD1L1</i>	7	2260715	T	T	T	T
7_2269552_G_T	<i>MAD1L1</i>	7	2269552	T	T	T	G
7_2278663_T_A	<i>FTSJ2</i>	7	2278663	T	T	A	A
7_2279482_A_C	<i>FTSJ2</i>	7	2279482	A	A	C	C
7_2279851_T_C	<i>FTSJ2</i>	7	2279851	T	T	C	C
7_2281754_C_T	<i>FTSJ2</i>	7	2281754	C	C	C	T
7_2290328_C_T	<i>SNX8</i>	7	2290328	C	C	C	T
7_2290336_T_G	<i>SNX8</i>	7	2290336	T	T	T	G
7_2290372_C_T	<i>SNX8</i>	7	2290372	C	C	T	T
7_2290522_C_T	<i>SNX8</i>	7	2290522	T	T	C	C
7_2292881_-TG	<i>SNX8</i>	7	2292881	-	-	TG	TG
7_2296493_A_G	<i>SNX8</i>	7	2296493	G	G	G	G
7_2297006_A_G	<i>SNX8</i>	7	2297006	A	A	G	G
7_2303000_G_A	<i>SNX8</i>	7	2303000	A	A	G	G
7_2303109_G_A	<i>SNX8</i>	7	2303109	G	G	G	A
7_2303986_T_C	<i>SNX8</i>	7	2303986	C	C	C	C
7_2304191_A_C	<i>SNX8</i>	7	2304191	A	A	C	C
7_2309159_-A	<i>SNX8</i>	7	2309159	-	-	-	A
7_2309174_G_A	<i>SNX8</i>	7	2309174	G	G	G	A
7_2309185_G_A	<i>SNX8</i>	7	2309185	G	G	G	A
7_2309209_C_T	<i>SNX8</i>	7	2309209	C	C	C	T
7_2314652_G_A	<i>SNX8</i>	7	2314652	G	G	G	A
7_2315008_G_C	<i>SNX8</i>	7	2315008	G	G	C	C
7_2318254_ACTC_-	<i>SNX8</i>	7	2318254	-	-	ACTC	ACTC
7_2318268_C_T	<i>SNX8</i>	7	2318268	T	T	C	C
7_2318463_C_T	<i>SNX8</i>	7	2318463	C	T	C	C
7_2318472_C_A	<i>SNX8</i>	7	2318472	C	A	C	C
7_2349470_T_C	<i>SNX8</i>	7	2349470	T	T	T	C
7_2394991_C_T	<i>EIF3B</i>	7	2394991	C	C	C	T
7_2400306_T_C	<i>EIF3B</i>	7	2400306	C	C	C	T
7_2405885_C_A	<i>EIF3B</i>	7	2405885	A	A	A	C
7_2409035_C_T	<i>EIF3B</i>	7	2409035	C	C	C	T
7_2414142_A_G	<i>EIF3B</i>	7	2414142	G	G	G	A
7_2414365_A_G	<i>EIF3B</i>	7	2414365	A	A	A	G
7_2416822_A_G	<i>EIF3B</i>	7	2416822	A	A	A	G
7_2418645_A_G	<i>EIF3B</i>	7	2418645	A	A	A	G
7_2433729_C_A	NA	7	2433729	C	C	C	A
7_2434522_G_A	NA	7	2434522	G	G	A	A
7_2434596_TT_-	NA	7	2434596	TT	TT	TT	-
7_2434597_T_-	NA	7	2434597	T	T	T	-

7_2472429_C_A	CHST12	7	2472429	C	C	C	A
7_2472455_A_T	CHST12	7	2472455	A	A	A	T
7_2489311_C_G	NA	7	2489311	C	C	C	G
7_2489336_GTGTGT GCACCT_-	NA	7	2489336	GTGTGTGCACCT	GTGTGTGCACCT	GTGTGTGCACCT	-
7_2489382_A_G	NA	7	2489382	A	A	A	G
7_2515382_-C	NA	7	2515382	-	-	C	C
7_2517473_C_T	NA	7	2517473	C	C	C	T
7_2552986_A_C	LFNG	7	2552986	A	A	A	C
7_2553030_T_C	LFNG	7	2553030	T	T	T	C
7_2557266_T_C	LFNG	7	2557266	T	T	C	C
7_2557377_T_C	LFNG	7	2557377	T	T	T	C
7_2565268_C_A	LFNG/MIR4648	7	2565268	C	C	A	A
7_2566433_G_A	LFNG/MIR4648	7	2566433	G	G	G	A
7_2577691_C_G	BRAT1	7	2577691	C	C	C	G
7_2577781_T_C	BRAT1	7	2577781	T	T	T	C
7_2578237_T_C	BRAT1	7	2578237	T	T	T	C
7_2578238_C_T	BRAT1	7	2578238	C	C	C	T
7_2578434_G_T	BRAT1	7	2578434	G	G	G	T
7_2578455_G_A	BRAT1	7	2578455	G	G	G	A
7_2580914_C_T	BRAT1	7	2580914	C	C	C	T
7_2581994_T_C	BRAT1	7	2581994	T	T	T	C
7_2582154_- _CCCCAGCCTCCCG GGTGT	BRAT1	7	2582154	CCCCAGCCTCC CGGGTGT	CCCCAGCCTCC CGGGTGT	-	-
7_2582233_A_G	BRAT1	7	2582233	A	A	A	G
7_2582381_T_A	BRAT1	7	2582381	T	T	T	A
7_2582584_G_A	BRAT1	7	2582584	G	G	G	A
7_2583165_G_A	BRAT1	7	2583165	G	G	G	A
7_2583328_C_T	BRAT1	7	2583328	C	C	C	T
7_2584523_T_C	BRAT1	7	2584523	T	T	T	C
7_2587122_T_C	BRAT1	7	2587122	T	T	T	C
7_2598757_T_C	IQCE	7	2598757	T	T	T	C
7_2604768_GTGT_-	IQCE	7	2604768	GTGT	GTGT	GTGT	-
7_2604769_T_*	IQCE	7	2604769	T	T	T	*
7_2604816_C_T	IQCE	7	2604816	C	C	C	T
7_2611878_T_C	IQCE	7	2611878	T	T	T	C
7_2612225_G_A	IQCE	7	2612225	A	A	G	G
7_2613042_G_C	IQCE	7	2613042	G	G	G	C
7_2618070_-A	IQCE	7	2618070	-	-	-	A
7_2619341_GT_-	IQCE	7	2619341	GT	GT	GT	-
7_2619363_G_C	IQCE	7	2619363	G	G	G	C
7_2623351_G_A	IQCE	7	2623351	G	G	G	A
7_2623801_G_T	IQCE	7	2623801	G	G	G	T
7_2625907_TGTCCC GGAG_-	IQCE	7	2625907	-	-	-	TGTCCCGGAG
7_2626039_CAGGGA ATGG_-	IQCE	7	2626039	CAGGGAATGG	CAGGGAATGG	CAGGGAATGG	-
7_2627310_T_A	IQCE	7	2627310	T	T	T	A
7_2627579_AG_-	IQCE	7	2627579	AG	AG	-	-
7_2632477_A_G	IQCE	7	2632477	A	A	A	G

7_2634518_AGAG_-	<i>IQCE</i>	7	2634518	AGAG	AGAG	AGAG	-
7_2636689_C_T	<i>IQCE</i>	7	2636689	C	C	C	T
7_2638347_C_G	<i>IQCE</i>	7	2638347	C	C	C	G
7_2640985_G_C	<i>IQCE</i>	7	2640985	G	G	G	C
7_2647048_C_G	<i>IQCE</i>	7	2647048	C	C	C	G
7_2691201_C_A	<i>TTYH3</i>	7	2691201	A	A	C	C
7_2699687_-_G	<i>TTYH3</i>	7	2699687	-	-	G	G
7_2701809_T_C	<i>TTYH3</i>	7	2701809	T	C	T	T
7_2742097_T_C	<i>AMZ1</i>	7	2742097	C	C	T	T
7_2749565_G_A	<i>AMZ1</i>	7	2749565	G	G	G	A
7_2752152_G_A	<i>AMZ1</i>	7	2752152	G	G	G	A
7_2752487_G_A	<i>AMZ1</i>	7	2752487	A	A	G	G
7_2773035_-_AT	<i>GNA12</i>	7	2773035	-	-	AT	AT
7_2802173_T_C	<i>GNA12</i>	7	2802173	T	T	T	C
7_2802522_T_C	<i>GNA12</i>	7	2802522	T	T	T	C
7_2834869_C_T	<i>GNA12</i>	7	2834869	C	C	T	T
7_2854010_G_A	<i>GNA12</i>	7	2854010	G	G	A	A
7_2946461_T_C	<i>CARD11</i>	7	2946461	C	C	C	T
7_2952912_G_A	<i>CARD11</i>	7	2952912	G	G	G	A
7_2957005_T_C	<i>CARD11</i>	7	2957005	C	C	C	T
7_2958298_C_T	<i>CARD11</i>	7	2958298	C	C	T	T
7_2962241_G_A	<i>CARD11</i>	7	2962241	A	A	G	G
7_2962753_G_A	<i>CARD11</i>	7	2962753	G	G	G	A
7_2966445_T_G	<i>CARD11</i>	7	2966445	T	T	T	G
7_2966466_GT_-	<i>CARD11</i>	7	2966466	GT	GT	GT	-
7_2968195_G_A	<i>CARD11</i>	7	2968195	G	G	G	A
7_2968361_G_A	<i>CARD11</i>	7	2968361	G	G	G	A
7_2968486_G_C	<i>CARD11</i>	7	2968486	G	G	G	C
7_2985364_AC_-	<i>CARD11</i>	7	2985364	A	C	-	-

**Supplementary Table S5.** Differentially expressed transcripts in A.IV-1 fibroblasts.

Gene name	Ensembl gene id	Description	log2(FC)	P-value	Adjusted p-value
<i>ABCA8</i>	ENSG00000141338	ATP-binding cassette, sub-family A (ABC1), member 8	-4.205829134	3.52E-05	0.019362186
<i>AC011294.3</i>	ENSG00000233539	Uncharacterized protein	3.130226421	2.55E-07	0.000427239
<i>ACCS</i>	ENSG00000110455	1-aminocyclopropane-1-carboxylate synthase homolog (Arabidopsis)(non-functional)	-0.953037549	1.28E-07	0.00025011
<i>ADAMTSL4</i>	ENSG00000143382	ADAMTS-like 4	-3.153983754	5.00E-07	0.000732904
<i>ALDH1A1</i>	ENSG00000165092	aldehyde dehydrogenase 1 family, member A1	-6.206001924	1.99E-08	5.40E-05
<i>ALPL</i>	ENSG00000162551	alkaline phosphatase, liver/bone/kidney	-3.039649649	0.000113858	0.038536558
<i>APBA2</i>	ENSG00000034053	amyloid beta (A4) precursor protein-binding, family A, member 2	-2.732442847	6.79E-05	0.028790694
<i>ARHGAP28</i>	ENSG00000088756	Rho GTPase activating protein 28	2.800539497	1.58E-05	0.010684692
<i>C10orf54</i>	ENSG00000107738	chromosome 10 open reading frame 54	-1.64124697	9.82E-05	0.035643913
<i>C1QTNF9B-AS1</i>	ENSG00000205861	C1QTNF9B antisense RNA 1	-2.047191404	3.88E-05	0.019780947
<i>C2CD4A</i>	ENSG00000198535	C2 calcium-dependent domain containing 4A	5.96982663	8.95E-07	0.001085914
<i>C4A</i>	ENSG00000244731	complement component 4A (Rodgers blood group)	-6.329532092	5.21E-11	3.66E-07
<i>CCDC180</i>	ENSG00000197816	coiled-coil domain containing 180	2.351169896	3.81E-05	0.019780947
<i>CD200</i>	ENSG00000091972	CD200 molecule	-4.053753384	0.000135641	0.044622222
<i>CHI3L1</i>	ENSG00000133048	chitinase 3-like 1 (cartilage glycoprotein-39)	4.852100597	6.43E-07	0.000837885
<i>CHN1</i>	ENSG00000128656	chimerin 1	-1.617295158	2.25E-09	1.32E-05
<i>CHRD2</i>	ENSG00000054938	chordin-like 2	-4.241249176	0.000101018	0.03591767
<i>CNTN1</i>	ENSG00000018236	contactin 1	-4.135254047	8.76E-05	0.032879001
<i>COL11A1</i>	ENSG00000060718	collagen, type XI, alpha 1	-4.558619592	6.30E-05	0.027357715
<i>CSF2RA</i>	ENSG00000198223	colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage)	3.072895701	6.90E-05	0.028906369
<i>CYP21A1P</i>	ENSG00000204338	cytochrome P450, family 21, subfamily A, polypeptide 1 pseudogene	-4.61429374	6.95E-06	0.005439088
<i>DUSP2</i>	ENSG00000158050	dual specificity phosphatase 2	3.527490256	1.70E-05	0.01128586
<i>ECM1</i>	ENSG00000143369	extracellular matrix protein 1	-1.261675626	5.81E-07	0.000786019
<i>EMCN</i>	ENSG00000164035	endomucin	3.361027658	4.95E-05	0.0229427
<i>EYA4</i>	ENSG00000112319	eyes absent homolog 4 (Drosophila)	-4.200503943	4.26E-05	0.020550766
<i>FAM21B</i>	ENSG00000152726	family with sequence similarity 21, member B	-4.995162913	7.93E-08	0.000164108
<i>FGF9</i>	ENSG00000102678	fibroblast growth factor 9	4.555259473	5.95E-08	0.000133394
<i>FOSB</i>	ENSG00000125740	FBJ murine osteosarcoma viral oncogene homolog B	3.628551087	8.80E-05	0.032879001
<i>FYB</i>	ENSG00000082074	FYN binding protein	-4.90648728	3.03E-05	0.016914087
<i>GRPR</i>	ENSG00000126010	gastrin-releasing peptide receptor	-4.422707449	0.000144048	0.046095412
<i>GSTT1</i>	ENSG00000184674	glutathione S-transferase theta 1	-9.843482395	6.94E-60	2.44E-55
<i>HIVEP3</i>	ENSG00000127124	human immunodeficiency virus type I enhancer binding protein 3	-1.599082154	6.64E-05	0.028486515
<i>HOXC10</i>	ENSG00000180818	homeobox C10	-4.942455716	2.36E-05	0.014065434
<i>HOXC-AS3</i>	ENSG00000251151	HOXC cluster antisense RNA 3	-5.14737488	6.01E-05	0.026775571
<i>IGF2</i>	ENSG00000167244	insulin-like growth factor 2 (somatomedin A)	-4.063105964	3.83E-05	0.019780947
<i>IGFBP1</i>	ENSG00000146678	insulin-like growth factor binding protein 1	2.926653654	1.58E-05	0.010684692



<i>IGFBP3</i>	ENSG00000146674	insulin-like growth factor binding protein 3	1.932377253	2.71E-06	0.002574523
<i>IL17RB</i>	ENSG00000056736	interleukin 17 receptor B	3.133016005	8.91E-05	0.032879001
<i>IL20RA</i>	ENSG00000016402	interleukin 20 receptor, alpha	3.104556086	3.12E-06	0.002749475
<i>IQCE</i>	ENSG00000106012	IQ motif containing E	-2.198435274	1.75E-43	3.08E-39
<i>IRF6</i>	ENSG00000117595	interferon regulatory factor 6	-3.931349479	7.74E-05	0.030946745
<i>IRX1</i>	ENSG00000170549	iroquois homeobox 1	1.496134127	2.70E-05	0.015596484
<i>IRX2</i>	ENSG00000170561	iroquois homeobox 2	1.161748374	4.51E-09	1.77E-05
<i>JAZF1-AS1</i>	ENSG00000234336	JAZF1 antisense RNA 1	-2.132735824	0.000107563	0.037119743
<i>KLKB1</i>	ENSG00000164344	kallikrein B, plasma (Fletcher factor) 1	2.478417669	8.84E-05	0.032879001
<i>LAMC1</i>	ENSG00000135862	laminin, gamma 1 (formerly LAMB2)	-1.201150815	2.84E-07	0.000453983
<i>LHX9</i>	ENSG00000143355	LIM homeobox 9	4.298162126	1.45E-06	0.001650048
<i>LINC00537</i>	ENSG00000232815	long intergenic non-protein coding RNA 537	-3.494641854	3.83E-05	0.019780947
<i>LINC00578</i>	ENSG00000228221	long intergenic non-protein coding RNA 578	-4.541164267	7.46E-09	2.39E-05
<i>LINC01048</i>	ENSG00000230390	long intergenic non-protein coding RNA 1048	2.214309895	3.99E-05	0.019780947
<i>LINC01117</i>	ENSG00000224577	long intergenic non-protein coding RNA 1117	-2.17227763	2.17E-05	0.01316806
<i>LRRC16A</i>	ENSG00000079691	leucine rich repeat containing 16A	-1.784218256	2.43E-05	0.014245821
<i>MAP6</i>	ENSG00000171533	microtubule-associated protein 6	-1.79656788	1.30E-06	0.001530605
<i>MAP7D2</i>	ENSG00000184368	MAP7 domain containing 2	2.359306267	2.66E-06	0.002574523
<i>MBP</i>	ENSG00000197971	myelin basic protein	-3.402601353	7.14E-05	0.029220727
<i>MCCC1</i>	ENSG00000078070	methylcrotonoyl-CoA carboxylase 1 (alpha)	-0.554719912	0.000128238	0.042584679
<i>MECOM</i>	ENSG00000085276	MDS1 and EVI1 complex locus	-4.511721791	3.14E-09	1.58E-05
<i>METTL9</i>	ENSG00000197006	methyltransferase like 9	0.906519962	3.86E-08	9.70E-05
<i>MFGE8</i>	ENSG00000140545	milk fat globule-EGF factor 8 protein	-2.694104724	7.44E-07	0.000934855
<i>MLLT3</i>	ENSG00000171843	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, <i>Drosophila</i> ); translocated to, 3	0.969013225	2.10E-05	0.01316806
<i>MRV11-AS1</i>	ENSG00000177112	MRV11 antisense RNA 1	-3.884388307	1.04E-05	0.007509783
<i>MYO16</i>	ENSG00000041515	myosin XVI	-1.962481605	0.000117031	0.039233207
<i>NDP</i>	ENSG00000124479	Norrie disease (pseudoglioma)	3.545635871	1.83E-05	0.011727901
<i>NR5A2</i>	ENSG00000116833	nuclear receptor subfamily 5, group A, member 2	-4.138425423	8.61E-06	0.006447287
<i>NTRK3</i>	ENSG00000140538	neurotrophic tyrosine kinase, receptor, type 3	2.777060262	2.95E-05	0.016740854
<i>PKIA</i>	ENSG00000171033	protein kinase (cAMP-dependent, catalytic) inhibitor alpha	-2.779337899	4.26E-05	0.020550766
<i>PNLDC1</i>	ENSG00000146453	poly(A)-specific ribonuclease (PARN)-like domain containing 1	2.95490495	2.27E-07	0.000399745
<i>PNMA6C</i>	ENSG00000235961	paraneoplastic Ma antigen family member 6C	-4.033899706	0.000151809	0.047958078
<i>POMZP3</i>	ENSG00000146707	POM121 and ZP3 fusion	-1.717492119	3.12E-06	0.002749475
<i>POSTN</i>	ENSG00000133110	periostin, osteoblast specific factor	2.83484857	2.14E-05	0.01316806
<i>PRKAR1B</i>	ENSG00000188191	protein kinase, cAMP-dependent, regulatory, type I, beta	-1.732349926	1.78E-05	0.011579369
<i>PRLR</i>	ENSG00000113494	prolactin receptor	3.300099077	0.000140582	0.045398939
<i>PTK2B</i>	ENSG00000120899	protein tyrosine kinase 2 beta	-1.47679232	6.10E-05	0.026827045
<i>RGCC</i>	ENSG00000102760	regulator of cell cycle	-3.93550099	7.06E-05	0.029220727
<i>RHBDD3</i>	ENSG00000100263	rhomboid domain containing 3	-0.996433377	1.05E-05	0.007509783
<i>RSPO1</i>	ENSG00000169218	R-spondin 1	-4.313022549	5.98E-05	0.026775571
<i>S1PR1</i>	ENSG00000170989	sphingosine-1-phosphate receptor 1	-3.093512898	5.52E-06	0.00452094

<i>SCN1A</i>	ENSG00000144285	sodium channel, voltage-gated, type I, alpha subunit	-4.906276861	3.80E-06	0.003184547
<i>SCN9A</i>	ENSG00000169432	sodium channel, voltage-gated, type IX, alpha subunit	-3.142530704	1.83E-06	0.001898572
<i>SDK1</i>	ENSG00000146555	sidekick cell adhesion molecule 1	3.188897352	4.55E-05	0.021372366
<i>SFRP1</i>	ENSG00000104332	secreted frizzled-related protein 1	-4.10661345	5.21E-07	0.000732904
<i>SHE</i>	ENSG00000169291	Src homology 2 domain containing E	3.455259022	1.63E-06	0.001733572
<i>SLA</i>	ENSG00000155926	Src-like-adaptor	-2.701260288	3.99E-05	0.019780947
<i>SLC16A11</i>	ENSG00000174326	solute carrier family 16, member 11	-2.677392243	4.37E-05	0.020765473
<i>SLC30A8</i>	ENSG00000164756	solute carrier family 30 (zinc transporter), member 8	6.104148946	1.35E-07	0.000250798
<i>SLC35D3</i>	ENSG00000182747	solute carrier family 35, member D3	9.179494017	2.46E-31	2.89E-27
<i>SLC4A3</i>	ENSG00000114923	solute carrier family 4 (anion exchanger), member 3	-2.366325383	7.28E-05	0.029469479
<i>SMOC1</i>	ENSG00000198732	SPARC related modular calcium binding 1	-5.03813069	0.000104067	0.036268803
<i>SPON2</i>	ENSG00000159674	spondin 2, extracellular matrix protein	-2.639592845	0.000112611	0.038484373
<i>SRD5A2</i>	ENSG00000049319	steroid-5-alpha-reductase, alpha polypeptide 2 (3-oxo-5 alpha-steroid delta 4-dehydrogenase alpha 2)	4.448074003	8.97E-05	0.032879001
<i>STAG3L2</i>	ENSG00000160828	stromal antigen 3-like 2 (pseudogene)	-0.881204831	9.94E-05	0.035716372
<i>TCF21</i>	ENSG00000118526	transcription factor 21	-6.04609974	1.52E-06	0.001676294
<i>ULK4</i>	ENSG00000168038	unc-51 like kinase 4	1.147447551	6.46E-09	2.28E-05
<i>WASF3</i>	ENSG00000132970	WAS protein family, member 3	0.972046987	2.81E-06	0.002603617
<i>XKR9</i>	ENSG00000221947	XK, Kell blood group complex subunit-related family, member 9	-1.989973823	0.000159916	0.049814665
<i>ZBTB16</i>	ENSG00000109906	zinc finger and BTB domain containing 16	-4.844890791	7.83E-05	0.030946745
<i>ZFPM2</i>	ENSG00000169946	zinc finger protein, FOG family member 2	-3.60011069	5.25E-14	4.62E-10
<i>ZNF804A</i>	ENSG00000170396	zinc finger protein 804A	-5.483825868	4.34E-09	1.77E-05

**Supplementary Table S6.** GO enrichment analysis. Show the distribution of GO terms exhibiting statistical significant differences of mis-regulated transcripts in A.IV-1 fibroblast using WebGestalt 2019 (Wang, Vasaikar, Shi, Greer, & Zhang, 2017).

GO terms	Description	Gene set size	Overlap	Expected Value	Enrichment Ratio	pvalue	FDR	overlap_id	user_id
GO:0009888	tissue development	1613	24	7.625171279802685	3.1474702822178497	1.528766712111107e-7	0.0012670418509976855	249;1116;1301;1844;1893;1901;2185;2254;2494;3484;3664;4300;4916;6422;6943;7704;10631;23414;25884;28984;54507;59269;79192;153572	TCF21;ZBTB16;COL11A1;CHRD1;NR5A2;SFRP1;RGCC;IRF6;ZFPM2;ADAMTSL4;S1PR1;ALPL;HIVEP3;PTK2B;ECM1;MLL3;IRX2;IRX1;NTRK3;POSTN;IGFBP1;DUSP2;FGF9;CHI3L1
GO:0001503	ossification	339	11	1.6025623458481775	6.864007524261469	4.940246167706164e-7	0.0020472380118974343	249;1301;1893;1901;2185;2254;3486;6422;7704;25884;64093	SMOC1;ZBTB16;COL11A1;CHRD1;SFRP1;S1PR1;ALPL;PTK2B;ECM1;IGFBP3;FGF9
GO:0030038	contractile actin filament bundle assembly	76	5	0.35927651411345574	13.916857360793287	2.9031950272262463e-5	0.060154200964127824	1901;2185;6422;28984;55604	SFRP1;RGCC;S1PR1;CARMIL1;PTK2B
GO:0043149	stress fiber assembly	76	5	0.35927651411345574	13.916857360793287	2.9031950272262463e-5	0.060154200964127824	1901;2185;6422;28984;55604	SFRP1;RGCC;S1PR1;CARMIL1;PTK2B
GO:0048584	positive regulation of response to stimulus	1768	21	8.35790627569197	2.512590989573087	4.138463705938289e-5	0.06101540527461857	1116;1123;1893;1901;2010;2185;2254;2533;3486;3818;4300;4916;5618;6422;6503;10417;28984;53832;55540;126669;284654	FYB;RSP01;EYAA4;SFRP1;RGCC;S1PR1;SLA;SPON2;CHN1;PTK2B;ECM1;MLL3;IGFBP3;KLB1;NTRK3;IL2RA;IL17RB;PLLR;SHE;FGF9;CHI3L1
GO:0007378	segment specification	15	3	0.07090983831186628	42.3072463768116	4.417138412737831e-5	0.06101540527461857	4300;79192;153572	MLL3;IRX2;IRX1
GO:0030278	regulation of ossification	169	6	0.7980175116470265	7.510162078723952	1.4457514683041683e-4	0.12546232977013955	1893;1901;2185;6422;7704;64093	SMOC1;ZBTB16;SFRP1;S1PR1;PTK2B;ECM1
GO:0048646	anatomical structure formation involved in morphogenesis	876	13	4.14113455741299	3.1392363179141025	2.015277408380367e-4	0.12546232977013955	1116;1301;1844;1893;1901;2185;2254;6422;6943;28984;79192;153572;221935	TCF21;COL11A1;SFRP1;RGCC;S1PR1;PTK2B;ECM1;IRX2;IRX1;SDK1;DUSP2;FGF9;CHI3L1
GO:0001501	skeletal system development	439	9	2.075294801260619	4.33673368327226	2.1232290176698143e-4	0.12546232977013955	249;1116;1301;2254;3226;6422;7704;10631;25884	HOXC10;ZBTB16;COL11A1;CHRD1;SFRP1;ALPL;POSTN;FGF9;CHI3L1
GO:0072047	proximal/distal pattern formation involved in nephron development	5	2	0.023636612770622088	84.61449275362318	2.182366244216638e-4	0.12546232977013955	79192;153572	IRX2;IRX1

**Supplementary Table S7. Summary of transcriptome mapping results.**

Sample ID	#total reads	#aligned reads	% aligned reads	#uniquely mapped	% uniquely aligned reads
A.IV-1.1	306,265,974	303,190,132	99.00	293,189,972	96.70
A.IV-1.2	344,782,588	341,370,324	99.01	330,182,914	96.72
ARN1.1	338,129,498	335,557,854	99.24	325,618,112	97.04
ARN1.2	301,339,526	298,096,006	98.92	289,521,446	97.12
ARN2.1	311,917,536	307,641,094	98.63	296,837,798	96.49
ARN2.2	336,751,354	333,728,602	99.10	322,903,810	96.76
ARN3.1	403,718,870	400,433,942	99.19	383,740,470	95.83
ARN3.2	299,879,030	296,763,108	98.96	286,203,188	96.44
ARN4.1	318,588,494	315,510,058	99.03	304,438,002	96.49
ARN4.2	320,107,070	316,729,460	98.94	304,559,776	96.16
ARN5.1	363,584,984	360,496,138	99.15	349,081,194	96.83
ARN6.1	324,984,406	321,703,444	98.99	311,631,678	96.87
ARN6.2	364,763,088	361,575,516	99.13	349,853,468	96.76

## SUPPLEMENTARY REFERENCES

- Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., & Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*, *14*(4), R36. doi:10.1186/gb-2013-14-4-r36
- Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*, *10*(3), R25. doi:10.1186/gb-2009-10-3-r25
- Nevers, Y., Prasad, M. K., Poidevin, L., Chennen, K., Allot, A., Kress, A., . . . Lecompte, O. (2017). Insights into Ciliary Genes and Evolution from Multi-Level Phylogenetic Profiling. *Mol Biol Evol*, *34*(8), 2016-2034. doi:10.1093/molbev/msx146
- Wang, J., Vasaikar, S., Shi, Z., Greer, M., & Zhang, B. (2017). WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Research*, *45*(W1), W130-W137. doi:10.1093/nar/gkx356

Le projet **IV**

portant sur *BBS5*





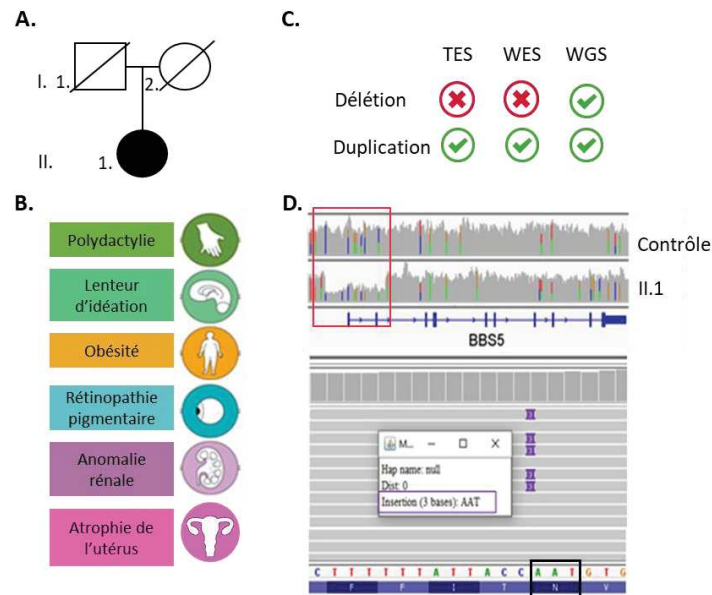
## I. Le synopsis

En 2019, le laboratoire a identifié deux mutations dans le gène *BBS5* chez l'enfant unique d'une famille caucasienne non consanguine. La patiente (II.1) présente les manifestations cliniques typiques du BBS c'est-à-dire une rétinite pigmentaire, une obésité, une polydactylie, une insuffisance rénale, une agénésie rénale, une déficience intellectuelle et une atrophie de l'utérus (Figure 45A et B). La recherche de mutation a été réalisée à la fois au Laboratoire de diagnostic génétique (HUS) et au Laboratoire de génétique médicale (UMR\_S 1112). Le matériel biologique des parents n'était pas disponible et ils sont décédés depuis. Nous avons eu accès aux cellules de la patiente (fibroblastes issus d'une biopsie de peau) maintenues en culture au laboratoire et dans le cas de cette famille, seul l'ADN de la patiente a été analysé.

La recherche de mutation a débuté par un séquençage ciblé (Sanger) des mutations récurrentes du BBS qui s'est révélé négatif. Ensuite, un panel de gènes (TES) par séquençage haut débit incluant 58 gènes (responsables de BBS et d'autres ciliopathies telles que le syndrome de Joubert et celui de Meckel-Gubler) a été réalisé aux HUS. Une duplication du triplet AAT (c.550\_552dupAAT, p.N184dup) dans l'exon 6 de *BBS5*, à l'état hétérozygote, a alors été observée. Ce premier variant est dit de signification inconnue (VSI, classe 3) et aucune autre variation, à l'état hétérozygote, n'a été mise en évidence dans *BBS5*. En particulier aucune délétion d'un ou plusieurs exons n'a pu être identifiée suite à l'analyse de la zone de *BBS5* sous IGV (ou avec d'autres logiciels). Ce type de variant est difficilement analysable en l'état car peu d'éléments sont disponibles pour son analyse. Cette première approche n'a donc pas permis de conclure sur ce cas.

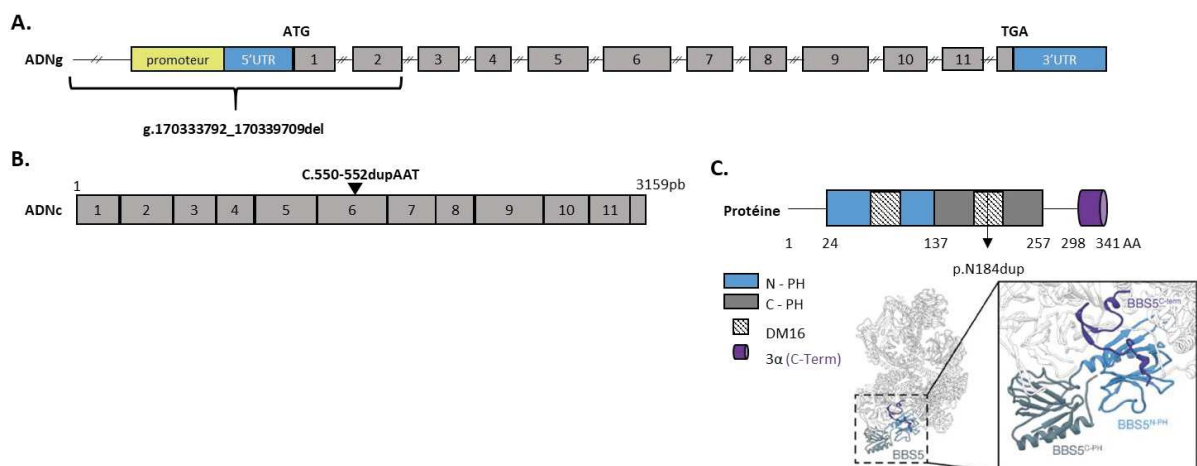
Un séquençage de l'exome complet (WES) a ensuite été mis en œuvre au sein du laboratoire de recherche (UMR\_S 1112). Seule la duplication à l'état hétérozygote du triplet AAT (c.550\_552dupAAT) dans l'exon 6 de *BBS5* a de nouveau été identifiée.

C'est finalement le génome (WGS) qui a permis de mettre en évidence un second évènement, à l'état hétérozygote, correspondant à une large délétion de 5920 pb englobant les exons 1 et 2 de *BBS5* (chr2: g.170333792\_170339709del). Puisque la délétion englobe également le promoteur, aucune expression de l'allèle en question n'est attendue. Ce 2<sup>ème</sup> variant conduisant très probablement à une perte de fonction, il a donc été classé « pathogène » (variant de classe 4) (Figure 45C et D).

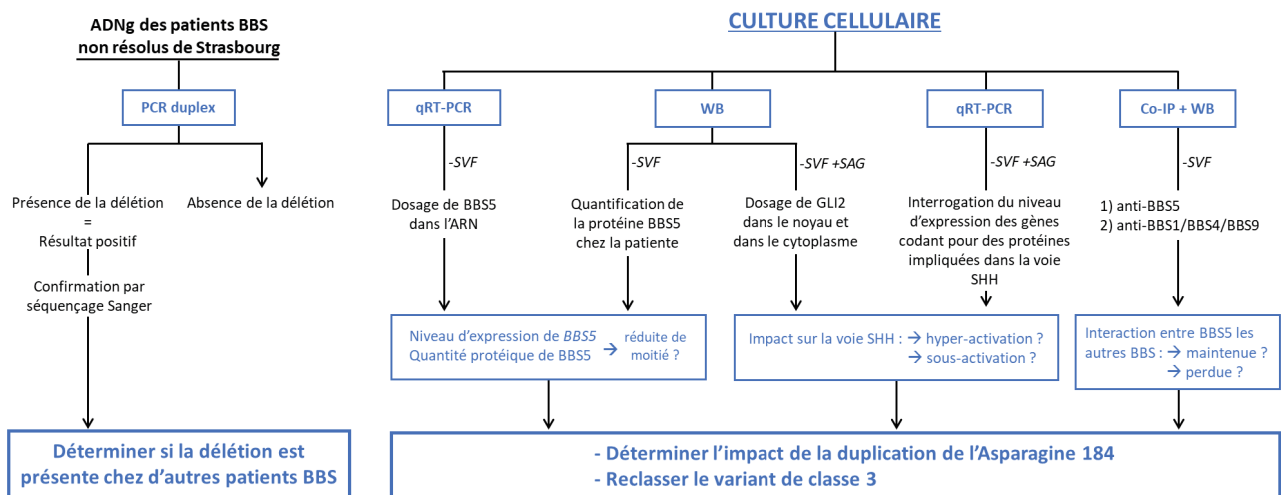


**Figure 45 : Description de la famille et des mutations identifiées.** A. Arbre de la famille. B. Le phénotype de la patiente II.1 est typique du BBS. C. Les technologies ayant permis la détection des mutations. D. Les visualisations de la délétion à l'état hétérozygote qui emporte les exons 1 et 2 du gène et de la duplication à l'état hétérozygote dans l'exon 6 de BBS5 dans IGV (Integrative Genomics Viewer). TES : target exome sequencing, WES : whole exome sequencing, WGS : whole genome sequencing

Les deux variations n'ont jamais été rapportées auparavant et la pathogénicité de l'une d'elles est incertaine. Le principal objectif de ce projet est de déterminer la pathogénicité de la duplication et de mieux comprendre la fonction de BBS5 en analysant l'impact de la mutation faux-sens. Un logigramme d'expériences a été pensé puis mis en œuvre pour répondre aux différentes problématiques (Figure 47). L'ensemble du matériel biologique (ADN, ARN, protéines) nécessaire aux expériences provient des cellules de la patiente ou de contrôles cultivées dans différentes conditions expérimentales.



**Figure 46 : Localisation des mutations.** A. Représentation du gène BBS5 avec le positionnement de la délétion. B. Représentation de l'ADNc de BBS5 avec l'ajout de la duplication. C. Représentation de la protéine BBS5 avec les domaines connus et la localisation de la mutation. Représentation 3D de la position de BBS5 à la base du BBsome. Visualisation de la structure de BBS5 avec les domaines d'homologie à la pleckstrine (BBS5N-PH et BBS5C-PH) et de l'extrémité C-terminale étendue (Singh et al., 2020).



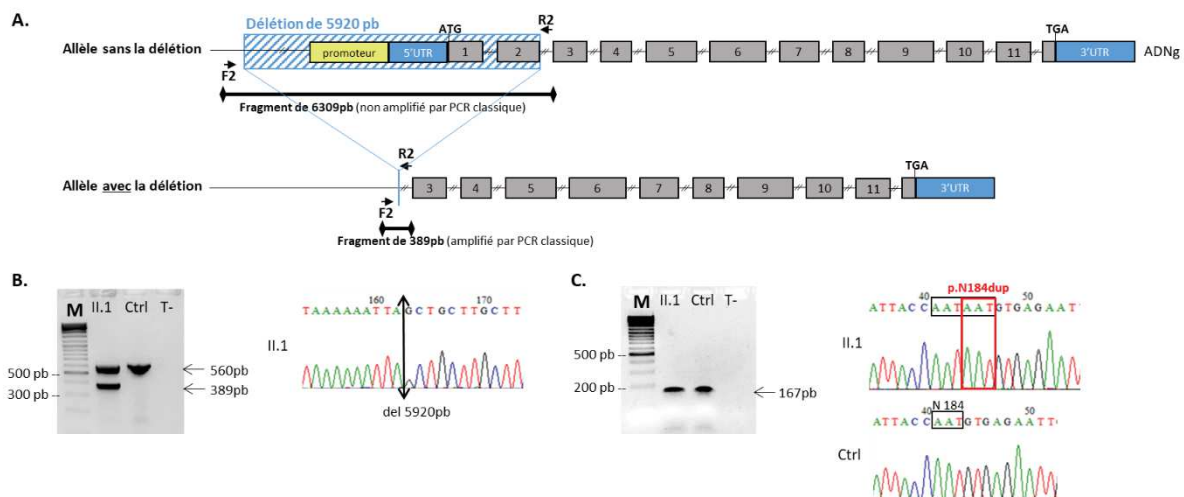
**Figure 47 : Logigramme.** Résumé des expériences mises en œuvre pour répondre aux questions posées et aux objectifs du projet. Mon travail est représenté en bleu.

## II. Les résultats

### 1. Les expériences portant sur la délétion

#### A) Détection par PCR et séquençage Sanger

Afin de confirmer l'existence de la délétion des 2 premiers exons et comme nous avons séquencé le génome complet de la patiente, nous avons obtenu précisément le point de cassure de cet événement. Nous avons alors dessiné des amorces F2/R2 dont la localisation est expliquée Figure 48A. Le premier objectif était de confirmer la présence de la mutation chez la patiente avec un test de PCR duplex sur son ADNg. Dans le cas où la délétion est présente, un fragment de 389 pb est amplifié. Dans un cas normal (sans délétion), il n'y a pas d'amplification en raison d'une taille trop importante du fragment (>6000 pb) non amplifiable dans les conditions de PCR appliquées (Figure 48A). Un second couple d'amorces contrôle (taille de la bande de 560 pb) ciblant un autre gène a été ajouté dans le mix PCR pour s'assurer que la PCR fonctionne bien dans les conditions expérimentales utilisées (Figure 48B). Ainsi, un individu porteur de la délétion a un profil avec 2 bandes (comme chez la patiente II.1) alors qu'en l'absence de la délétion, une seule bande est observée (celle de 560 pb, comme chez le contrôle). Les résultats de la PCR duplex montrent effectivement un fragment de 389 pb seulement chez la patiente confirmant ainsi la présence de la délétion. Le séquençage Sanger de l'ADNg confirme la présence d'une large délétion de 5920 pb.



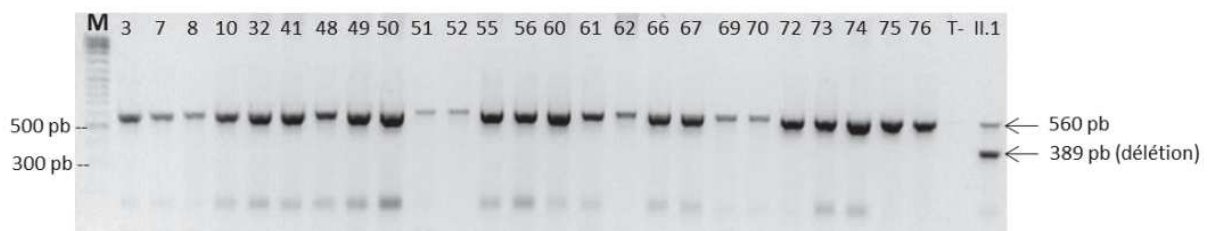
**Figure 48 : Détection des deux mutations par PCR et séquençage Sanger.** **A.** Positionnement du couple d'amorces F2/R2 sur les allèles porteurs ou non de la délétion. **B.** Résultats de la PCR duplex réalisée sur l'ADNg de la patiente et d'un contrôle. Résultats du séquençage sanger pour le fragment de la patiente : présence de la délétion **C.** Résultats de la PCR classique sur l'ADNg avec les amorces exon 6-7(F)-exon 8(R) et du séquençage sanger du fragment de la patiente et de celui d'un contrôle : confirmation de la duplication. M= marqueur, II.1= la patiente, Ctrl= contrôle, T-= témoin négatif.

### B) Exploration de la cohorte par PCR duplex

En raison de la difficulté de détecter la délétion des exons 1 et 2 par les techniques classiques et comme peu de génomes complets sont encore réalisés dans notre laboratoire, il y a une forte probabilité que d'autres patients de notre cohorte soient également porteurs de cette délétion. Nous avons donc utilisé le test de PCR duplex précédemment expliqué afin de réaliser une exploration de la cohorte BBS négative de Strasbourg. Par contre, il faut noter que nous n'avons pas recherché spécifiquement la 2<sup>ème</sup> mutation. Elle aurait été détectée par toutes les techniques NGS déjà mises en place dans les 2 laboratoires (hospitalier HUS et recherche LGM).

Au laboratoire de génétique médicale (LGM), nous avons testé une cohorte de 62 patients BBS non résolus pour lesquels l'ADNg était disponible. Les résultats des PCRs duplex n'ont montré aucun patient portant la délétion de l'exon 1 et 2 du gène *BBS5* (pas d'amplification d'un fragment de 389 pb) hormis notre patiente (Figure 49). En parallèle, le laboratoire de diagnostic génétique (HUS) a testé 156 ADNg de patients non résolus atteints de RP isolés et 222 ADNg de patients BBS non résolus. Les doublons possibles entre les 2 laboratoires ont été mis en évidence et écartés avant le début de l'expérience. Le résultat ne montre pas non plus d'autres patients portant la délétion, hormis notre patiente. Malgré l'exploration de 444 patients non résolus atteints de ciliopathie dans la cohorte de Strasbourg, la mutation (chr2: g.170333792\_170339709del) a été détectée uniquement chez la patiente qui fait l'objet de cette étude.

Afin de déterminer la fréquence de cette délétion dans la population générale, un screening bioinformatique a également été mis en place dans 863 génomes accessibles au laboratoire. La délétion n'a été détectée dans aucun des génomes analysés et n'est pas présente dans la base de données gnomAD SV.



**Figure 49 : Screening de la cohorte de patients BBS non résolus (LGM).** Résultats des PCR duplex réalisées sur un échantillon de 25 ADNg sur les 62 testés. M : marqueur, II.1 : la patiente. T- : témoin négatif

### C) L'étude de l'expression de l'ARNm et protéique pour BBS5

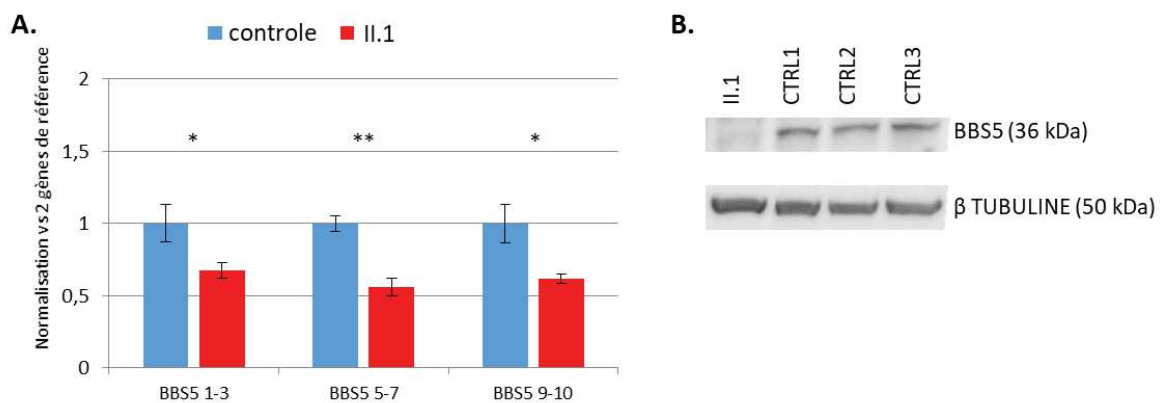
Puisque la délétion englobe le promoteur du gène *BBS5*, l'expression de l'allèle porteur devrait être nulle : le seul allèle exprimé devrait être celui avec la duplication. La quantité d'ARN transcrite chez la

patiente devrait être diminuée de moitié par rapport à un contrôle, tout comme la quantité de protéine BBS5.

Pour vérifier nos hypothèses, nous avons étudié le niveau d'expression de l'ARNm de *BBS5* par qRT-PCR en utilisant trois couples d'amorces distribués le long du gène afin d'évaluer si les niveaux d'expressions à différents endroits du gène concordent. Les résultats de la quantification relative (par rapport aux deux gènes de référence *GAPDH* et *HPRT*) montrent une réduction de moitié de l'expression de l'ARNm de *BBS5* chez la patiente par rapport aux contrôles (Figure 50A). Donc, chez la patiente, un seul allèle est exprimé : l'allèle porteur de la duplication. Comme attendu, l'allèle porteur de la délétion n'est pas exprimé : la mutation abolit l'expression du gène car elle englobe le promoteur.

Pour poursuivre nos investigations et confirmer ces premiers résultats, nous avons déterminé le niveau d'expression protéique de BBS5 chez la patiente par rapport à des contrôles par WB, en utilisant comme protéine de référence la  $\beta$ -TUBULINE. En théorie, nous attendons 2 fois moins de BBS5 chez la patiente puisqu'un seul allèle est exprimé. Les résultats montrent que l'expression de BBS5 est fortement diminuée chez la patiente par rapport aux contrôles (Figure 50B).

Avec cette première vague d'expériences, nous avons confirmé qu'un seul allèle est transcrit et traduit chez la patiente : celui porteur de la duplication. Cette situation sera utilisée à notre avantage puisque les expériences réalisées par la suite vont révéler l'impact du seul allèle BBS5 exprimé : celui porteur de la duplication.



**Figure 50 : Un seul allèle est exprimé.** **A.** Résultats de la qRT-PCR réalisée pour déterminer le niveau d'expression de *BBS5* chez la patiente (triplicat technique) par rapport à des contrôles (triplicats biologiques). Quantification relative réalisée par rapport aux 2 gènes de références *GAPDH* et *HPRT*. **B.** Résultats du WB (n=3) détectant BBS5 et la protéine de référence  $\beta$  TUBULINE. Quantification de BBS5 chez II.1 (n=3) par rapport à la  $\beta$  TUBULINE. Application d'un T-test ( $p < 0,05$  \*,  $p < 0,01$  \*\*). M : marqueur, II.1 : la patiente. CTRL : contrôle

## 2. Les expériences portant sur la duplication

### A) Confirmation de la co-ségrégation familiale

Afin de confirmer la présence d'une mutation identifiée par TES, WES et WGS, une seconde technique est appliquée, à savoir ici, le séquençage Sanger.

Habituellement, cette étape permet de confirmer la co-ségrégation familiale en définissant le statut des mutations l'une par rapport à l'autre : en *trans* ou en *cis*. Normalement, les ADN des parents ou des apparentés sont séquencés. Dans le cas de 2 mutations en *trans* : chacun des parents est porteur d'un des 2 allèles mutés (hétérozygote) et a transmis l'allèle à son enfant qui se retrouve hétérozygote composite avec 2 mutations différentes dans le même gène. Dans le cas où les 2 mutations sont en *cis* : elles sont situées sur le même allèle et n'ont été transmises que par un seul parent. Ce cas de figure n'est pas envisagé dans le cadre des maladies récessives puisqu'il faut que les 2 allèles soient mutés pour que le phénotype se développe.

Dans le cas de cette famille, le matériel biologique des parents n'est pas disponible. Nous nous sommes alors intéressés à l'ARN du cas index. Un couple d'amorces choisi pour amplifier l'ADNc de part et d'autre de la duplication a été utilisé. Le résultat de la PCR ne permet pas de discriminer le fragment de 167 pb amplifié chez le contrôle et celui de la patiente de 170 pb, en cas de duplication. Les résultats du séquençage Sanger confirment la présence de la duplication du triplet AAT dans l'exon 6 (Figure 48C) et la co-ségrégation familiale. Un seul pic est observé au niveau de la duplication ce qui montre que la duplication est en *trans* par rapport à la délétion des exons 1 et 2. En effet, en s'intéressant à l'ADNc, seuls les allèles exprimés sont étudiés. Nous avons déterminé qu'un seul allèle est exprimé. Si la duplication avait été *en cis* de la délétion, c'est-à-dire, sur le même allèle, alors la duplication n'aurait pas été visualisée avec le sanger. Un allèle avec un seul triplet AAT comme chez le contrôle aurait été observé. Si jamais, les 2 allèles avaient été exprimés, une superposition de pics à partir de la duplication aurait été observée. Dans ce cas, il aurait été impossible de déterminer la position des mutations sur les allèles et l'accès au matériel biologique des parents aurait été indispensable.

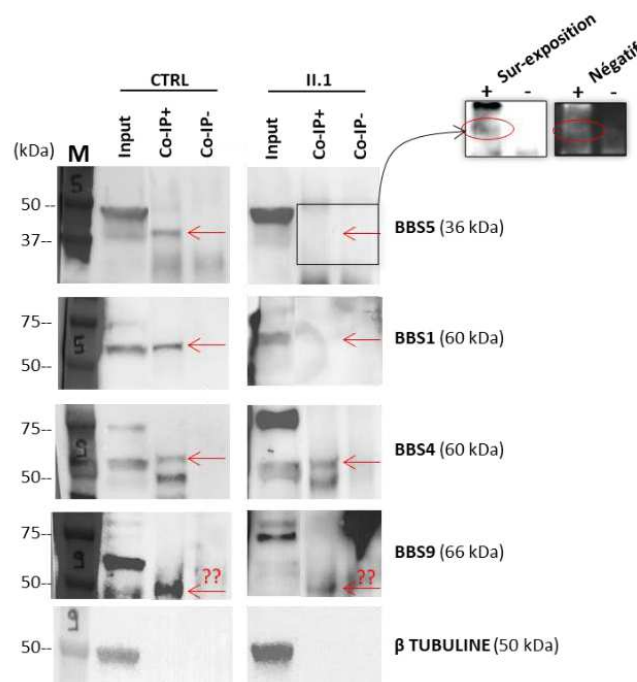
### B) Etude de l'interaction entre les sous-unités du BBSome

L'absence d'une des 8 sous unités du BBSome ou la perte d'interaction entre celles-ci peut déstabiliser l'ensemble du complexe. Nous voulions contrôler que le BBSome s'assemble correctement et qu'il est toujours présent chez la patiente. Il a été montré très récemment que BBS5 et BBS9 sont en interaction directe au sein du BBSome (Klink et al. 2020b). Nous avons donc envisagé la possibilité que la duplication impacte l'interaction entre ses 2 sous unités. Nous avons donc recherché de potentielles pertes d'interaction entre BBS5 et les autres sous unités du BBSome par Co-IP. Nous avons d'abord immunoprécipité la protéine BBS5 puis réalisé un WB avec des AC dirigés contre BBS9, BBS1 et BBS4 (Figure 51).



Pour le contrôle, un marquage de BBS5 est présent sur la membrane. Il s'agit d'un témoin positif qui montre que la protéine appât (BBS5 en l'occurrence) a été immunoprécipitée par l'anticorps. Le marquage des autres sous unités du BBSome telles que BBS1 et BBS4 confirme les interactions au niveau de ce complexe : elles ont été co-immunoprécipitées avec BBS5. Par contre, lors du marquage de BBS9, j'ai obtenu une bande à 50 kDa alors que la taille attendue est de 66 kDa. Cela peut être dû à la présence de plusieurs isoformes de BBS9 ou à une mauvaise qualité de l'AC.

Pour la patiente, le marquage de BBS5 pose problème. En effet, avec une exposition normale aucune bande correspondant à BBS5 n'est visible. Par contre, en surexposant la membrane et en mettant le négatif de l'image, une bande de BBS5 semble apparaître faiblement. Pour expliquer ce résultat, nous avons plusieurs hypothèses. Soit, l'expérience n'a pas marché (BBS5 n'a pas été immunoprécipitée). Soit, l'expérience a fonctionné mais la quantité de BBS5 immunoprécipitée est très faible car la quantité de BBS5 chez la patiente est très réduite (Figure 51B) et en conséquence, les protéines co-immunoprécipitées sont également en faible quantité. Dans tous les cas, les résultats ne sont pas concluants, il faudra refaire cette expérience en augmentant la quantité de protéines totales utilisée pour la Co-IP à 500µg par exemple (contre 300µg initialement). Il faudra faire une Co-IP en utilisant comme protéine appât, une sous unité du BBSome dont la quantité est normale chez la patiente comme chez les contrôles : BBS1 par exemple.



**Figure 51 : Les interactions entre BBS5 et les autres BBS.** Résultats des Co-IP anti-BBS5 (protéine appât) réalisées chez un contrôle et chez la patiente. Résultats du WB après marquage des protéines BBS5, BBS1, BBS4, BBS9 : 4 des 8 sous-unités du BBSome. M : marqueur, II.1 : la patiente, CTRL : contrôle

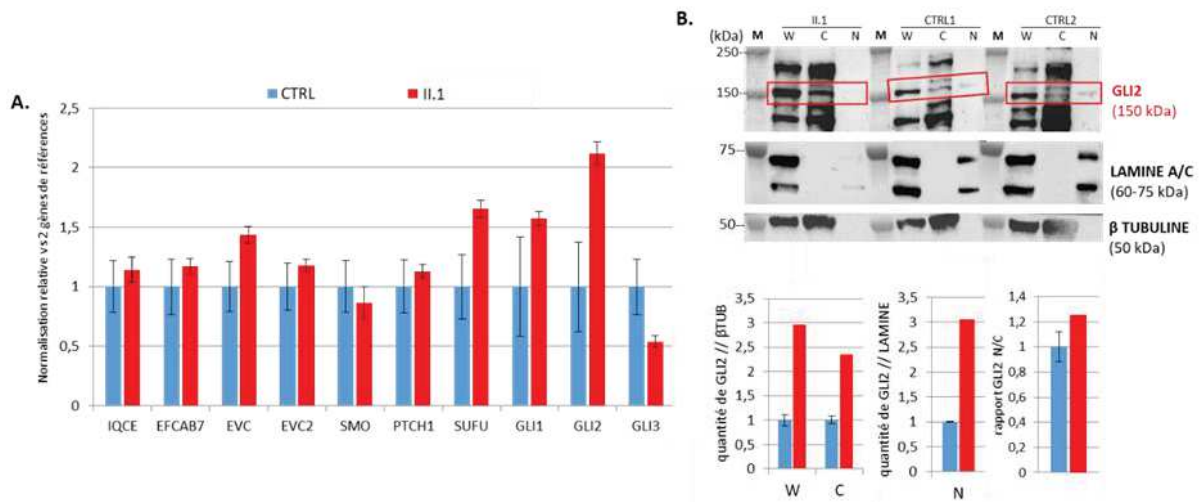
C) L'étude du fonctionnement de la voie SHH

Des mutations dans certains gènes BBS peuvent perturber l'activation de voies de signalisations majeures telle que la voie SHH. Nous avons étudié l'impact de la duplication sur l'activation de cette voie en cultivant les cellules dans un milieu déprivé en sérum (-SVF) mais en présence de SAG (agoniste de la voie SHH, SMO agonist) pendant 24h. A l'issue du traitement, nous avons réalisé une qRTPCR pour interroger l'expression de gènes dont les produits protéiques sont impliqués dans la cascade de signalisation de cette voie. Nous avons procédé à une extraction différentielle des protéines du noyau et celles du cytoplasme. L'activation de la voie SHH entraîne la translocation de GLI2 depuis le cytoplasme vers le noyau où ce facteur de transcription active l'expression de gènes cibles. Nous avons réalisé un WB pour détecter la présence de la protéine GLI2 dans chacun des compartiments. Une présence majoritaire de GLI2 dans le noyau indiquera qu'il a effectivement été transloqué dans ce dernier comme attendu quand la voie SHH est activée avec le SAG. En revanche, si la quantité nucléaire de GLI2 est moindre chez la patiente, nous pourrions conclure que la duplication entraîne une sous-activation de la voie SHH.

En interrogeant la voie SHH, nous apporterons une explication au phénotype typique de BBS de la patiente en mettant en évidence une perturbation au sein d'une des voies majeures du développement. L'analyse de la quantification relative aux deux gènes de références *GAPDH* et *HPRT* montre chez la patiente une surexpression de *GLI1*, *SUFU* et *GLI2* et une sous-expression de *GLI3*. Le niveau d'expression des autres gènes interrogés ne semble pas varier par rapport aux contrôles (Figure 52A). Ces résultats suggèrent, pour ces conditions de culture, une suractivation de la voie SHH chez la patiente par rapport aux contrôles. La quantité protéique de GLI2 dans le cytoplasme et le noyau a été déterminée par WB par rapport à la  $\beta$ -TUBULINE et à la LAMINE A/C (protéines de références cytoplasmique et nucléaire respectivement). Les résultats montrent que les quantités de GLI2 totales (W), cytoplasmiques (C) et nucléaires (N) chez la patiente sont supérieures par rapport à celles des contrôles. En plus, le rapport GLI2 nucléaire sur GLI2 cytoplasmique chez la patiente est supérieur à 1 : la quantité de GLI2 nucléaire est supérieure à celle de GLI2 cytoplasmique (Figure 52B). Ces résultats vont dans le sens d'une translocation nucléaire de GLI2 plus importante chez la patiente suggérant ainsi une suractivation de la voie SHH, dans ces conditions de culture. Cette conclusion concorde avec celle issue des données de qRTPCR.

L'expérience devra être refaite et un travail de mise au point sera à prévoir pour minimiser les bandes aspécifiques présente sur la membrane. En recourant au SAG, nous avons activé la voie SHH : GLI2 doit être présent dans toutes les fractions. La bande GLI2 considérée est celle à 150 kDa (PM attendu) qui apparait dans la fraction nucléaire. Il faudra déposer plusieurs extractions différentielles à partir de plusieurs culots cellulaires de la patiente (triplicats biologiques) pour confirmer que les résultats

obtenus sont fiables et reproductibles. Il faudra refaire les 2 expériences avec des culots de cellules cultivées en conditions -SVF seul afin de déterminer si la suractivation mise en évidence est également présente en conditions basales chez la patiente.



**Figure 52 : Résultats des expériences menées sur la voie SHH.** A. Résultats de la qRTPCR réalisée pour déterminer le niveau d'expression de 10 gènes codants pour des protéines impliquées dans la régulation de la voie chez la patiente (triplicats techniques) par rapport à des contrôles (triplicats biologiques). Quantification relative réalisée par rapport aux 2 gènes de références *GAPDH* et *HPRT*. B. Résultats du WB après le marquage de la protéine GLI2 et de la protéine de référence cytoplasmique ( $\beta$ -TUBULINE) et nucléaire (LAMINE A/C). Résultats des quantifications de GLI2 par rapport à la  $\beta$ -TUBULINE, à la LAMINE A/C et le rapport de GLI2 N/C, chez II.1 en normalisant par rapport aux contrôles. M : marqueur, II.1 : la patiente, CTRL : contrôle, W : whole /protéines totales, C : cytoplasme, N : noyau

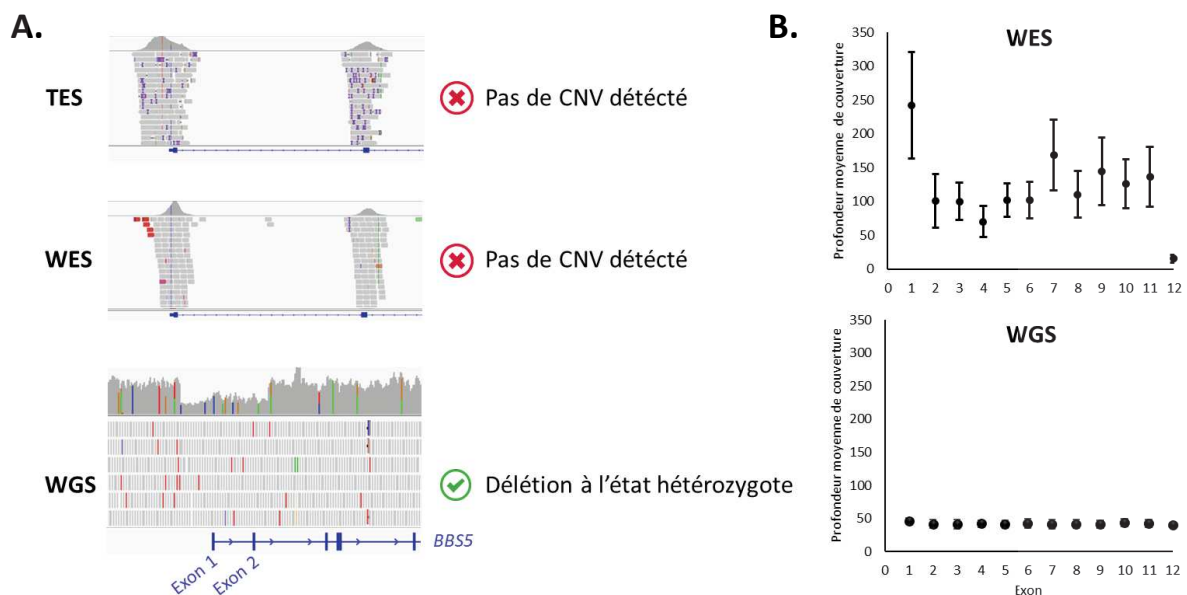
### III. La discussion

Les technologies de séquençage à haut débit ont largement contribué à réduire les difficultés rencontrées lors du diagnostic moléculaire des patients atteints de BBS. La mise en place de panel de gènes en premier intention, permet le séquençage simultané de tous les gènes BBS incluant parfois d'autres gènes responsables de phénotypes proches. Le séquençage de l'exome (WES) et maintenant du génome (WGS) permettent d'étendre cette recherche d'anomalies moléculaires et de réduire l'errance diagnostic pour de nombreuses familles. Des études ont déjà montré que certaines mutations peuvent être détectées par des technologies et manquées par d'autres (panel vs WES vs WGS) (Geoffroy et al. 2018). C'est le cas de la délétion dans ce projet, détectée par WGS seulement.

Les mutations dans le gène *BBS5* sont rares et ne correspondent qu'à 2% des cas de BBS (Imani et al. 2019). Jusqu'à présent, plusieurs mutations ont été rapportées dans ce gène comme responsables des BBS et sont référencées dans le Tableau 8. La patiente rapportée dans cette étude est hétérozygote composite : elle est porteuse d'une délétion de l'exon 1 et 2 du gène *BBS5* situé *en trans* d'une duplication d'un triplet AAT (c.550\_552dupAAT) dans l'exon 6 de *BBS5* conduisant à la duplication de

l'asparagine 184 (p.N184dup). Les deux variations n'ont jamais été rapportées auparavant et ne sont pas présentes dans les bases de données.

La duplication avait été observée dès le début par panel puis par WES et enfin par WGS. Cependant il s'agit d'une duplication d'un acide aminé, une mutation dont la pathogénicité est difficile à démontrer contrairement à d'autres mutations (les non-sens, par exemple). De plus, avec le panel et le WES aucune autre mutation dans le gène BBS5 avait été mise en évidence. BBS étant une pathologie à transmission autosomique récessive la présence de mutations sur les 2 allèles du gène sont requises. Un séquençage du génome complet (WGS) a été réalisé dans ce but de détecter une mutation intronique profonde ou une variation structurale (CNV). La délétion a alors été identifiée. Le fait qu'elle n'ait pas été détectée par les 2 premières techniques peut être expliqué par la richesse en GC de la zone génomique où sont situés les exons 1 et 2 de BBS5 : des régions plus difficilement capturées et séquencées. Ce problème a été résolu par le WGS car cette technique assure une couverture plus homogène des fragments d'ADN (absence de capture) comme le résume la Figure 53.



**Figure 53 : Détection des mutations en fonction de la technologie NGS utilisée.** A. Comparaison des impressions d'écran IGV pour les 3 techniques. B. Graphique représentant la profondeur moyenne de couverture pour chaque exon de BBS5. CNV : Variation du nombre de copie.

La délétion englobe le promoteur du gène, nous avons confirmé l'hypothèse initiale qu'il n'y aurait pas de transcription de cet allèle ni de protéine traduite. Les données de qRT-PCR et de WB montrent une réduction de moitié de la quantité d'ARN et une forte réduction de la quantité de protéine BBS5. Pour être sûre que les 50% d'ARN exprimés ne sont pas divisés en une proportion de 25% de chaque allèle, nous avons vérifié sur le génome que le gène BBS5 ne présente aucun promoteur alternatif qui aurait entraîné l'expression d'un allèle porteur de la délétion. Nous avons observé la mutation N184dup à l'état homozygote sur le profil Sanger de l'ADNc ce qui confirme la co-ségrégation familiale. Les

mutations sont situées sur 2 allèles différents, chacun hérité d'un des parents. Ces 2 résultats nous permettent finalement de conclure que l'unique allèle exprimé de *BBS5* est celui porteur de la duplication N184.

Pour aller encore plus loin, nous pourrions recourir à une qRT-PCR avec la technologie Taqman afin de discriminer le niveau d'expression de chacun des 2 allèles individuellement. Une quantité protéique normale de *BBS5* aurait pu être observée malgré l'expression d'un seul des deux allèles allant dans le sens d'une probable compensation au niveau protéique. Mais, le résultat du WB montre une forte réduction de *BBS5* et prouve qu'il n'y a pas eu de compensation. Chez la patiente, la protéine *BBS5* résiduelle est mutée. Pour confirmer que la bande très légèrement marquée chez la patiente n'est pas un artéfact, il faudrait refaire un WB en déposant des quantités croissantes de protéines. Si l'intensité du marquage et la quantité de protéine déposée varient de manière proportionnelle alors nous pourrions conclure que la bande observée correspond effectivement à *BBS5*. Alternativement, s'il n'y a pas de protéine *BBS5*, cela confirme l'effet délétère de la duplication N184. L'effet probable de la duplication serait une déstabilisation la structure 3D de la protéine qui conduirait à la dégradation de la protéine.

Dans notre cas, la seule protéine *BBS5* synthétisée est porteuse de la duplication N184 au niveau de son 2<sup>ème</sup> domaine DM16. Il s'agit d'un domaine de fonction encore inconnue mais pour lequel la séquence est très conservée au cours de l'évolution (J. B. Li et al. 2004). Ce haut niveau de conservation laisse donc penser qu'une mutation dans ce domaine serait pathogène.

La protéine *BBS5* est l'une des huit sous unités composant le BBSome, localisé à la base du cil primaire. La mutation a potentiellement créé un changement dans la structure/le repliement de la protéine *BBS5* impactant alors potentiellement sa conformation, sa localisation, sa stabilité et les interactions qu'elle établit (Klink et al. 2020b; S. K. Singh et al. 2020b). Nous avons donc pensé déterminer la localisation de la protéine *BBS5* mutée au niveau du cil par IF. Le problème majeur que nous avons rencontré est inerrant à l'anticorps utilisé. En effet, il existe plusieurs études publiées ayant déjà essayé de marquer la protéine *BBS5*. Cependant, certaines ont rapporté un marquage à la base du cil (Kazatskaya et al. 2017) et d'autres, aucun marquage (Eguether et al. 2014b) sur des cellules contrôles ciliées. Toutes les équipes ont utilisé la même référence d'anticorps de chez Proteintech (14569-1AP). Il est possible que ces différences de marquages soient dues à des conditions de fixations ou de perméabilisations différentes ou encore à la qualité intrinsèque de l'AC. Nous allons devoir faire des mises au point pour optimiser le protocole. L'article de Eguether *et al.* 2014 montre qu'un marquage de *BBS5* est obtenu pour des cellules mutées dans *IFT27*. En effet, l'absence d'une protéine *IFT27* fonctionnelle empêche le recyclage de *BBS5* hors du cil : la protéine *BBS5* s'accumule et permet l'apparition d'un marquage ciliaire de *BBS5* par IF. Il se trouve qu'au laboratoire, nous avons à disposition de telles cellules. Je comptais donc les utiliser, en condition ciliée, pour mettre au point l'IF. Une autre possibilité pour

localiser la protéine BBS5 normale ou mutée dans les cellules serait d'exprimer la protéine avec un « tag » fluorescent de type GFP, par exemple et de confronter les localisations observées.

Pour déterminer l'impact de la mutation sur l'assemblage du BBSome, nous avons recherché une possible perte d'interaction entre BBS5 et les autres protéines du BBSome en réalisant une Co-IP dirigée contre BBS5. Les résultats se sont avérés concluants chez un contrôle car les résultats ont confirmé une interaction entre BBS5 et les protéines BBS1 et BBS4. Le marquage de BBS9 a posé problème : une bande à 50 kDa a été révélée alors que la taille attendue est de 66 kDa. Cette bande peut être expliquée par le fait que BBS9 présente plusieurs isoformes. De toute manière, ce marquage n'est pas fiable et il faudrait recourir à un autre AC pour pouvoir conclure.

Chez la patiente, les résultats de la Co-IP posent problème : un léger marquage de la bande BBS5 est apparu suite à une surexposition de la membrane. Etant donné le faible marquage de BBS5 déjà observé chez la patiente lors des WB précédents, il est difficile de conclure. Il est fort probable que la quantité de la protéine appât ayant été immunoprécipitée avec l'anticorps soit très faible. Cette expérience devra être refaite. Il serait judicieux d'utiliser une quantité de protéines à immunoprécipiter supérieure aux 300 µg initiaux.

Pour comprendre si la duplication affecte l'assemblage général du BBSome et pas seulement les interactions que BBS5 établit avec les autres sous-unités, nous avons prévu de faire une deuxième Co-IP avec un AC anti-BBS1 suivi d'un WB avec des AC anti-BBS4 et anti-BBS5. Faute de temps et de matériel biologique, cette expérience n'a pas pu être réalisée. Mais, elle nous permettra de déterminer si oui ou non BBS1 interagit avec BBS5 chez la patiente, sans qu'il y ait de problème de quantité protéique de la protéine appât. En effet, pour cette Co-IP, la quantité de BBS1 est normalement équivalente chez la patiente et chez les contrôles. Seule la quantité de BBS5 est diminuée de moitié chez la patiente. La présence d'une bande pour BBS5, chez la patiente montrerait que l'interaction entre les 2 sous-unités est maintenue. En revanche, la disparition de la bande suggérerait une perte d'interaction entre BBS1 et BBS5 et donc un impact de la duplication sur l'assemblage global du BBSome.

La corrélation génotype-phénotype permettant de relier le phénotype de BBS à des mutations dans le gène *BBS5* peut être établie à plusieurs niveaux. Au niveau cellulaire, les gènes responsables de ciliopathies sont impliqués dans le fonctionnement ciliaire. En cas de mutations, des changements dans la structure des cils ou le mauvais adressage de molécules impliquées dans les voies de signalisation sont mis en évidence. BBS5 est impliqué dans la régulation de la voie SHH (Imani et al. 2019), nous avons donc cherché à déterminer l'impact de la mutation sur la régulation de cette voie. Pour ces expériences, nous avons étudié des cellules cultivées dans des conditions -SVF+SAG : la voie SHH est activée pour tous les individus. Nous cherchons à mettre en évidence une suractivation ou une sous-activation chez la patiente par rapport aux contrôles. Du coup, des gènes codants pour des protéines

impliquées dans la cascade de signalisation ont été interrogés par qRT-PCR. Une surexpression de *GLI1* et *GLI2* (GLI activateurs) et une sous-expression de *GLI3* (GLI répresseur) ont été observées. Ces résultats suggèrent une possible suractivation de la voie SHH. Une surexpression du régulateur négatif *SUFU* a également été mise en évidence. Cette variation peut être expliquée par le fait que la cellule cherche à garder l'homéostasie en augmentant l'expression de *SUFU* afin d'avoir un rétrocontrôle négatif et permettre un retour à un niveau normal d'activation.

Au niveau protéique, nous avons appliqué un protocole d'extraction différentielle basée celui de Estrada-Cuzcano et al. 2020. Finalement, par cette technique, nous avons obtenu des enrichissements de compartiments qui sédimentent ensemble. Par exemple, la fraction dite « nucléaire » correspond en fait à un enrichissement des protéines nucléaires mais qui contient possiblement des portions de membrane du RE. Afin de réaliser une extraction nucléaire pure, il faudrait procéder à une séparation de l'enrichissement d'abord par gradient de sucrose puis à l'aide d'une colonne d'affinité par exemple. Toutefois, l'analyse des résultats obtenus montrent une augmentation de la quantité de *GLI2* dans les fractions « nucléaires » (N) et cytoplasmiques (C) chez la patiente. Le rapport *GLI2* N/C est supérieur à 1 chez la patiente : *GLI2* subirait une translocation nucléaire en quantité plus importante par rapport aux contrôles. Ces résultats de WB confortent ceux de la qRT-PCR suggérant, pour la condition de culture -SVF+SAG, une possible suractivation de la voie SHH, même s'il s'agit de résultats préliminaires. Il faudra refaire cette expérience en procédant à l'optimisation du protocole et en effectuant des triplicats pour la patiente. Il faudrait également réaliser des conditions contrôles sur les cellules de la patiente et des mêmes contrôles en conditions -SVF seul et +SVF+SAG afin de déterminer le profil correspondant dans ces conditions basales. Des études de la signalisation SHH dans un certain nombre d'organismes ont montré que SHH est un « messenger » sécrété par les cellules pouvant agir sur les cellules environnantes pour réguler leur prolifération et influencer leur différenciation cellulaire. La famille des protéines GLI (*GLI1*, 2, 3) peuvent, par exemple, réguler le processus de développement des membres (Tickle et Towers 2017). Par exemple, les maladies congénitales humaines associées à *GLI3* incluent toutes des malformations des membres (Hill, Götz, et Rütger 2009). L'hyperactivation de la voie SHH a été associée à la polydactylie dans plusieurs études (Xu et al. 2020). Il s'agit effectivement d'un signe clinique typique du BBS présent chez la patiente.



## IV. La conclusion et les perspectives

Nous rapportons deux nouvelles mutations au niveau du gène *BBS5*. La première est une délétion englobant la région promotrice de l'allèle et conduisant à une perte d'expression de celui-ci (variant pathogène : classe 4). La seconde correspond à une duplication de l'Asp184. Nous avons montré que la quantité de protéine *BBS5* de la patiente est fortement réduite. Malheureusement, face au manque de fiabilité des résultats de la Co-IP anti-*BBS5* chez la patiente, nous ne sommes pas en mesure de conclure totalement sur l'impact de cette duplication. Pour l'instant, sa pathogénicité doit être confirmée, le variant ne peut pas être reclassé ni en variation bénigne (classe 2) ni en variation pathogène (classe 4). La dérégulation de la voie SHH mise en évidence pour la patiente ne peut pas être reliée à cette mutation. Cette donnée ne fait que confirmer le lien, déjà connu, entre une hyperactivation de la voie SHH et le BBS. Nous voudrions pouvoir établir le lien entre la possible suractivation mise en évidence et la duplication identifiée dans *BBS5*.

Pour mener à bien ce projet et répondre aux objectifs, certaines expériences devront être refaites et optimisées. D'autres devront être mises en œuvre. Afin de confirmer que le phénotype BBS observé chez la patiente est effectivement dû aux mutations du gène *BBS5*, un travail complémentaire sera à réaliser. Par exemple, une modélisation bioinformatique basée sur la structure atomique récemment publiée du BBSome en ajoutant l'Asp184 dupliquée à la protéine *BBS5* pourrait aider à mieux comprendre les effets de cette mutation sur la structure 3D de la protéine et les interactions établies au sein du BBSome. Au niveau cellulaire, une expérience de rescue/sauvetage pourrait être mise en place : il s'agirait de réexprimer la protéine normale *BBS5* dans les cellules de la patiente et d'observer s'il y a une restauration du phénotype cellulaire. Une transfection avec un plasmide porteur de la version mutée de *BBS5* peut être réalisée sur des cellules contrôles dont l'expression du gène *BBS5* normal a été abolie. Si le phénotype observé pour les cellules transfectées correspond à celui observé pour les cellules de la patiente, nous pourrions conclure que les mutations causent le phénotype. Nous aurons alors la preuve de pathogénicité et nous pourrions établir une corrélation génotype-phénotype. Nous pourrions faire exprimer des protéines *BBS5* avec un « tag » de GFP pour localiser les protéines sans recourir aux anticorps du marché.

L'objectif ultime est de pouvoir revenir vers la patiente avec un diagnostic moléculaire robuste sur les causes de sa maladie. Au regard de la communauté scientifique, une fois notre travail de validation terminé, ces mutations seront ajoutées aux bases de données de variations telle que ClinVar afin de pouvoir partager notre analyse et nos conclusions. Ces informations pourront être utilisées afin de mieux comprendre le fonctionnement de *BBS5* et du BBSome.

## V. Les données supplémentaires.

Nomenclature de la mutation au niveau de l'ADNc	Nomenclature de la mutation au niveau de la protéine	Référence
c.1A>T	p.M1?	Muller (2010) Hum Genet 127, 583
c.2T>A	p.M1?	Harville (2010) J Med Genet 47, 262
c.149T>G	p.L50R	Chen (2011) Invest Ophthalmol Vis Sci 52, 5317
c.148C>A	p.L50I	Wang (2014) Hum Genet 133, 331
c.158C>T	p.T53I	Abu-Safieh (2012) Eur J Hum Genet 20, 420
c.166A>G	p.R56G	Muller (2010) Hum Genet 127, 583
c.214G>A	p.G72S	Hjortshoj (2008) Am J Med Genet A 146A, 517
c.265C>T	p.R89*	Hirano (2015) PLoS One 10, e0136317
c.412C>T	p.R138C	Carss (2017) Am J Hum Genet 100, 75
c.413G>A	p.R138H	Redin (2012) J Med Genet 49, 502
c.425T>G	p.L142*	Chandrasekar (2018) Indian J Med Res 147, 177
c.532G>A	p.G178R	Shaheen (2016) Genome Biol 17, 242
c.547A>G	p.T183A	Hjortshoj (2008) Am J Med Genet A 146A, 517
c.551A>G	p.N184S	Li (2004) Cell 117, 541
c.584A>G	p.D195G	Lindstrand (2016) Am J Hum Genet 99, 318
c.620G>A	p.R207H	Li (2004) Cell 117, 541
c.790G>A	p.G264R	Weisschuh (2016) PLoS One 11, e0145951
c.889G>A	p.D297N	Feuillan (2011) J Clin Endocrinol Metab 96, E528
c.208+2T>C		Imani (2019) Biosci Rep epub, epub
c.522+3A>G		Li (2004) Cell 117, 541
c.619-27T>G		Hirano (2015) PLoS One 10, e0136317
c.619-1G>C		Feuillan (2011) J Clin Endocrinol Metab 96, E528
c.900G>C		Carss (2017) Am J Hum Genet 100, 75
c.123delA	p.(Gly42Glufs*11)	Smaoui (2006) Invest Ophthalmol Vis Sci 47, 3487
c.734_744del11	p.(Glu245Glyfs*18)	Maria (2016) Sci Rep 6, 34764
c.955_957delGAA	p.(Glu319del)	
c.54dupC	p.(Ala19Argfs*14)	Feuillan (2011) J Clin Endocrinol Metab 96, E528
c.966dupT	p.(Ala323Cysfs*57)	
c.263_271delTACGAGGCCinsGCTCTTA	p.(Leu88Cysfs*3)	
ex. 6-7, c.387-?_618+?del232		Deveault (2011) Hum Mutat 32, 610
ex. 9-12		Nishimura (2005) Am J Hum Genet 77, 1021
incl. ex. 8-12		

**Tableau 8** : Listes des mutations de *BBS5* répertoriées.

# IV.

Partie 2 :

La caractérisation

phénotypique du nouveau

modèle murin *Vps15*<sup>R998Q</sup>

## I. Le synopsis

### 1. Remise dans le contexte :

La mutation faux-sens c.2993G>A, p.R998Q du gène *VPS15* (*PIK3R4*) a été identifiée par le laboratoire comme responsable d'une ciliopathie pour une fratrie avec 3 patients présentant principalement une dégénérescence rétinienne précoce et une atteinte rénale sévère (Stoetzel et al. 2016). Avant cette publication, la protéine VPS15 était seulement connue comme une sous unité régulatrice de la Phosphatidyl-Inositol 3 kinase (VPS34) impliquée dans le processus d'autophagie et dans le transport vésiculaire. En 2016, le laboratoire a rapporté un nouveau rôle de VPS15 : son implication dans le fonctionnement du cil primaire. A l'époque, les validations fonctionnelles de la mutation ont été réalisées en recourant aux fibroblastes des patients, aux levures humanisées et aux poisson-zèbres *Vps15*<sup>MO</sup>. Pour parfaire la compréhension des rôles de *Vps15*, un modèle murin Knock-In (KI) de *Vps15* porteur de la même mutation que les patients (*Vps15*<sup>R998Q/R998Q</sup>) a été généré au laboratoire en 2017. La partie 2 de ce manuscrit porte sur la caractérisation phénotypique de ce nouveau modèle au cours de mes 3 années de thèse. En générant un tel modèle, nous avons envisagé la possibilité d'obtenir un nouveau modèle de ciliopathie. Il s'agissait (1) de déterminer les potentielles anomalies présentes dans les différents organes des souris mutées et (2) de pouvoir étudier la protéine *Vps15* mutée dans un contexte pathologique grâce à ce modèle *in vivo*. En 2017, quand j'ai débuté ma thèse, le seul modèle murin de *Vps15* disponible était la souris KO conditionnelle publiée par Nemazanyy et al. 2013 (le modèle KO total avait été rapporté létal par la même équipe). Ce n'est qu'un an plus tard que la publication de Gstrein et al. en 2018 a décrit un nouveau modèle murin hypomorphe pour *Vps15* avec des anomalies cérébrales.

### 2. La démarche scientifique

Une première vague d'expérience avait pour but de déterminer si ce modèle murin était un bon modèle d'étude de ciliopathie. Pour cela, nous avons d'abord porté un intérêt majeur aux organes principalement atteints chez les patients : la rétine et le rein. Dans ce cas, nous avons mené des études fonctionnelles, histologiques et moléculaires. Nous avons mis en culture les fibroblastes des souris afin d'étudier le phénotype à une échelle plus réduite. Puis, une première analyse du cerveau des souris (biométrie) a été réalisée en 2018 par l'équipe de Binnaz Yalcin.

Nous avons approfondi nos travaux sur le modèle KI *Vps15* à la recherche de troubles cognitifs et de marqueurs histologiques classiques de neurodégénérescence pour plusieurs raisons : (1) *Vps15* est une protéine impliquée dans le processus d'autophagie et plusieurs pathologies neurodégénératives comme la maladie d'Alzheimer ont déjà été associées à des perturbations de cette voie. (2) Les modèles murins de *Vps15* rapportés par Gstrein *et al*, 2018 montraient des atteintes cérébrales. La deuxième vague d'expériences avait pour but d'explorer plus précisément le phénotype cérébral des

souris afin de déterminer si les principaux critères de la maladie d'Alzheimer (plaques amyloïdes, mort cellulaire, hyperphosphorylation de la protéine Tau, troubles cognitifs) sont présents chez les souris *Vps15*.

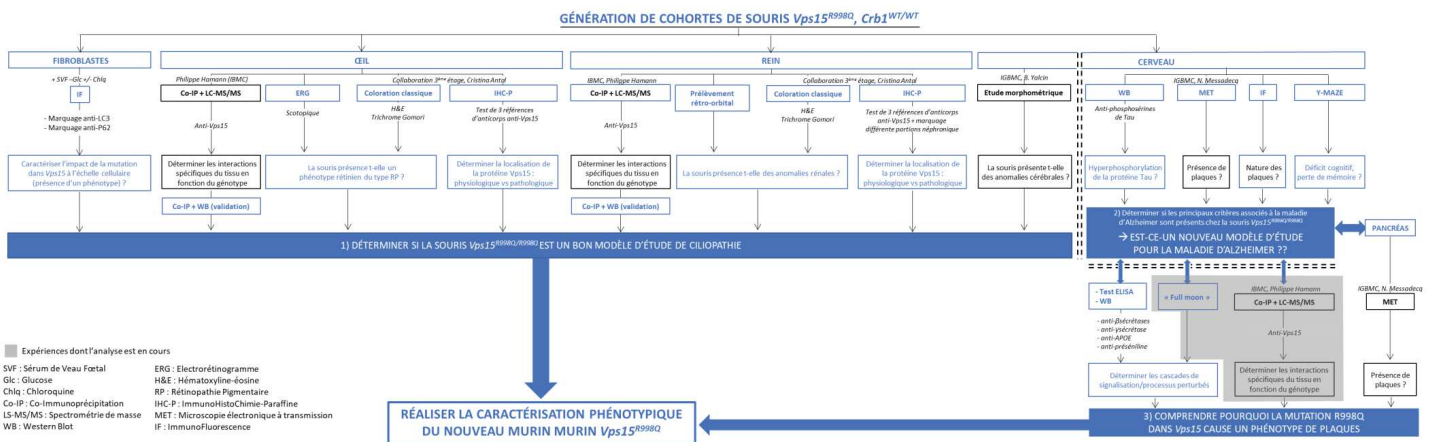
Finalement, une 3<sup>ème</sup> et dernière vague d'expériences a été menée pour relier la mutation R998Q dans *Vps15* au phénotype cérébral observé. La maladie d'Alzheimer fait partie des amyloïdoses : des maladies caractérisées par la formation de plaques dans de multiples organes. Ainsi, nous avons étendu nos recherches à d'autres organes et plus particulièrement au pancréas.

Le logigramme (Figure 54) résume les expériences mises en place pour aboutir à l'objectif principal de ce projet : réaliser la caractérisation phénotypique de ce modèle murin *Vps15*.

### 3. Le changement de fond génétique

Les souris ont été générées en 2017 grâce au financement de la Fondation maladie rare (FRM), par l'ICS, dans le cadre du projet « Phenomin ». Initialement, cette lignée a été générée sur un fond génétique C57BL6N, par la technique de « point de mutation » et l'utilisation du système « Cre-lox » (voir matériel et méthodes). La génération de la lignée sur le fond génétique C57BL6J avait échoué.

Mattapallil et al. 2012 a rapporté que le fond génétique C57BL6N est porteur de la mutation Rd8 (c.3481delC) dans le gène *Crb1*, une protéine impliquée dans l'établissement de jonctions serrées entre les cellules de Muller et les photorécepteurs de la rétine. Cette mutation provoque un phénotype rétinien avec la présence de pseudo-rosettes. Ce phénotype est en revanche absent chez les animaux de fond génétique C57BL6J. Etant donné le potentiel phénotype rétinien attendu chez les souris *Vps15*, nous avons génotypé les animaux et comme attendu, ils étaient tous porteurs de la mutation Rd8 à l'état homozygote. Nous avons donc décidé très rapidement, dès le début du projet, de passer les animaux sur un fond génétique C57BL6J afin de ne pas biaiser les possibles résultats. Nous avons réalisé une série de 8 backcross en croisant les souris *Vps15*<sup>R998Q/WT</sup> avec des animaux WT de fond génétique C57BL6J et confirmé par génotypage que tous les animaux étudiés par la suite sont indemnes de Rd8 (*Crb1*<sup>WT/WT</sup>).



**Figure 54 :** Logigramme. Résumé des 3 vagues d'expériences mises en œuvre pour répondre aux différents objectifs du projet. Mon implication est indiquée en bleu.

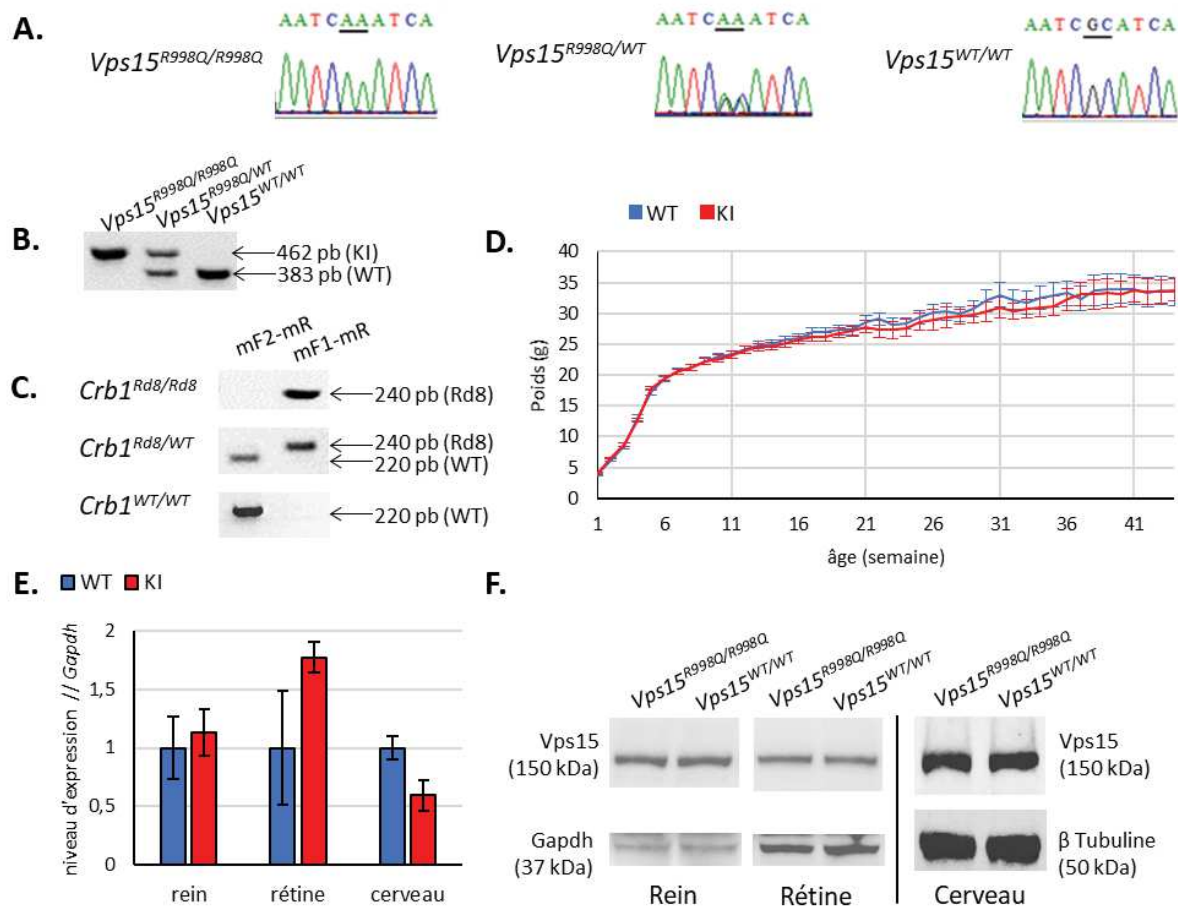
## II. Les résultats

### 1. Présentation de la lignée

Les souris *Vps15*<sup>R998Q/R998Q</sup> sont porteuses de la mutation c.3430\_3431delinsAA ; p.R998Q dans le gène *Vps15*, une mutation confirmée par séquençage Sanger sur l'ADNg extrait des bouts de doigts des souris (couple Mf-Mr) (Figure 55A). Nous avons également confirmé l'expression de l'allèle porteur de la mutation dans les 3 organes d'intérêt (rétine, rein et cerveau) par séquençage Sanger de l'ADNc. Le génotypage de routine des animaux utilise le couple d'amorce Ef<sup>2</sup>-Er<sup>2</sup> (Figure 55B). La bande supérieure de 462 pb correspond à l'amplification d'un fragment porteur du site Lox-P associé à la mutation R998Q : bande unique amplifiée chez un animal *Vps15*<sup>R998Q/R998Q</sup> (Figure 55A). En revanche, un fragment de 383 pb correspond au fragment WT sans insertion du Lox-P : bande unique amplifiée chez un animal *Vps15*<sup>WT/WT</sup>. La présence des 2 bandes est retrouvée chez un animal hétérozygote (*Vps15*<sup>R998Q/WT</sup>). Le génotypage de la mutation Rd8 du gène *Crb1* suit le protocole de (Mattapallil et al. 2012) (Figure 55D) pour lequel un fragment unique de 240 pb obtenu avec le couple mF1-mR est porteur de la délétion : animal *Crb1*<sup>Rd8/Rd8</sup>. Un fragment unique de 220 pb obtenu avec le couple mF2-mR est WT : animal *Crb1*<sup>WT/WT</sup>. Un animal hétérozygote *Crb1*<sup>Rd8/WT</sup> présente les 2 fragments. Après changement de fond génétique, nous avons toujours comparé des souris *Vps15*<sup>R998Q/R998Q</sup> ; *Crb1*<sup>WT/WT</sup> à leurs congénères *Vps15*<sup>WT/WT</sup> ; *Crb1*<sup>WT/WT</sup>. Tous les animaux, quel que soit le sexe, sont fertiles. Les proportions de Mendel sont respectées à l'issue des croisements entre hétérozygotes et le ratio femelles/mâles est équilibré à 50:50. La courbe de poids générée ne montre pas de différence de poids significative en fonction du génotype (Figure 55D).

Pour les 3 tissus étudiés dans ce projet, nous avons confirmé par qRT-PCR que le niveau d'expression de l'ARNm de *Vps15* porteur de la mutation R998Q n'est pas significativement différent de l'ARNm WT (Figure 55E). De la même manière, au niveau protéique, par WB, nous n'avons pas observé de différence ni dans la quantité de *Vps15* (par rapport aux protéines de référence : Gapdh et  $\beta$ Tubuline) ni dans la taille de la protéine (150 kDa) (Figure 55F). Ces résultats sont cohérents avec ceux obtenus à partir des fibroblastes des patients (Stoetzel et al. 2016).





**Figure 55 : Caractérisation de base de la lignée.** **A.** Confirmation par séquençage Sanger de la présence de la mutation dans *Vps15* pour des souris homozygotes, hétérozygotes ou WT. **B.** Résultats du génotypage par PCR de routine pour les 3 statuts par rapport à la mutation dans *Vps15*. **C.** Résultats du génotypage par PCR pour les 3 statuts par rapport à la mutation Rd8 dans *Crb1*. **D.** Courbe représentant l'évolution du poids des souris en fonction de l'âge (n=41, sexes confondus). **E.** Niveau d'expression de *Vps15* en fonction du génotype dans les 3 tissus d'intérêt (rein, rétine et cerveau). **F.** Résultats de WB du marquage de la protéine *Vps15* en fonction du génotype dans les 3 tissus et des protéines de référence Gapdh ou βTubuline.

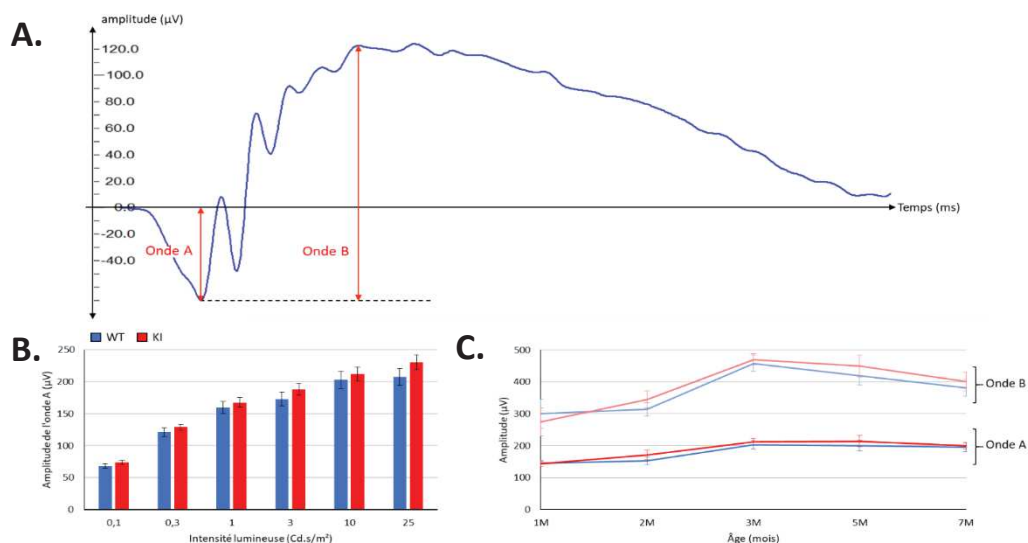
## 2. Analyse portant sur les yeux :

Les patients porteurs de la mutation R998Q dans le gène *Vps15* présentent tous une atteinte visuelle sévère ayant mené à la cécité précoce. En effet, leur électrorétinogramme (ERG) s'est avéré plat dès l'âge de 10 ans. Pour déterminer si un phénotype similaire se développait chez les souris, nous avons réalisé un test fonctionnel de type ERG. Nous avons également réalisé une étude histologique en collaboration avec Cristina Antal (Institut d'histologie, Faculté de médecine, Strasbourg). Des expériences de Co-IP suivies de spectrométrie de masse, réalisées en collaboration avec Philippe Hammann (plateforme de protéomique, IBMC, Strasbourg) ont permis de déterminer les interactants de *Vps15* dans la rétine.

A) Etude fonctionnelle par ERG

Afin de déterminer si les souris *Vps15*<sup>R998Q/R998Q</sup> présentent des anomalies visuelles, nous avons réalisé une série d'ERG en condition scotopique sur les souris (n=7, tous sexes confondus). Nous n'avons pas fait de distinction de sexes car il n'existe pas de dysmorphisme sexuel au niveau des yeux des souris. L'ERG est un test fonctionnel permettant de déterminer la capacité de réponses des cellules de la rétine (par enregistrement de courant électrique) à des stimuli lumineux d'intensités croissantes. La condition scotopique nécessite une adaptation au noir pendant 12h avant le test. Elle a pour objectif de déterminer la capacité de réponse de l'ensemble des photorécepteurs (bâtonnets et cônes). La condition photopique ne nécessite pas d'adaptation préalable au noir (contrairement à la condition scotopique) et permet de tester la capacité de réponse des cônes uniquement puisque les bâtonnets sont saturés dans ces conditions. Nous avons réalisé une étude longitudinale sur la même cohorte de souris à des âges croissants (1, 2, 3, 5 et 7 mois) pour mettre en évidence une potentielle dégénérescence progressive et déterminer à quel âge elle apparaît.

Sur un tracé d'ERG, l'onde A correspond à l'hyperpolarisation des photorécepteurs et l'onde B correspond à la dépolarisation des cellules ganglionnaires, en réponse aux flashes lumineux (Figure 56A). Nous n'avons pas observé de différence d'amplitude au niveau de l'onde A, en fonction du génotype des souris, même à 7M, l'âge étudié le plus avancé (Figure 56B). Globalement, les amplitudes des 2 ondes ne sont pas significativement différentes en fonction du génotype (Figure 56C). Les souris *Vps15*<sup>R998Q/R998Q</sup> ne présentent pas d'atteintes rétiniennes. Nous n'avons pas observé d'aplatissements progressifs des ondes comme c'est le cas pour des modèles murins BBS (par exemple *BBS10*<sup>-/-</sup> ou *Cep290*<sup>-/-</sup> dans Brun et al. 2019).

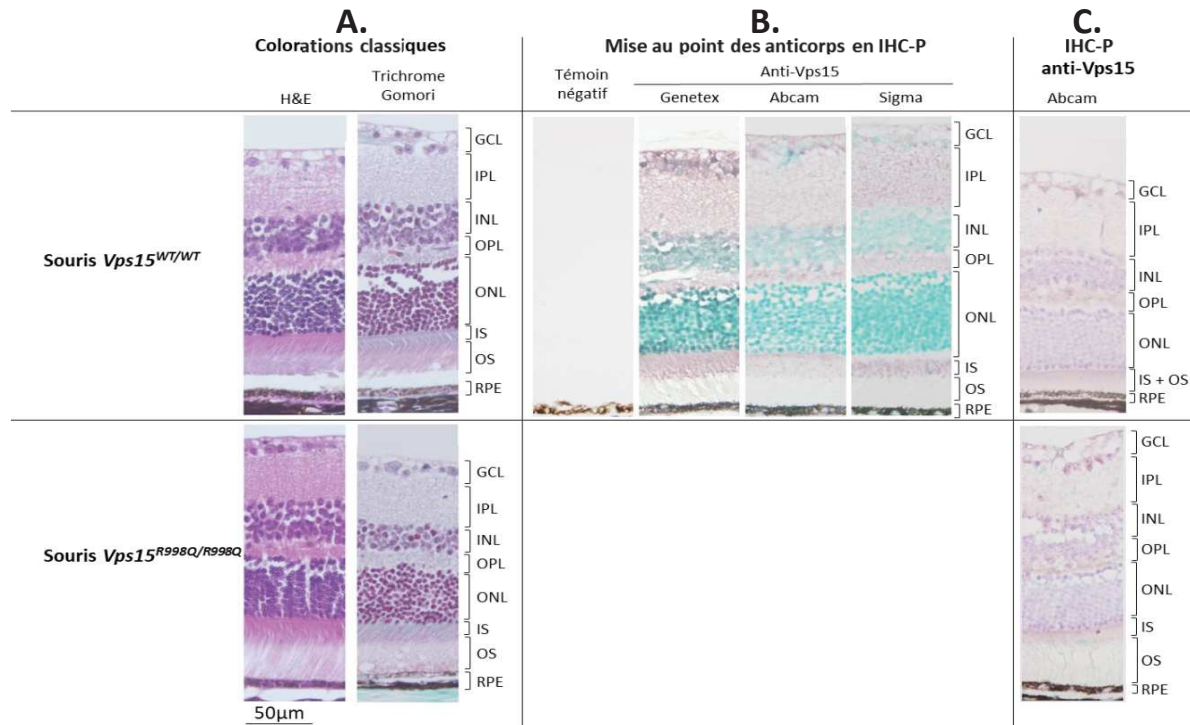


**Figure 56 : Résultats d'ERG.** A. Tracé typique obtenu à l'issue d'un ERG en condition scotopique : annotations des ondes A et B. B. Histogramme représentant les amplitudes de l'onde A en fonction de l'intensité du flash lumineux dans le cadre de l'ERG à 7 mois (n=7, sexes confondus). C. Représentation des amplitudes moyennes des ondes A et B en fonction de l'âge des souris.

B) Etude histologique des yeux

Nous avons d'abord mené une étude histologique sur les yeux des souris de 5 mois (n=6 pour chacun des sexes) afin de mettre en évidence une anomalie d'organisation du tissu. Nous avons appliqué des colorations classiques de type H&E pour laquelle l'hématoxyline : un colorant dit cationique ou basique marque les structures dites anioniques ou acides en violet tels que les noyaux par exemple. A l'inverse, l'éosine est un colorant anionique qui marque les structures cationiques en rose tel que le cytoplasme (Figure 57A). Le Trichrome de Gomori est une coloration qui permet la visualisation des fibres de collagène en bleu et des noyaux en violet. La fibrose peut par exemple être détectée par cette coloration. Dans les 2 cas, nous n'avons pas observé de différences en fonction du génotype : l'histologie rétinienne des souris *Vps15*<sup>R998Q/R998Q</sup> est similaire à celle des souris *Vps15*<sup>WT/WT</sup>.

Nous avons ensuite cherché à déterminer le patron d'expression de *Vps15* dans la rétine par Immunohistochimie paraffine (IHC-P) en utilisant plusieurs références d'anticorps dirigés contre cette protéine (Abcam, Sigma et Genetex). L'intensité du marquage est proportionnelle au niveau d'expression de la protéine. Comme attendu, à partir de résultats d'HIS (hybridation in situ) réalisée contre l'ARNm de *Vps15* sur des rétines de souris adultes (Blackshaw et al. 2004), *Vps15* est exprimé dans l'ensemble des couches de la rétine. Il n'y a pas de différence en fonction du génotype (Figure 57C).



**Figure 57 : Etude histologique de la rétine.** A. Colorations classiques H&E (hématoxyline et éosine) et trichrome de Gomori pour des coupes d'yeux de souris de 5M (n=6/sexes/génotypes). B. Mise au point des 3 références d'anticorps anti-Vps15 sur les coupes des yeux de souris WT. C. Résultats des IHC-P réalisées avec l'anticorps anti-Vps15 (Abcam). GCL : couche des cellules ganglionnaires, IPL : couche plexiforme interne, INL : couche nucléaire interne, OPL : couche plexiforme externe, ONL : couche nucléaire externe, IS : segment interne/OS : segment externe des photorécepteurs, RPE ; épithélium pigmentaire rétinien.

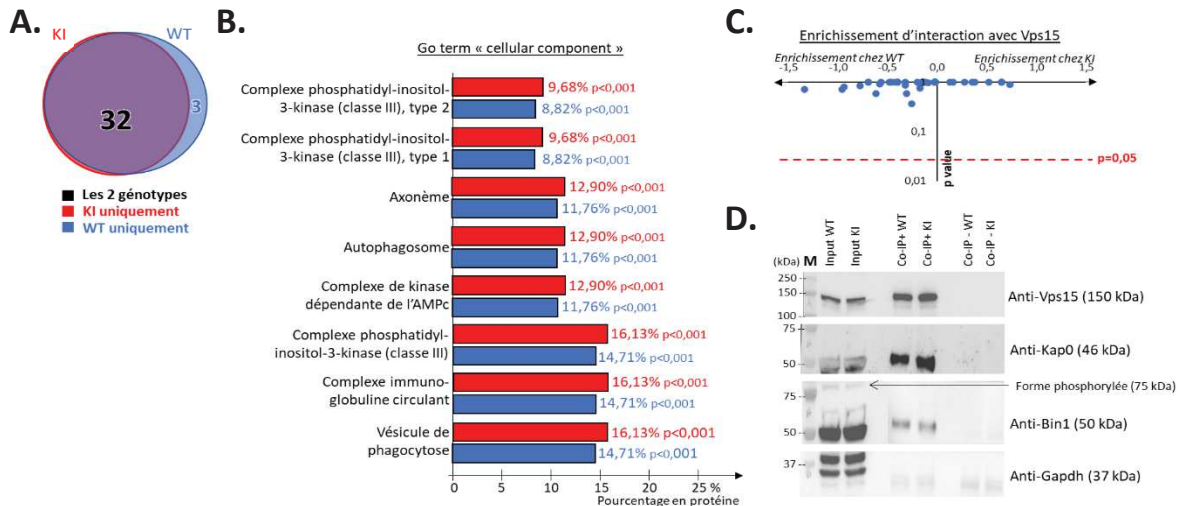
C) Etude de l'interactome de Vps15

Afin de déterminer, d'une part, si la mutation R998Q impacte les interactions établies entre Vps15 et d'autres protéines et afin de mettre en évidence, d'autre part, l'interactome dans la rétine des souris *Vps15*, nous avons réalisé une co-immunoprécipitation anti-Vps15 suivi d'une analyse par spectrométrie de masse (n=3).

Pour analyser le tableau de résultats (Figure 58), j'ai appliqué des filtres pour sélectionner uniquement les protéines dont le nombre de spectres identifiés était  $\geq 2$ , absent ou peu présent ( $\leq 1$ ) dans les Co-ip négatives. Ainsi, 32 protéines ont été sélectionnées dans les rétines KI : elles sont toutes présentes chez les WT. Seulement 3 protéines sont en interaction avec Vps15 uniquement dans la rétine WT (Figure 59A). Les analyses permettant de mettre en évidence les quantités de protéine en interaction avec Vps15 dans un des géotypes par rapport à l'autre n'ont donné aucun résultat significatif (Figure 59C). Par contre, les analyses d'enrichissement de voies de signalisation cellulaire réalisées avec les listes de 32 protéines pour les KI et 35 protéines pour les WT montrent des enrichissements dans les voies associées au transport vésiculaire et au complexe PIK3, nécessaire au processus d'autophagie. Un enrichissement d'environ 12% a été mis en évidence dans la voie d'activation dépendante de l'AMPc (PKA). En effet, les sous-unités catalytiques (KapcA et KapcB), régulatrices (Kap0 et Kap1) ainsi que la protéine d'ancrage à la membrane (Shkap) ont été identifiées en interaction avec Vps15. De plus, la protéine Bin1 (amphiphysine-2) a également été identifiée comme un nouvel interactant de Vps15 dans la rétine. Les interactions de Vps15 et Kap0 et de Vps15 et Bin1 ont été validées par Co-IP suivie d'un Western blot (Figure 59D).

accession	BASIC Spectral Count (#spectra)												Average #spectra				RATIOS				Gene name
	CTRL-WT			CTRL-KO			VPS15-IP-WT			VPS15-IP-KO			CTRL		VPS15		VPS15/CTRL		VPS15		
	1	2	3	1	2	3	1	2	3	1	2	3	WT	KO	WT	KO	WT	KO		KO/WT	
Pi3R4_MOUSE							145	184	162	146	133	138			163.67	139.00	VPS15-WT	VPS15-KO	0.85	Pik3r4	
PK3C3_MOUSE							102	120	109	106	104	107			110.33	105.67	VPS15-WT	VPS15-KO	0.96	Pik3c3	
UVRAG_MOUSE							51	70	63	55	48	48			61.33	50.33	VPS15-WT	VPS15-KO	0.82	Uvrag	
BECN1_MOUSE							58	60	57	53	56	49			58.33	52.67	VPS15-WT	VPS15-KO	0.90	Becn1	
BAKOR_MOUSE							17	25	20	20	20	15			20.67	18.33	VPS15-WT	VPS15-KO	0.89	Atg14	
AMPH_MOUSE							16	23	22	22	17	22			20.33	20.33	VPS15-WT	VPS15-KO	1.00	Amph	
SPKAP_MOUSE							10	24	19	14	11	10			17.67	11.67	VPS15-WT	VPS15-KO	0.66	Sphkap	
NRF2_MOUSE							13	9	15	9	11	10			12.33	10.00	VPS15-WT	VPS15-KO	0.81	Nrf2	
HVM58_MOUSE							5	6	7	4	7	7			6.00	6.00	VPS15-WT	VPS15-KO	1.00		
KAPO_MOUSE							3	9	4	6	4	5			5.33	5.00	VPS15-WT	VPS15-KO	0.94	Prkar1a	
HVM53_MOUSE							3	3	5	2	4	4			3.67	3.33	VPS15-WT	VPS15-KO	0.91		
HVM63_MOUSE							3	3	5	2	4	4			3.67	3.33	VPS15-WT	VPS15-KO	0.91		
KCNT1_MOUSE							2	4	5	3	3	2			3.67	2.67	VPS15-WT	VPS15-KO	0.73	Kcnt1	
CUED2_MOUSE							1	4	4	3	4	5			3.00	4.00	VPS15-WT	VPS15-KO	1.33	Cuedc2	
KV5AB_MOUSE							3	1	4	1	2				2.67	1.00	VPS15-WT	VPS15-KO	0.38		
KAP1_MOUSE							7	7	1	2	2	2			2.67	1.33	VPS15-WT	VPS15-KO	0.50	Prkar1b	
HVM06_MOUSE							2	1	4	1	3	1			2.33	1.67	VPS15-WT	VPS15-KO	0.71		
SAHH_MOUSE							2	2	2	2	2	1			2.00	1.00	VPS15-WT	VPS15-KO	0.50	Ahcy	
KV2A7_MOUSE							3	1	1	1	2	2			1.67	1.67	VPS15-WT	VPS15-KO	1.00		
SARM1_MOUSE							2	2	3	2	1				1.67	1.00	VPS15-WT	VPS15-KO	0.60	Sarm1	
VDAC3_MOUSE							2	1	1	2	1	2			1.33	1.67	VPS15-WT	VPS15-KO	1.25	Vdac3	
ALBU_MOUSE							1	1	2	1	2	1			1.33	1.33	VPS15-WT	VPS15-KO	1.00	Alb	
HXK2_MOUSE							1	1	1	1	1	1			1.00	0.67	VPS15-WT	VPS15-KO	0.67	Hk2	
CNR6_MOUSE							1	1	1	1	1				1.00	0.67	VPS15-WT	VPS15-KO	0.67	Pde6g	
GBRR1_MOUSE							1	2	1	1	1	1			1.00	0.67	VPS15-WT	VPS15-KO	0.67	Gabr1	
CHM4B_MOUSE							1	1	1	1	1	1			0.67	1.00	VPS15-WT	VPS15-KO	1.50	Chmp4b	
BIN1_MOUSE				1			9	17	10	12	14	14	0.33	12.00	13.33	40.00	1.11	Bin1			
RP9_MOUSE					1	2	2	1	1	3	2	2	1.00	1.33	2.33	2.33	1.75	rp9			
RLA1_MOUSE					1		2	1	1	2	1	1	0.33	1.33	1.00	3.00	0.75	Rplp1			
SERF2_MOUSE					1		1	3		2	1		0.33	1.33	1.00	3.00	0.75	Serf2			
KAPCA_MOUSE							3	4	5	4	4	1	0.33	4.00	3.00	12.00	0.75	Prkaca			
KAPCB_MOUSE							2	2	5	2	3	1	0.33	3.00	2.00	9.00	0.67	Prkacb			
KV5A3_MOUSE							3	2	3	2	3	3	0.67	2.67	2.67	4.00	1.00				
RS12_MOUSE							2	2	1	1	1	2	0.33	1.67	1.33	5.00	0.80	Rps12			
CLP1_MOUSE							1		2	1	1	1	0.33	1.00	1.00	3.00	1.00	Clp1			

**Figure 58 :** Résultats de spectrométrie de masse pour la rétine. Tableau de résultats récapitulant les interactants connus de *Vps15* et les nouveaux interactants potentiels spécifiques de la rétine.



**Figure 59 :** Analyses d'enrichissements et validations d'interactions pour la rétine. **A.** Diagramme de Venn représentant le nombre de protéines identifiées par spectrométrie de masse partagées par les 2 génotypes. **B.** Histogramme montrant l'enrichissement protéique mis en évidence dans les protéines identifiées chez les WT et les KI. **C.** Volcano-plot représentant la quantité de chacune des protéines en interaction avec *Vps15* en fonction du génotype (KI vs WT). **D.** Résultats de la Co-IP anti-*Vps15* suivi du WB pour la validation des interactions.

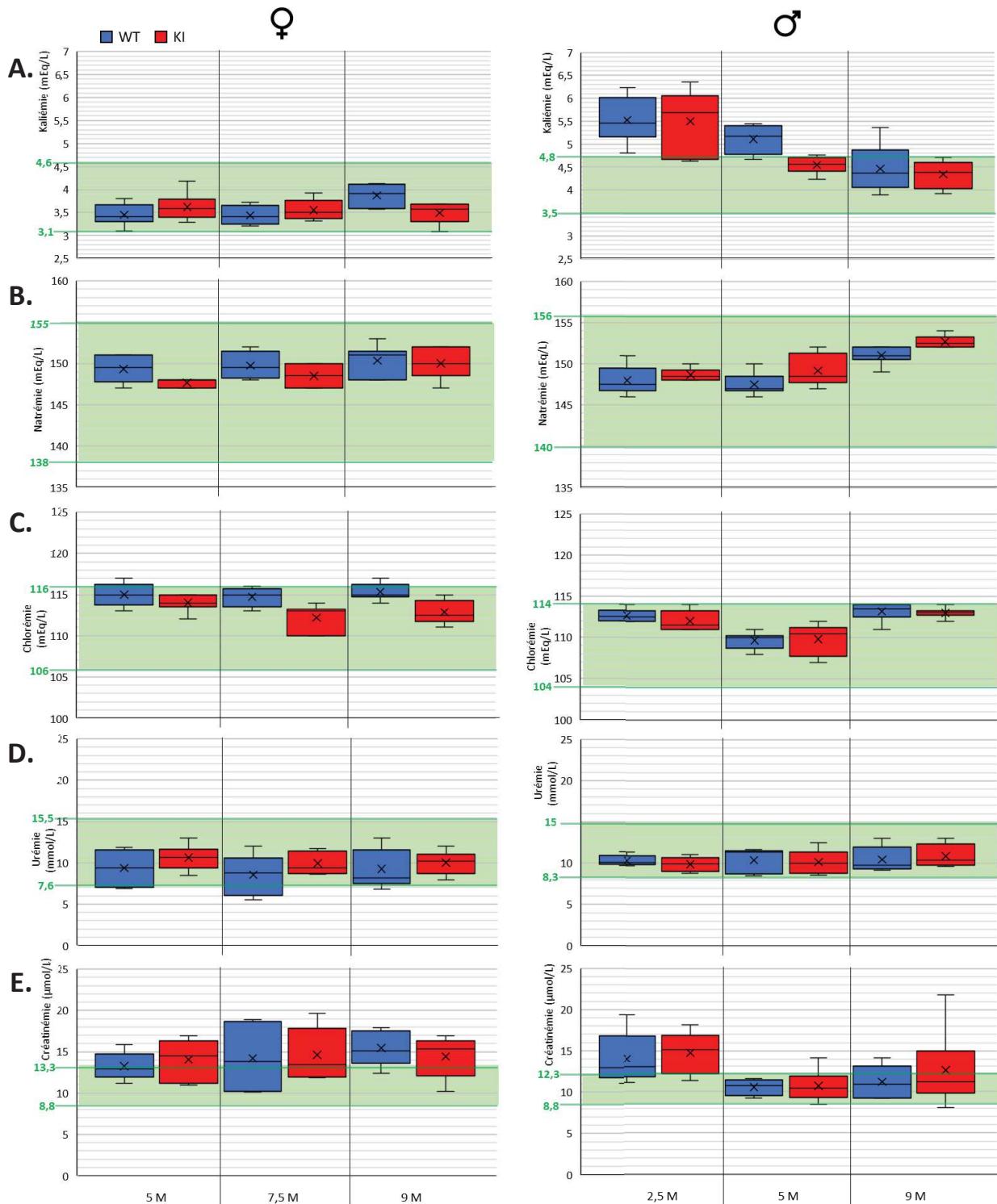


## 2. Analyse portant sur le rein :

### A) Prélèvement rétro-orbital et analyses biochimiques

Afin de déterminer si les souris présentent une anomalie du fonctionnement rénal, nous avons réalisé des analyses biochimiques du plasma des souris (n=6, sexes séparés) à des âges croissants. En effet, des concentrations sanguines anormales de substances éliminées dans l'urine (urée, créatinine, Na<sup>+</sup>, Cl<sup>-</sup>, K<sup>+</sup>) sont la preuve d'un mauvais fonctionnement rénal. Par exemple, la filtration du sang par les glomérules permet la formation d'une urine primaire dont la composition est normalisée. Les différentes zones du néphron sont le lieu de réabsorption/sécrétion d'ions et d'autres substances. Ainsi, les variations hors des normes connues pour les substances dosées devraient nous aider à déterminer la zone du néphron qui dysfonctionne.

Globalement, les valeurs de kaliémie, natrémie, chlorémie et d'urémie des souris (n=6 par sexe) sont toujours comprises dans les zones de référence (Figure 60A-D). A 9 mois, il ne semble pas y avoir de différences significatives en fonction du génotype. Pour les valeurs de créatinémie, les femelles WT et KI sont hors des valeurs normales mais aucune différence dépendante du génotype n'a été mise en évidence (Figure 60E). Pour les mâles KI, les valeurs sont très hétérogènes avec des valeurs très hautes pour certains animaux. Il faudrait augmenter l'effectif d'animaux et éventuellement faire les analyses biochimiques au-delà de 9 mois pour voir l'évolution des valeurs.

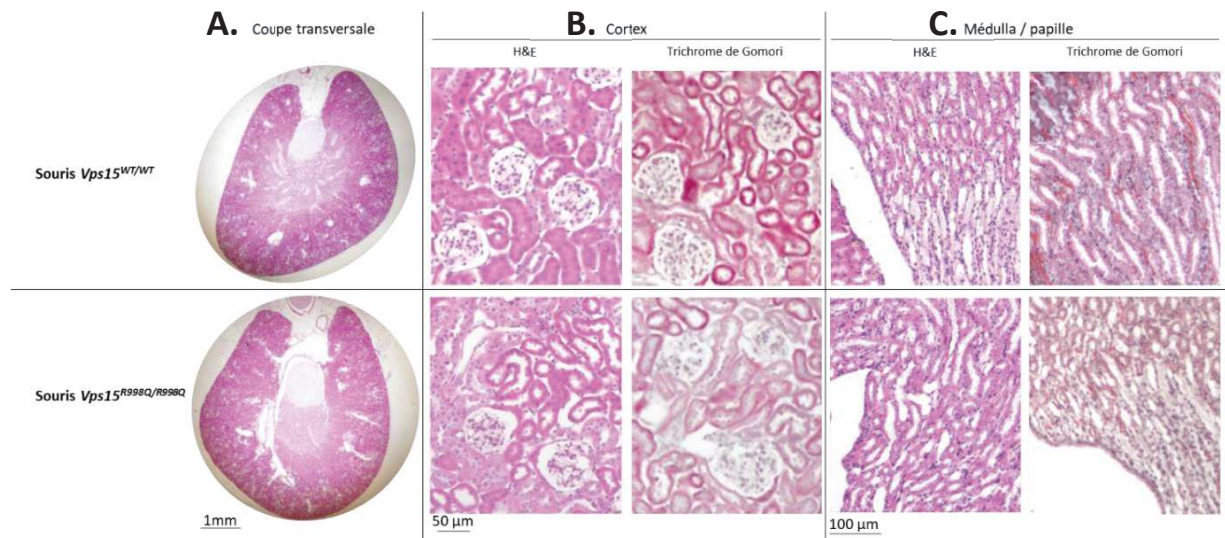


**Figure 60 : Analyses biochimiques plasmatiques.** Représentation, en fonction de l'âge, des concentrations plasmatiques des substances dosées : A. K+, B. Na+, C. Cl-, D. l'urée, E. la créatinine. Les valeurs de références tirées de Otto et al. 2016 et Mazzaccara et al. 2008 sont délimitées en vert. Les résultats pour les souris femelles (n=6) sont à gauche et ceux pour les mâles (n=6) à droite.



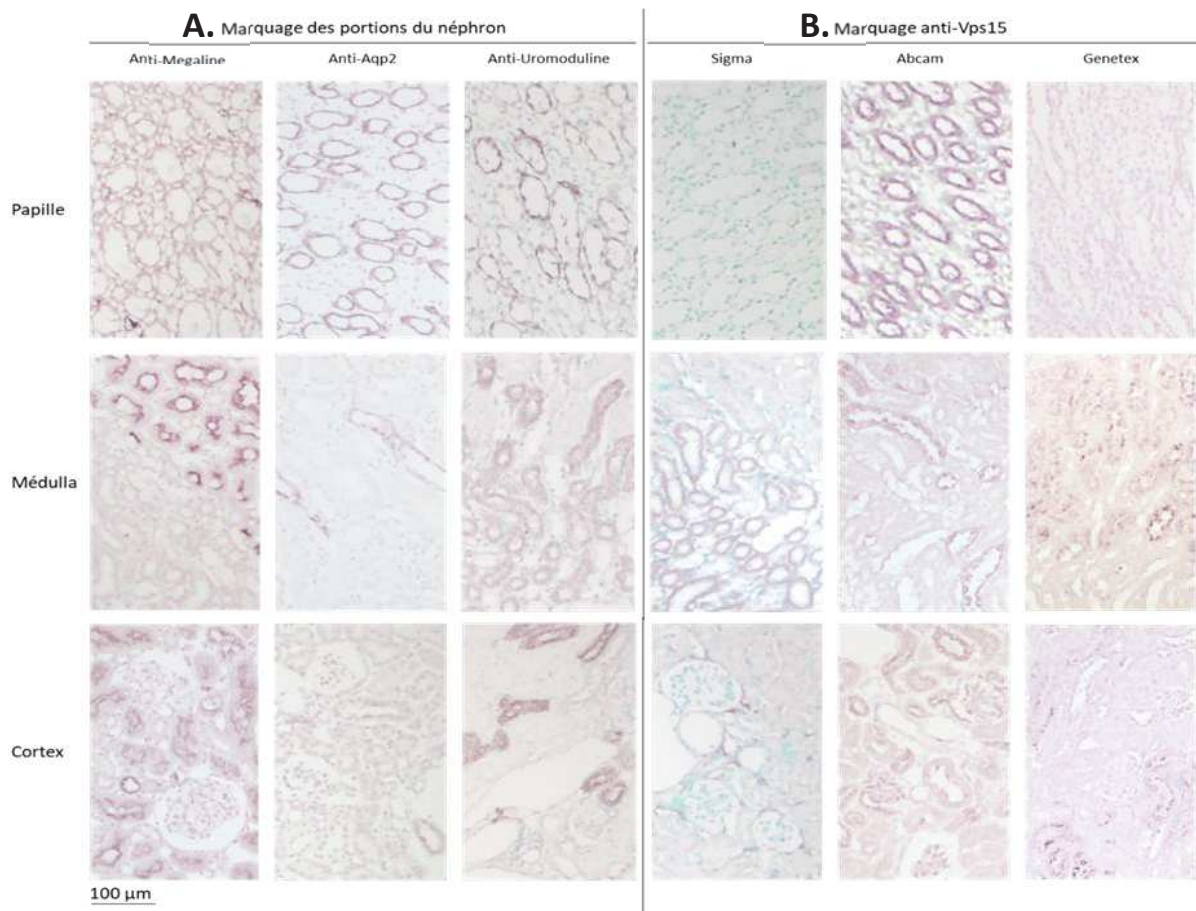
B) Etude histologique du rein

Nous avons mené une étude histologique sur les reins des souris de 5M (n=6 par sexe) afin de mettre en évidence une anomalie d'organisation du tissu. Comme pour l'œil, nous avons appliqué les mêmes colorations (H&E et trichrome de Gomori) (Figure 61B et C). Nous n'avons pas observé de différences en fonction du génotype : aucun kyste n'a été mis en évidence sur les coupes transversales (Figure 61). Aucune fibrose n'a été révélée ni dans le cortex, ni dans la médulla (Figure 61B et C). L'histologie rénale des souris *Vps15*<sup>R998Q/R998Q</sup> est similaire à celle des souris *Vps15*<sup>WT/WT</sup>.



**Figure 61 : Etude histologique du rein.** A. Photo des coupes transversales entières des reins des souris (n=6 /sexe/génotype) colorées par H&E. Photos B. du cortex et C. de la médulla pour les 2 colorations appliquées.

Afin de déterminer les différentes zones néphroniques, nous avons utilisé des anticorps reconnaissant des protéines marqueurs de chacune d'elles (Figure 62A). Un marquage de la mégaline/Lrp2 permet de distinguer les tubules proximaux. Un anticorps anti-Aqp2 révèle les tubes collecteurs. En marquant l'uromoduline/THP, les tubes contournés et les branches fines de l'anse de Henlé sont mis en évidence. Nous avons ensuite cherché à déterminer le patron d'expression de *Vps15* dans le rein par IHC-P en utilisant plusieurs anticorps (Abcam, Sigma et Genetex). Ils ont été appliqués sur des coupes sériées de reins (coupes obtenues à la suite les unes des autres). Les correspondances de marquages entre les différentes références ont été comparées pour déterminer les zones du rein où *Vps15* est exprimée. Sigma marque les tubes contournés. Abcam marque l'urothélium et les tubes collecteurs au niveau de la papille et du cortex et les podocytes du glomérule. Genetex marque légèrement les tubes collecteurs dans la papille et les tubes contournés dans le cortex. A la vue de ces observations, il est impossible de conclure car nous avons obtenu 3 patrons différents avec les 3 anticorps anti-*Vps15* (Figure 62B).



**Figure 62 :** IHC-P sur les coupes rénales de souris WT. **A.** Photos des différentes portions rénales avec des marquages anti-mégaline, anti-Aqp2 et anti-uromoduline. **B.** Photos des différentes portions rénales avec les marquages anti-Vps15 issus des 3 références d'anticorps utilisées.

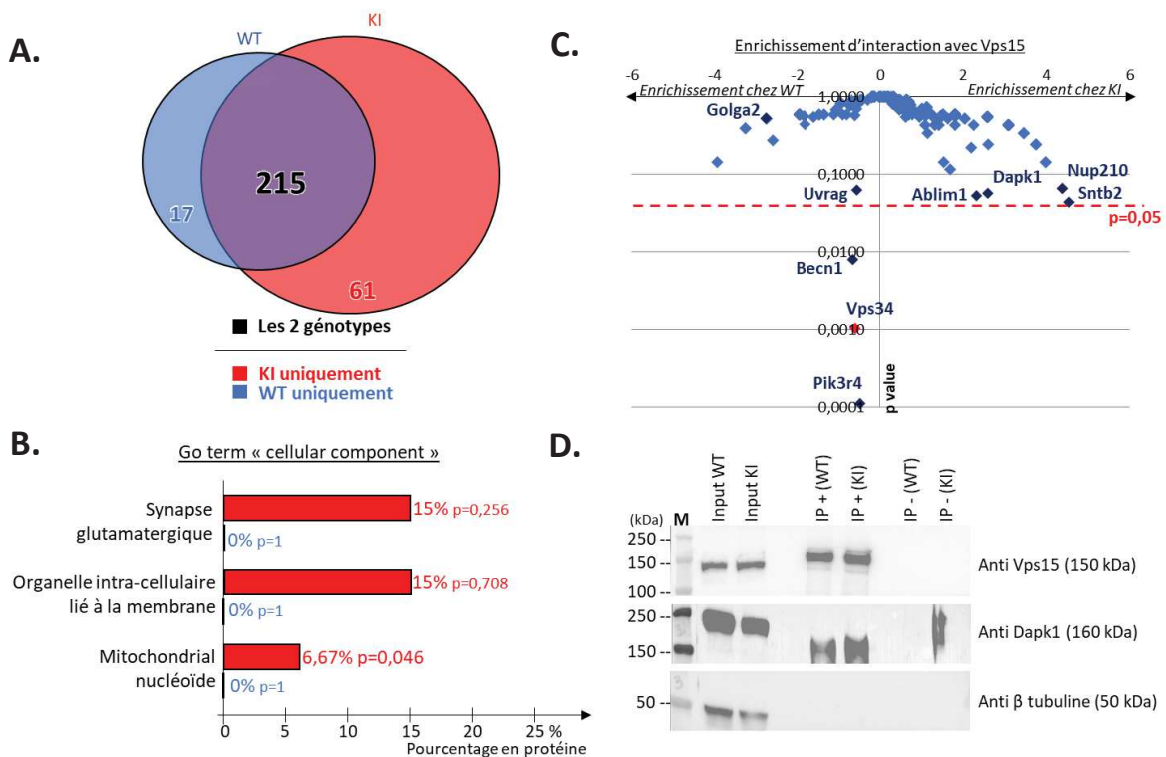
C) Etude de l'interactome de Vps15

Afin de déterminer si la mutation R998Q impacte les interactions établies entre Vps15 et d'autres protéines et de mettre en évidence l'interactome, cette fois dans le rein des souris *Vps15*, nous avons réalisé une Co-IP anti-Vps15 suivi d'une analyse par spectrométrie de masse (n=3).

Pour analyser le tableau de résultats (Figure 63), j'ai appliqué des filtres pour sélectionner uniquement les protéines dont le nombre de spectres identifiés était  $\geq 2$ , absent ou peu présent ( $\leq 1$ ) dans les co-ip négatives. Ainsi, 215 protéines sont communes entre les reins WT et KI. 17 protéines sont en interaction avec Vps15 uniquement dans les reins des souris WT. 61 protéines sont spécifiquement retrouvées en interaction avec Vps15 chez les souris KI (Figure 64A) et des analyses d'enrichissements des voies de signalisations cellulaires ont été réalisées sur cette liste (Figure 64B). Les analyses permettant de mettre en évidence une quantité de protéine en interaction avec Vps15 dans un des génotype par rapport à l'autre ont mis en évidence certaines protéines dont Dapk1 : sa quantité protéique en interaction avec Vps15 semble plus importante dans le rein KI (Figure 64C). La validation de ces résultats a été réalisée par Co-IP suivie d'un Western blot : l'interaction entre Vps15 et Dapk1 a été validée (Figure 64D). Il semble même que la bande correspondante à Dapk1 soit de taille plus importante chez le KI par rapport au WT après Co-IP alors que la grosseur des bandes est inversée dans l'input. Golga2 (= Gm130) a été identifiée parmi les protéines pour lesquelles la quantité en interaction avec Vps15 est appauvrie dans les reins KI (Figure 64C). Il s'agit d'une protéine qui avait été identifiée dans les co-ip réalisées sur les cellules des patients (Stoetzel et al. 2016) mais qui n'a pas été mise en évidence dans les rétines des souris.

Accession	2018_S40_Kidney_IP@VPS15										Moyennes #spectra				Gene name		
	CTRL_WT1	CTRL_WT2	CTRL_WT3	VPS15_WT1	VPS15_WT2	VPS15_WT3	CTRL_KI1	CTRL_KI2	CTRL_KI3	VPS15_KI1	VPS15_KI2	VPS15_KI3	CTRL_WT	VPS15_WT		CTRL_KI	VPS15_KI
Pik3r4_MOUSE	0	0	0	177	157	163	0	0	0	135	129	147	0,00	165,67	0,00	137,00	Pik3r4 Vps15
Pik3cb1_MOUSE	0	0	0	33	27	32	0	0	0	24	24	22	0,00	31,11	0,00	27,50	Pik3cb1 Vps15
Becn1_MOUSE	0	0	0	60	59	59	0	0	0	45	40	48	0,00	59,33	0,00	43,67	Becn1
Uvrag_MOUSE	0	0	0	52	51	58	0	0	0	39	44	44	0,00	53,67	0,00	42,33	Uvrag Uvrag1
Svl1_MOUSE	1	0	0	9	13	10	1	2	0	16	14	15	0,33	10,67	1,00	15,00	Svl1
Htrf1_MOUSE	0	0	0	6	13	9	2	1	0	16	17	9	0,00	9,00	1,00	14,00	Htrf1
Mirc3b_MOUSE	0	0	0	4	9	10	0	0	0	13	15	13	0,00	7,67	0,00	13,67	Mirc3b
Hpl1_MOUSE	0	0	0	7	12	7	2	1	0	15	14	10	0,33	8,67	1,00	13,00	Hpl1
Symc_MOUSE	0	0	0	9	9	9	0	0	1	12	12	8	0,00	9,00	0,33	10,67	Symc
Dendf1a_MOUSE	0	1	1	7	5	7	1	2	0	10	9	13	0,67	6,33	1,00	10,67	Dendf1a
Rl2f1_MOUSE	0	0	0	6	9	8	1	0	2	9	11	11	0,00	7,67	1,00	10,33	Rl2f1
Nckap1_MOUSE	0	0	0	4	8	9	1	1	0	7	13	10	0,00	7,00	0,67	10,00	Nckap1 Hem2 Kiaz
Rl2fa_MOUSE	0	0	2	7	8	9	1	1	0	11	9	10	0,67	8,00	0,67	10,00	Rl2fa
Linn1_MOUSE	1	0	0	2	5	1	1	0	2	9	11	6	0,33	2,67	1,00	9,33	Linn1
Gak_MOUSE	0	0	0	6	11	5	0	0	0	11	7	9	0,00	7,33	0,00	9,00	Gak
Co6a2_MOUSE	0	0	1	5	3	2	1	1	0	9	9	8	0,33	3,33	0,67	8,67	Co6a2
Ssf2_MOUSE	0	0	0	4	6	5	1	1	0	8	9	7	0,00	5,00	0,67	8,00	Ssf2
Dock2_MOUSE	1	0	0	4	9	4	1	2	0	9	6	9	0,33	5,67	1,00	8,00	Dock2 Gm30 Kiaz
Rrbp1_MOUSE	0	2	0	4	5	4	0	2	0	9	9	6	0,67	4,33	0,67	8,00	Rrbp1
Mpp1p_MOUSE	0	1	1	2	1	4	1	1	0	8	8	8	0,67	2,33	0,67	8,00	Mpp1p Kiaz0864 M
Dctn4_MOUSE	0	0	0	6	7	6	1	0	2	7	8	8	0,00	6,33	1,00	7,67	Dctn4
Gas2_MOUSE	0	0	0	7	4	6	1	0	0	8	8	7	0,00	5,67	0,33	7,67	Gas2 Gas-2
Rpl37a_MOUSE	0	0	1	7	9	8	0	0	3	7	9	7	0,33	8,00	1,00	7,67	Rpl37a
Paps2_MOUSE	0	0	0	1	5	0	0	1	1	7	8	8	0,00	2,00	0,67	7,67	Paps2 Atp5k2
Coro2b_MOUSE	0	0	0	5	7	2	0	0	0	6	6	8	0,00	4,67	0,00	7,33	Coro2b
Rargr1_MOUSE	0	0	0	8	12	9	0	0	0	7	6	8	0,00	9,33	0,00	7,00	Rargr1 Rgs4L D34
Pgsm_MOUSE	0	0	0	5	5	3	0	0	0	9	7	5	0,00	4,33	0,00	7,00	Pgsm
Tw1f1_MOUSE	0	0	0	2	4	2	1	0	0	7	7	6	0,00	2,67	0,33	6,67	Tw1f1 P19
Flnc_MOUSE	0	0	1	4	7	5	1	1	1	7	7	6	0,33	5,33	1,00	6,67	Flnc At3l P12
Sytl5_MOUSE	1	0	0	5	5	3	1	0	1	10	5	5	0,67	4,67	0,67	6,67	Sytl5
Fcho2_MOUSE	0	0	0	3	5	2	0	0	0	7	6	6	0,00	3,33	0,00	6,33	Fcho2
Oxda_MOUSE	1	0	1	4	5	5	1	1	1	7	7	5	0,67	4,67	1,00	6,33	Dao Dao1
Ablm1_MOUSE	0	0	0	1	1	1	0	0	0	8	7	4	0,00	1,00	0,00	6,33	Ablm1 Ablm Kiaz
Hsp27_MOUSE	0	0	0	3	2	5	0	0	0	8	5	4	0,00	3,33	0,00	5,67	Hsp27 Hsp27
S100a4_MOUSE	0	1	1	2	4	4	1	1	1	7	6	4	0,67	3,33	1,00	5,67	S100a4 Cap1 Mts1
M20a_MOUSE	0	0	0	4	8	1	0	1	1	6	5	5	0,00	4,33	0,67	5,33	Sk25a11
Coat1_MOUSE	0	0	0	3	3	3	2	0	0	6	4	6	0,00	3,00	0,67	5,33	Coat1
Tecp_MOUSE	0	0	0	5	5	2	0	0	0	7	5	4	0,00	2,33	0,00	5,33	Tecp
Dapk1_MOUSE	0	0	0	2	0	0	0	0	0	9	4	3	0,00	0,67	0,00	5,33	Dapk1
Arpc5_MOUSE	0	0	0	3	5	3	0	0	0	5	6	4	0,00	3,67	0,00	5,00	Arpc5
Htrf3_MOUSE	0	0	0	1	3	3	0	0	0	3	3	3	0,00	2,33	0,00	3,00	Htrf3

**Figure 63 :** Résultats de spectrométrie de masse pour le rein. Tableau de résultats récapitulant les interactants connus de *Vps15* et les nouveaux interactants potentiels spécifiques du rein.



**Figure 64 :** Analyses d'enrichissements et validations des interactions pour le rein. **A.** Diagramme de Venn représentant le nombre de protéines identifiées par spectrométrie de masse partagées ou spécifiques d'un génotype. **B.** Histogramme montrant l'enrichissement protéique à partir de la liste de 61 protéines uniquement identifiées chez les KI. **C.** Volcano-plot représentant la quantité de chacune des protéines en interaction avec *Vps15* en fonction du génotype (KI vs WT). **D.** Résultats de la Co-IP anti-*Vps15* suivi du WB pour la validation des interactions.

### 3. Analyse portant sur les cellules des souris

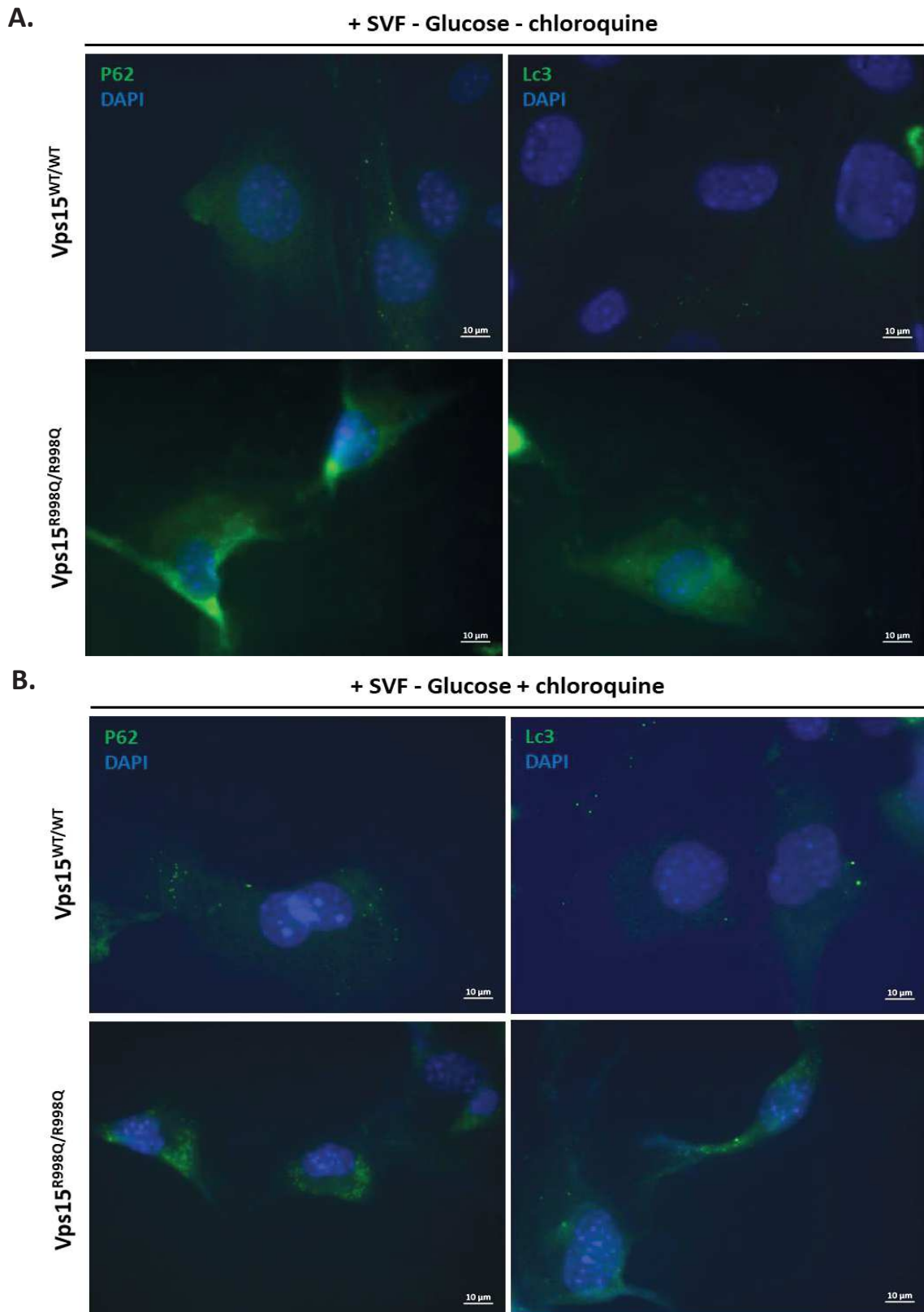
#### A) Marquage de LC3 et P62 avec/sans Chloroquine

D'un point de vue macroscopique, nous n'avons pas mis en évidence de preuve significative de la présence d'un phénotype typique de ciliopathie chez les souris *Vps15*. Afin de déterminer, le phénotype à l'échelle cellulaire, j'ai réalisé des explants de peau à partir des bouts d'oreilles des souris *Vps15*. Les fibroblastes murins (WT et KI) maintenus en culture ont été soumis à des conditions de cultures variées : -SVF pour induire la ciliogénèse, -Glc pour stimuler le processus d'autophagie. Les expériences menées sur les fibroblastes murins sont préliminaires : elles ont permis la mise en point et l'optimisation des protocoles pour la mise en œuvre d'une étude ultérieure à plus grande échelle. Les expériences menées sur les cellules en condition de ciliogénèse (-SVF pendant 24h) devront être conduites en triplicats. Le marquage du cil primaire a été réalisé par IF avec un anticorps dirigé contre l'alpha-tubuline-acétylée (photos non montrées). Des comptages du nombre de cellules ciliées et des mesures de la taille des cils devront être réalisés afin de déterminer si les fibroblastes des souris *Vps15*<sup>R998Q/R998Q</sup> présentent un phénotype de ciliopathie.

Les expériences menées sur les cellules en condition de déprivation (-Glc pendant 24h) devront également être conduites en triplicats mais les résultats de l'étude préliminaire sont encourageants. Cette condition de culture stimule le processus d'autophagie. Ainsi, par IF, les protéines P62 et Lc3 ont été marquées afin de déterminer le niveau basal d'activation de cette voie en fonction du génotype. Les résultats montrent un marquage plus intense dans les fibroblastes des souris *Vps15*<sup>R998Q/R998Q</sup> (Figure 65A). Le niveau basal d'autophagie semble donc plus important dans les cellules des souris KI par rapport aux WT.

L'ajout de chloroquine pendant 6h sur des cellules en condition -Glc (pour un total de 24h) permet de bloquer les étapes tardives du processus d'autophagie en empêchant la fusion des autophagosomes avec les lysosomes. Les autophagosomes sont détectables dans le cytoplasme en marquant les protéines P62 et Lc3. Ils s'accumulent dans le cytoplasme des cellules sous formes de « puncta » (Nemazanyy et al. 2013). Dans les fibroblastes des souris *Vps15*<sup>R998Q/R998Q</sup>, ces « puncta » sont présents en plus grande quantité par rapport aux WT (Figure 65B), révélant l'accumulation des autophagosomes. Le niveau basal d'autophagie dans les cellules des souris KI semble effectivement augmenté par rapport aux WT.





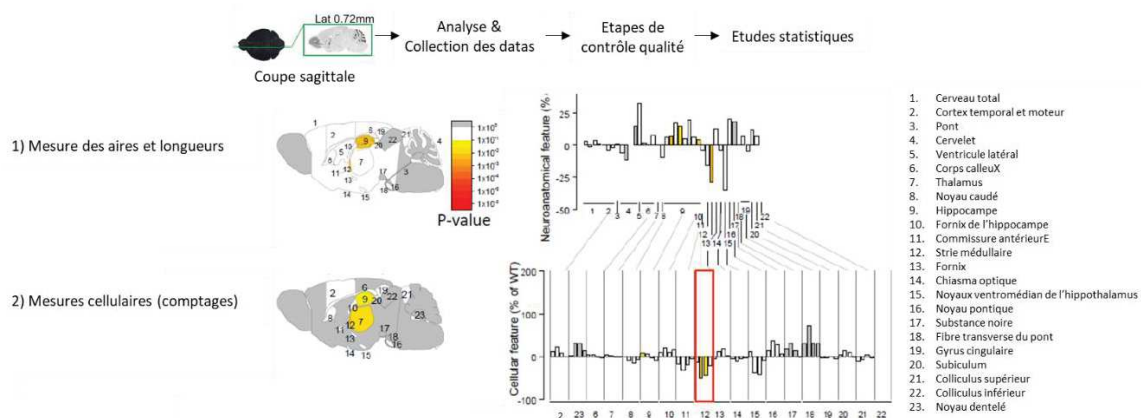
**Figure 65 :** Travail portant sur les fibroblastes murins. Résultats des IF réalisées sur les fibroblastes murins de 1M (n=1) : Marquage dirigé contre les protéines P62 et LC3b (vert) et le noyau (DAPI, bleu) dans des conditions de culture : +SVF, -glucose (24h), **A.** - chloroquine ou **B.** + chloroquine (50  $\mu$ M, 6h).

#### 4. Analyse portant sur le cerveau

Les patients *VPS15* présentent une déficience intellectuelle légère : un type d'anomalie cérébrale typiquement retrouvée dans le cadre des ciliopathies. Afin de déterminer la présence d'une atteinte cérébrale chez les souris, nous avons mis en place une collaboration avec le Dr Yalcin Binnaz (IGBMC, Illkirch) en 2017.

##### A) Etude préliminaire

L'étude morphométrique développée par l'équipe de Y. Binnaz a précédemment été appliquée à d'autres modèles murins pour déterminer les zones cérébrales dans lesquelles les tailles et les densités cellulaires varient par rapport aux souris WT (S. C. Collins et al. 2018). Dans cette optique, l'étude a été menée sur les femelles KI *Vps15* (n=3) de 5 mois et a révélé une réduction significative de la taille et de la densité cellulaire des régions associées à l'hippocampe (Figure 66B).



**Figure 66 :** Travaux réalisés en collaboration avec l'équipe de Y. Binnaz. Résultats de l'étude morphométrique réalisée sur une cohorte de femelles de 5 mois (n=3).



➤ Les 1<sup>ères</sup> conclusions

La 1<sup>ère</sup> vague d'expériences mise en place dans ce projet de caractérisation phénotypique avait pour objectif de déterminer si le nouveau modèle murin *Vps15<sup>R998Q/R998Q</sup>* était un bon modèle d'étude pour la ciliopathie ultra rare rapportée dans Stoetzel et al. 2016. Nous n'avons pas montré d'anomalies ni rétinienne ni rénale significatives aux âges étudiés, ce modèle murin ne semble donc pas être un bon modèle d'étude pour cette ciliopathie car il ne récapitule pas les traits phénotypiques typiques de ces pathologies.

Une étude préliminaire réalisée sur les cerveaux des souris en collaboration avec Yalcin Binnaz a mis en évidence des anomalies morphométriques au niveau des hippocampes des souris. Nous avons donc approfondi la caractérisation du phénotype cérébral de la souris KI *Vps15*.

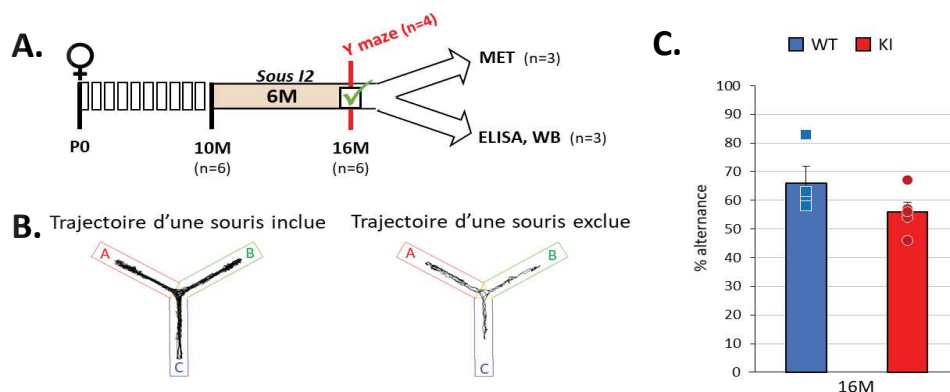
Des premiers résultats de MET réalisée sur quelques souris KI âgées de plus de 1 an ont révélé, de manière inattendue, la présence de plaques cérébrales dans l'hippocampe (photos non montrées). Le phénotype mis en évidence avec ce « coup de sonde » a déjà aidé à orienter le projet. Bien évidemment, étant donné le peu de souris étudiées, nous étions conscients qu'il faudrait confirmer ces observations avec un effectif plus important d'animaux. En plus, les recherches bibliographiques portant sur les maladies neurodégénératives associés à un phénotype de plaque, nous ont aidé à définir les expériences à mettre en œuvre afin de préciser le phénotype cérébral des souris *Vps15*. La 2<sup>ème</sup> vague d'expériences imaginée pour ce projet avait pour objectif de répondre à une nouvelle hypothèse de travail : Est-ce que le modèle murin *Vps15<sup>R998Q/R998Q</sup>* pourrait présenter les principaux critères de la maladie d'Alzheimer et être, finalement, un modèle d'étude pour cette pathologie ?

Une 2<sup>ème</sup> vague d'expérience a débuté en 2019 avec la constitution de plusieurs cohortes de souris, de sexes différents. Certaines étaient nourries avec une nourriture normale dite « chow diet » (cohorte 3). D'autres ont été soumises à une nourriture spéciale dite « high tout, I2 » afin de les « challenger » (cohorte 1 et 2). Sa composition riche en cholestérol, fructose et graisse est propice au développement des plaques d'athérosclérose et des plaques cérébrales comme rapporté dans Mancini et al. 2018. L'idée était de déterminer si, sous cette nourriture, les souris KI allaient présenter un phénotype plus sévère par rapport aux WT : prouvant un impact de la mutation sur le phénotype provoqué.

### B) Etude de la mémoire de travail

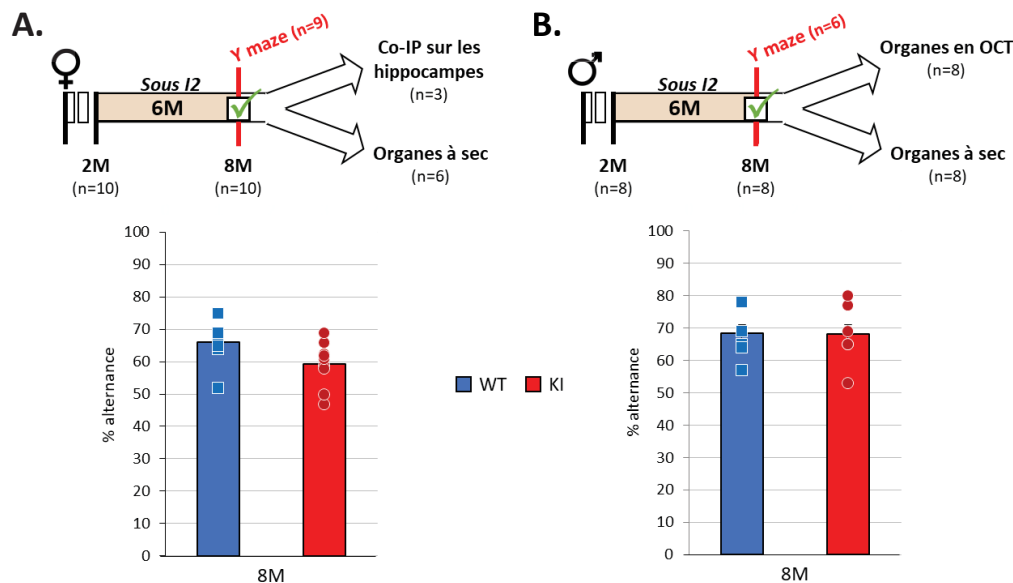
Toutes les cohortes de souris, quelque soit leur nourriture, ont été soumises au test de comportement Y-maze. Ce test de référence permet de calculer le pourcentage d'alternance entre les 3 bras du labyrinthe pour une souris donnée. Une valeur diminuée par rapport à un contrôle est synonyme d'une altération de la mémoire de travail à court terme. La souris « ne se souvient pas » qu'elle a déjà explorée ce bras et va retourner l'explorer. Ce type de trouble cognitif fait partie des critères associés à la maladie d'Alzheimer.

Une première cohorte (cohorte 1) composée de femelles, ont été mises sous I2 pendant 6 mois entre 10M et 16M. Elles ont été soumises au Y-Maze à 16M (Figure 67A). Lors du Y-Maze, 4 souris (n=2 WT et n=2KI) ont été exclues car elles n'avaient pas exploré un minimum de 10 bras. Il s'agit du seuil minimum à atteindre pour pouvoir considérer que la souris a exploré le labyrinthe. Ainsi, une souris n'ayant pas explorée (=exclue) est aisément visualisable sur les trajectoires obtenues à l'issue du test (Figure 67B). Pour cette cohorte, le pourcentage moyen d'alternance est diminué par rapport aux WT (Figure 67C) avec un effectif de n=4 seulement. En augmentant l'effectif, la diminution sera peut-être statistiquement significative.



**Figure 67 : Etude de la mémoire par Y-maze, Cohorte 1.** A. Représentation schématique de la chronologie des événements pour cette cohorte. B. Comparaisons des trajectoires entre une souris ayant explorée (inclue dans l'analyse) et une souris n'ayant pas explorée (exclue). C. Histogramme représentant les résultats du Y-maze pour n=4 femelles de 16M (sous I2 depuis 6 mois). Chaque animal est représenté par un point.

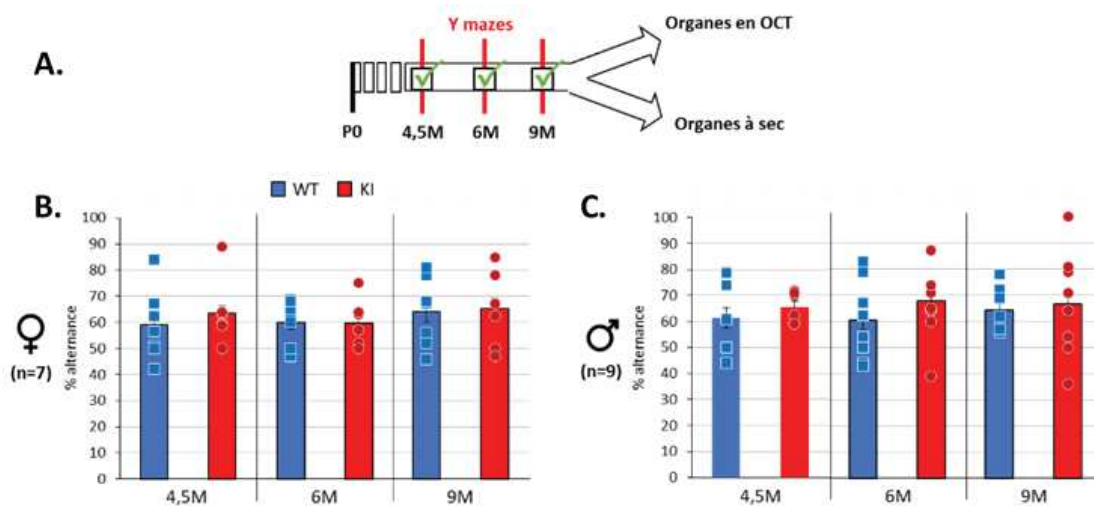
Des cohortes de mâles et de femelles (cohorte 2), nourries avec la nourriture I2 pendant une période de 6M dès leur 2M, ont été soumises au test de comportement Y-Maze à 8M. La valeur moyenne du pourcentage d'alternance pour les femelles KI est légèrement diminuée par rapport aux WT (Figure 68A). Pour les mâles, il n'y a pas de différences entre les 2 génotypes (Figure 68B). Il semblerait donc que l'impact cognitif soit plus important chez les femelles. Pour consolider cette hypothèse, il faudrait augmenter l'effectif pour réduire au maximum les variations intra-groupes.



**Figure 68 : Etude de la mémoire par Y-maze, Cohorte 2.** Représentation schématique de la chronologie des événements pour cette cohorte de souris placée 6 mois sous I2 et analysée à 8M. Histogramme représentant les résultats des Y-maze pour **A.** les femelles (n=9) et **B.** les mâles (n=6). Chaque animal est représenté par un point.

Les cohortes de mâles et femelles n'ayant pas été mises sous la nourriture I2 (cohorte 3) ont subi une série de Y-maze à des âges croissants (4,5M, 6M et 9M) afin de mettre en évidence l'apparition d'une potentielle atteinte cognitive au cours du temps (Figure 69A). Les pourcentages d'alternance au sein de chaque groupe sont très hétérogènes. L'ensemble des résultats ne montre pas de différence entre les génotypes, quelque-soit le sexe (Figure 69B et C).

Il semblerait donc que le trouble cognitif soit accentué et potentiellement observable seulement chez des femelles dans des conditions où elles sont nourries avec le I2. L'effet semble accentué quand l'âge des femelles est plus avancé. Ces résultats de comportements préliminaires sont encourageants et devront être reproduits avec des effectifs plus importants d'animaux pour émettre une conclusion plus solide.



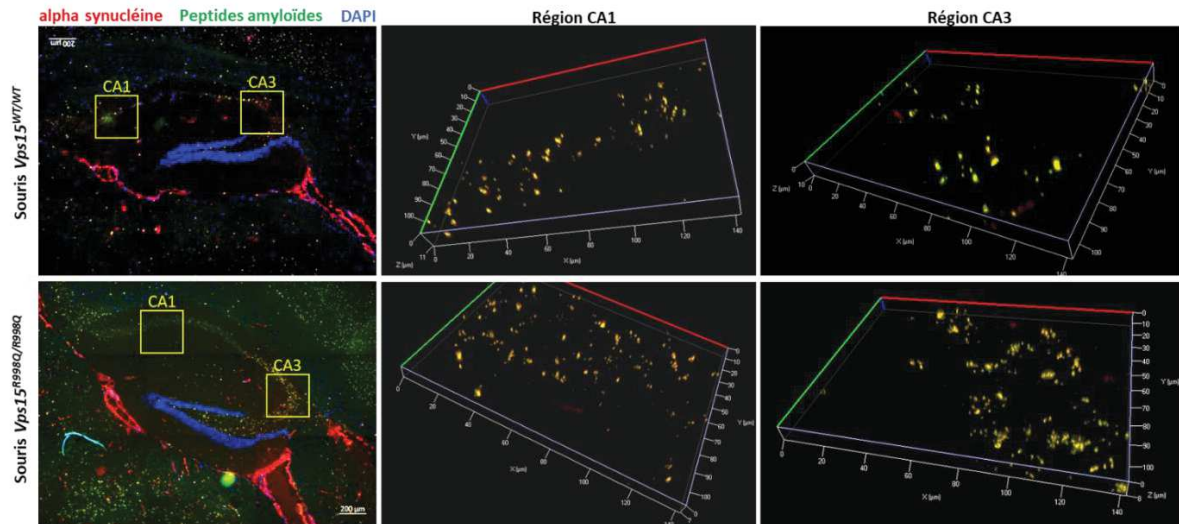
**Figure 69 : Etude de la mémoire par Y-maze, Cohorte 3.** A. Représentation schématique de la chronologie des événements pour cette cohorte de souris nourrie par « chow diet ». Histogramme représentant les résultats des Ymazes pour B. les femelles (n=7) et C. les mâles (n=9). Chaque animal est représenté par un point.

C) Analyse de l'ultrastructure des régions hippocampiques

Afin de mettre en évidence d'autres critères associés à la maladie d'Alzheimer, nous avons réalisé de la MET sur les hippocampes des femelles (n=3) de la cohorte 1 (Figure 67A). Les régions CA1 et CA3 ont été observées. Des plaques, une accumulation de filaments et des morts cellulaires de neurones ont été mises en évidence seulement chez les souris KI dans les 2 zones. (Figure 71).

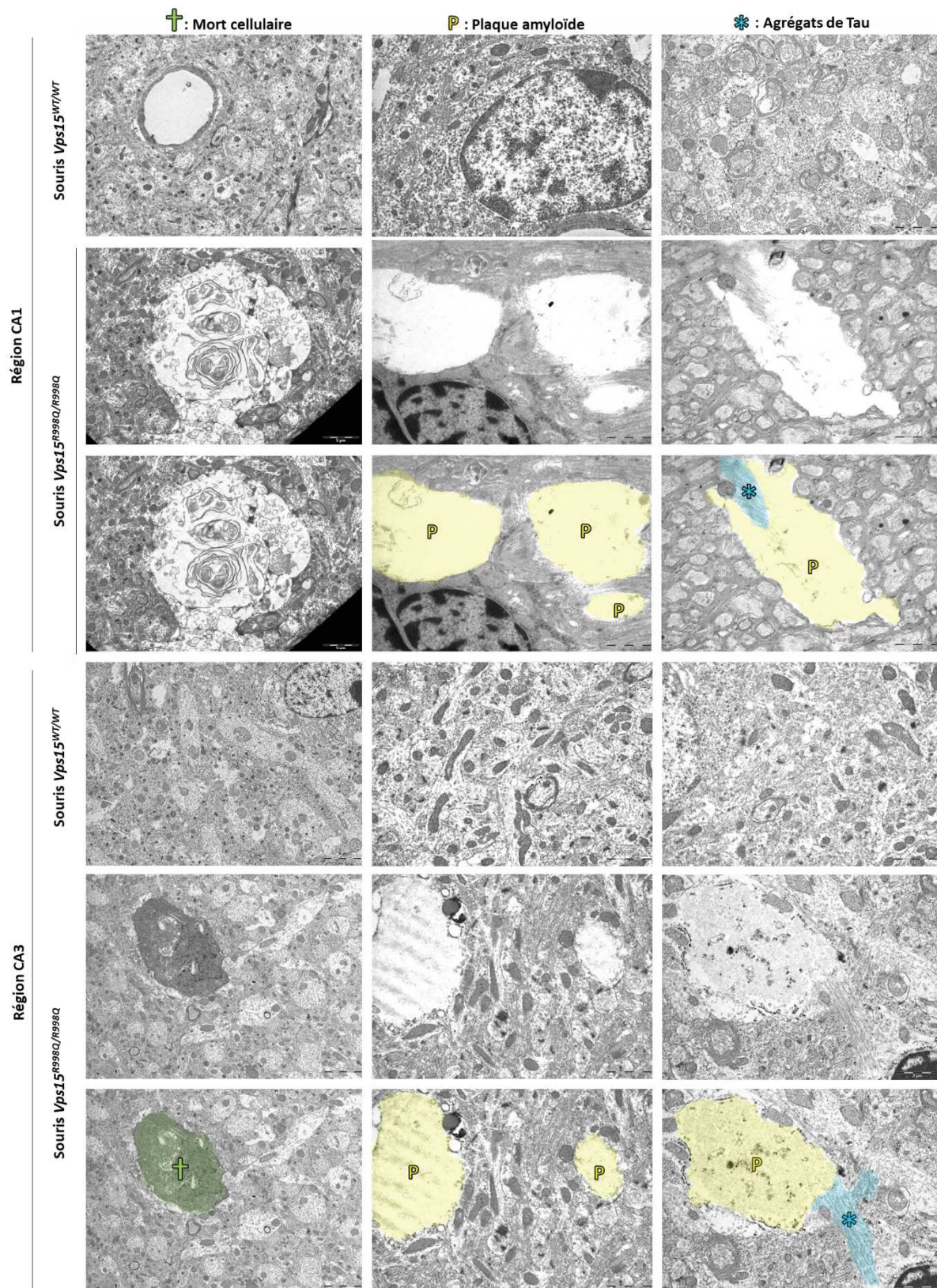
D) Déterminer la nature des plaques

Afin de déterminer la nature des plaques visualisés par MET, nous avons réalisé des IF sur des cryocoupes de cerveaux issus d'une cohorte de n=3 mâles de 1 an, nourris par « chow diet ». Nous avons marqué l' $\alpha$ -synucléine (une protéine ayant été rapportée dans la composition de plaque amyloïde de patients (Towhig et Nielsen 2019)) et les peptides A $\beta$ -amyloïdes (1-40, 1-42). Nous avons recherché les zones de colocalisations au niveau de l'hippocampe (Figure 70). Nous avons pris des photos plus précisément au niveau des zones CA1 et CA3 en mode Z stack. Nous avons ainsi visualisé la colocalisation des protéines marquées à la surface des cellules. Nous avons observé une accumulation des agrégats de colocalisations à la surface des cellules des souris KI *Vps15*. Il faudrait reproduire ce marquage sur les cryocoupes de cerveaux issus des cohortes décrites précédemment notamment celles sous I2 afin de confirmer ces premières observations.



**Figure 70 : Marquages protéiques sur des cryocoupes de cerveaux.** Résultats du co-marquage anti- $\alpha$ -synucléine et anti-peptides  $A\beta$ -amyloïdes par IF. Photos d'hippocampes obtenues après marquage de cryocoupes de cerveaux de souris mâles de 1 an ( $n=3$ ). Photos en mode Z stack à la surface des régions CA1 et CA3.

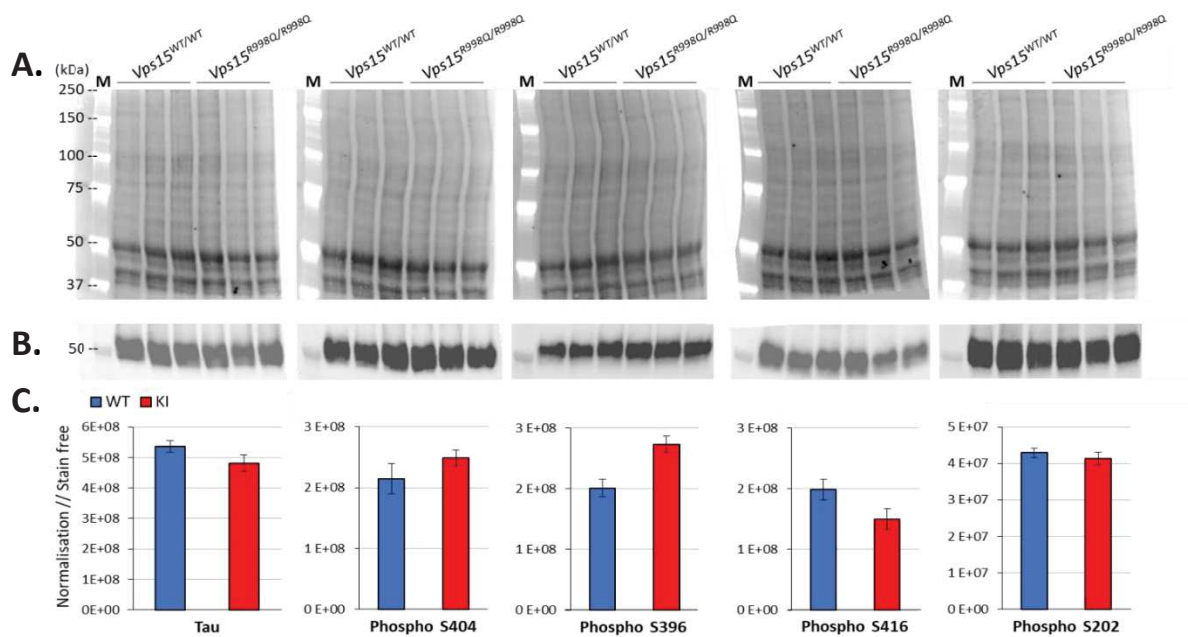




**Figure 71 : Observation de l'hippocampe par MET.** Résultats de MET sur les régions hippocampiques CA1 et CA3 des femelles de la cohorte 1 (n=3). Les critères histologiques observables sur les photos ont été surlignés par des marquages de couleurs.

E) Etude de la protéine Tau

L'hyperphosphorylation de la protéine Tau est l'un des critères associés à la maladie d'Alzheimer. J'ai réalisé des WB sur les protéines totales extraites à partir des cerveaux des femelles de la cohorte 1 (n=3). J'ai marqué la protéine Tau et certaines des sérines de cette protéine potentiellement phosphorylées (Figure 72B). J'ai fait une quantification en normalisant par rapport au stain free (Figure 72A) afin de déterminer le niveau de phosphorylation de chaque sérine en fonction du génotype (Figure 72C). Il est très difficile de conclure face à ces résultats car la quantité protéique déposée (40µg) est trop importante ce qui biaise les quantifications. Cette expérience est donc à refaire.



**Figure 72 : Tau et ses sérines phosphorylées.** Résultats des WB sur les protéines totales extraites des cerveaux des femelles de la cohorte 1 (n=3). **A.** Photos des 5 membranes « stain free » avant l'étape de révélation. **B.** Résultats des marquages de la protéine Tau et de ces sérines phosphorylées. **C.** Histogrammes représentant la quantité de protéine Tau et le niveau de phosphorylation des sérines en fonction du génotype.

➤ 2<sup>ème</sup> conclusion partielle sur le phénotype cérébral

Finalement, les résultats de cette 2<sup>ème</sup> vague d'expériences nous ont permis de « dégrossir » le sujet et d'effectivement mettre en évidence certains des critères associés à la maladie d'Alzheimer, notamment, la présence de plaques au niveau de l'hippocampe, révélées par MET. Les résultats préliminaires des IF montrent qu'il s'agirait d'agrégats de type  $\beta$ -amyloïdes au niveau de l'hippocampe. Les résultats de Y-Maze et WB restent très préliminaires et donnent seulement des tendances car ils ne sont pas significatifs. Ils devront être confirmés en augmentant l'effectif des cohortes analysées. Il faudrait notamment étoffer la cohorte 1. Il serait intéressant de réaliser le même schéma d'expériences avec des mâles du même âge nourris au I2 pour voir si le sexe joue un rôle dans l'apparition des signes.

Dans la troisième et dernière vague d'expériences, nous avons poursuivi nos investigations pour comprendre pourquoi la mutation R998Q dans le gène *Vps15* est responsable d'un phénotype de

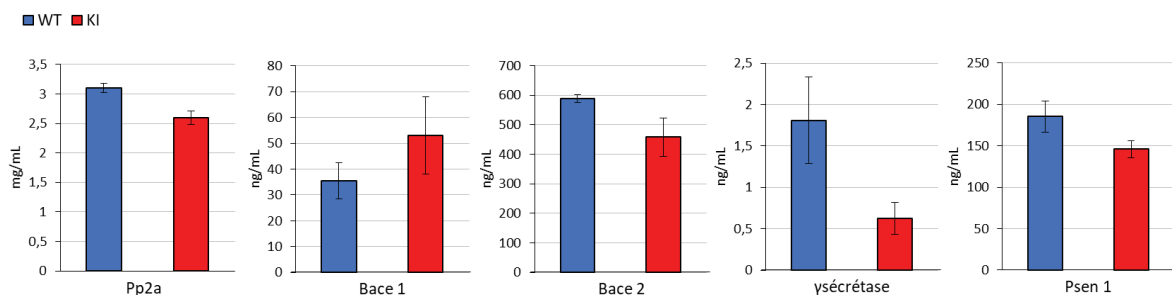


plaque chez les souris. Pour l'instant, nous avons effectué une série d'Elisa afin de doser les quantités protéiques des enzymes impliquées dans la synthèse des peptides A $\beta$  (à partir des cerveaux des femelles de la cohorte 1). D'autres expériences sont en cours d'analyse à l'heure actuelle (Full moon et Co-IP). Les amyloïdoses sont des maladies dont la maladie d'Alzheimer fait partie. Elles sont caractérisées par l'accumulation de protéines agrégées sous forme de plaques dans une multitude d'organes. Par exemple, le diabète de type 2 peut être provoqué par la présence, autour des îlots de Langerhans, d'une plaque composée d'amyline qui empêche le relargage dans le sang des hormones synthétisées par les cellules  $\alpha$  (glucagon) et  $\beta$  (insuline) (Lorenzo et al. 1994). Nous avons donc étendu notre exploration au pancréas pour voir si une plaque pouvait effectivement être visualisée par MET.

## 5. Analyses exploratoires

### A) Dosages protéiques cérébrales

Afin de préciser le mécanisme pouvant expliquer le développement de la plaque observée en MET, nous avons réalisé une série de tests Elisa afin de doser dans le cerveau des souris de 16M (cohorte 1) les quantités de protéines impliquées dans la formation des peptides A $\beta$  amyloïdes 1-40 et 1-42 à partir du précurseur APP. Nous avons donc déterminé la quantité de Bace 1 et 2 ( $\beta$  secrétases) et de la  $\gamma$  secrétase ainsi que celle de sa sous unité catalytique Psen1 (préséniline 1). Le résultat obtenu pour Bace 1 n'est pas significatif entre les 2 génotypes : les barres d'erreur se chevauchent. Les quantités de Bace 2, de  $\gamma$  secrétase et de Psen1 semblent réduites chez le KI par rapport au WT. Nous avons également dosé la quantité protéique de Pp2a, une phosphatase qui réalise la déphosphorylation des sérines de la protéine Tau. Sa quantité semble moindre chez les souris *Vps15* (Figure 73). En complément de ces tests Elisa, il serait intéressant de déterminer le niveau d'activité enzymatique de ces protéines. Ce niveau pourrait être augmenté chez les KI pour compenser l'écart dû à la réduction de la quantité protéique et maintenir un niveau d'activité normal.



**Figure 73 : Dosages protéiques.** Résultats des tests Elisa réalisés sur les protéines extraites des cerveaux des femelles de la cohorte 1 (n=3).

### B) Etude protéique (en cours)

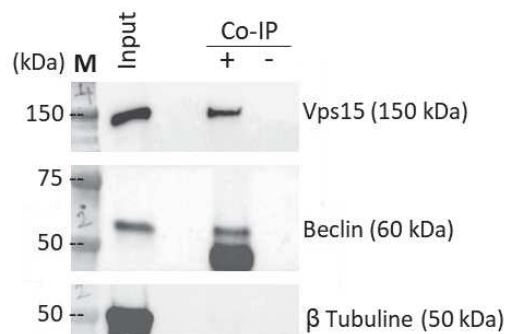
Le Full moon est une expérience qui permet l'exploration de la quantité de nombreuses protéines (Elisa) et de la phosphorylation de certaines protéines sur une lame (IF). J'ai réalisé l'expérience

« phosphoexplorer » sur les protéines extraites de cerveau des souris de 16M mais je n'ai pas finalisé l'analyse du grand nombre de données obtenues.

### C) Etude de l'interactome de Vps15 (en cours)

Afin de comprendre plus précisément le mécanisme reliant le génotype au phénotype cérébral obtenu, nous avons réalisé des Co-IP anti-Vps15 sur des tranches de cerveaux de 3mm d'épaisseur contenant l'hippocampe. Nous avons utilisé les cerveaux des femelles de la cohorte 2 (Figure 68A). Le cerveau a été découpé au moyen d'un moule appelé « brain slicer » dédié aux souris adultes permettant de découper transversalement tous les 1mm des tranches de cerveau de manière reproductible.

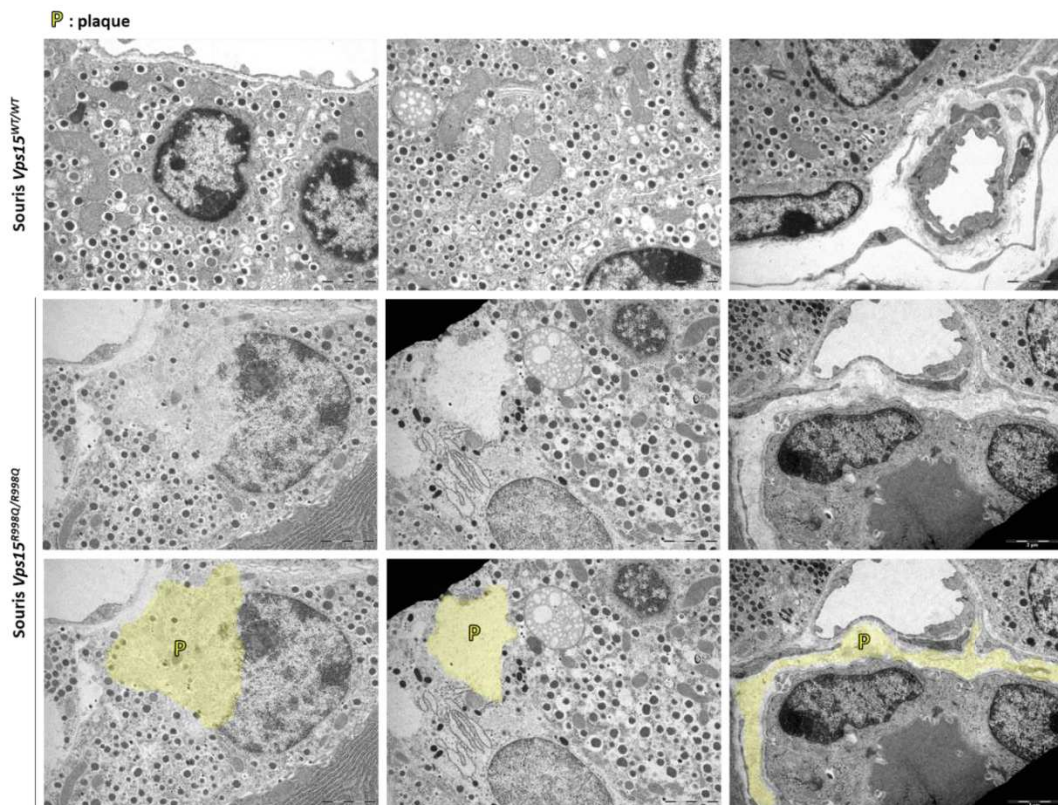
Pour déterminer (1) les conditions optimales pour la mise en place de l'expérience et (2) si l'anticorps anti-Vps15 utilisé dans les 2 Co-IP précédentes serait applicable pour cette expérience sur les cerveaux, j'ai effectué une première mise au point de Co-IP avec une tranche d'hippocampe contrôle. Sur le WB, j'ai révélé dans la Co-IP+ la protéine appât Vps15 et Beclin-1 (une protéine connue pour être en interaction avec Vps15) (Figure 74). Nous avons ensuite mis en place l'expérience de Co-IP suivie de spectrométrie de masse sur n=3 cerveaux des femelles de la cohorte 2 en collaboration avec l'équipe de Philippe Hamann (IBMC, Strasbourg). Pour l'instant, nous n'avons pas pu analyser les résultats récemment reçus. Ils devraient pouvoir nous aiguiller sur les interactions spécifiques à cette zone du cerveau (comprenant l'hippocampe) et peut être montrer les interactions perdues ou gagnées selon le génotype.



**Figure 74 : Mise au point de Co-ip anti Vps15 dans le cerveau.** Résultats obtenus sur les protéines de cerveaux extraites à partir d'une tranche d'hippocampe de souris contrôle. Photos de la membrane marquée avec des anticorps dirigés contre Vps15, contre Beclin-1 : un interactant connu de Vps15 (contrôle positif) et contre la βtubuline : une protéine n'interagissant pas avec Vps15 (contrôle négatif).

D) Analyses de l'ultrastructure du pancréas

Nous avons réalisé de la MET sur les pancréas des femelles de la cohorte 1 (n=3). Une plaque est visualisable autour des îlots de Langerhans chez les souris KI (Figure 75). Ce résultat vient conforter le phénotype cérébral mis en évidence. Etant donné que nous avons révélé la présence de plaques au niveau du cerveau et du pancréas des femelles de la cohorte 1, il serait intéressant d'étendre nos recherches à d'autres organes.



**Figure 75 : Observation du pancréas par MET.** Résultats de MET sur les pancréas des femelles de la cohorte 1 (n=3) : La plaque a été surlignée par une couleur jaune sur les photos.

➤ Dernière conclusion sur la 3<sup>ème</sup> vague

Dans le cadre de cette dernière vague, les expériences de Co-IP et de Full-moon sont en cours d'analyses et leurs interprétations devraient nous aiguiller sur le possible mécanisme physiopathologique expliquant le phénotype. Nous avons montré la présence de plaques au niveau des pancréas des souris *Vps15* par MET. Ce résultat nous permet d'étoffer nos premières conclusions et d'envisager le modèle murin comme un modèle d'amyloïdoses.

### III. La discussion et les perspectives

Lorsque j'ai débuté le projet portant sur le modèle murin *Vps15*, le principal objectif était de procéder à la caractérisation phénotypique de la souris en dirigeant nos efforts sur les 2 organes principalement touchés chez les patients : l'œil et le rein. Nous voulions savoir si cette souris était un bon modèle d'étude pour la ciliopathie. Pour chaque organe, nous avons essayé de suivre un schéma parallèle associant des analyses de type ERG ou des analyses biochimiques du plasma pour déterminer si le fonctionnement de l'organe en question était altéré. Nous avons également mené des études moléculaires pour déterminer si des interactions étaient perdues ou gagnées (Co-ip) ou si les patrons d'expression au niveau tissulaire étaient modifiés (histologie). Au fur et à mesure des résultats obtenus il a fallu préciser les objectifs scientifiques du projet ce qui explique pourquoi 3 vagues d'expériences se sont succédées.

#### 1. Est-ce un bon modèle d'étude des ciliopathies ?

A l'issue de la première vague d'expériences, aucune conclusion franche ne se dégage des résultats obtenus. Certaines des expériences telles que la détermination du patron d'expression de la protéine *Vps15* dans l'œil et le rein n'ont pas abouties. Nous nous sommes retrouvés dans une impasse. Pour surmonter le problème, nous mettrons en place une collaboration pour réaliser de l'hybridation *in situ* contre *Vps15* parce que nous n'avons pas le matériel nécessaire pour réaliser ce type d'expérience au laboratoire. Cette technique nous permettrait de déterminer sur des coupes de tissus, les cellules qui expriment l'ARNm de *Vps15*. Ce sont celles qui seraient potentiellement impactées par la mutation R998Q.

##### A) L'interactome de *Vps15* en fonction des organes

Pour la rétine et pour le rein, nous avons réalisé des Co-IP anti-*Vps15* suivies d'analyses par spectrométrie de masse afin de déterminer les interactants de cette protéine dans les tissus de souris de 1M. Parce que nous n'avons pas mis en évidence d'anomalies histologiques dans les tissus, nous avons choisi d'analyser des souris de cet âge pour 2 raisons : (1) les animaux sont adultes, (2) il fallait prendre en considération le besoin de générer et maintenir un effectif important de souris (concernant la rétine) pour atteindre les 3mg de protéines requises pour la réalisation de l'expérience. Nous avons ainsi déterminé des interactants spécifiques pour chaque tissu.

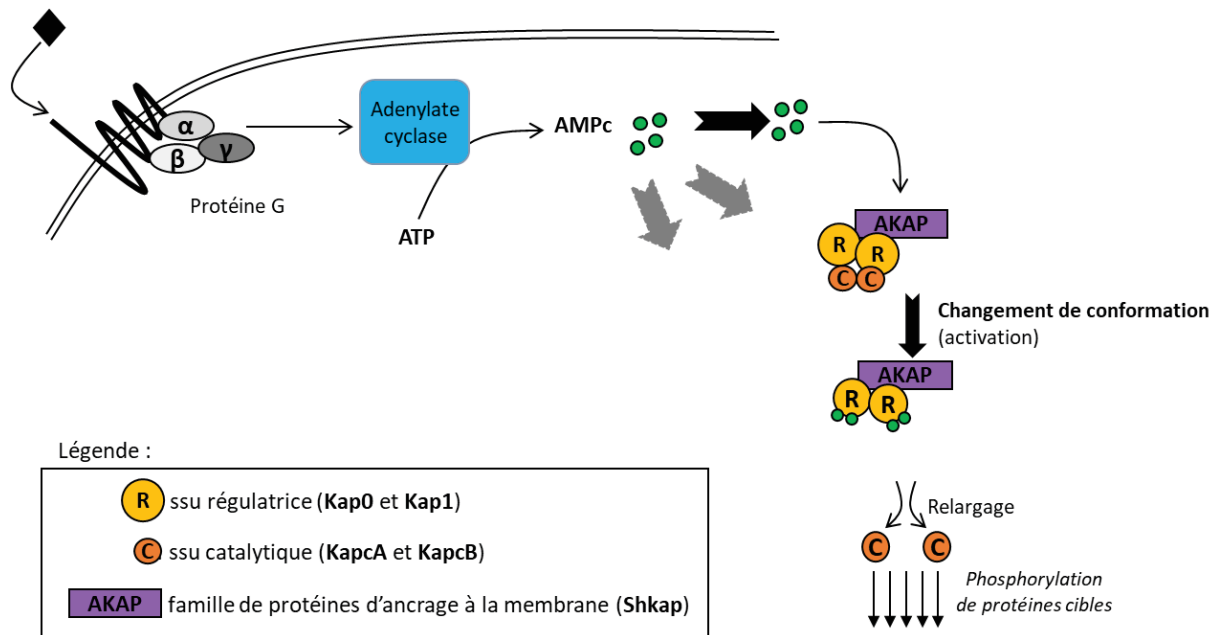
Dans la rétine, nous avons identifié les 5 sous unités (2 sous unités catalytiques, 2 sous unités régulatrices et une protéine d'ancrage) composant l'ensemble du complexe PKA (protéine kinase A) qui appartient à la famille des protéines kinases dépendantes de l'AMPc. Nous avons confirmé par WB l'interaction de *Vps15* avec Kap0, une des sous unités régulatrices. Comme le résume la Figure 76, quand PKA est inactive, les sous unités régulatrices et catalytiques interagissent. L'activation de PKA est dépendante de la fixation d'AMPc au niveau des sous unités catalytiques qui entraîne un

changement de conformation du complexe et permet le relargage des sous unités catalytiques. En conséquence, elles phosphorylent une multitude de protéines cibles et participent à la mise en place de multiples réponses cellulaires. Le premier lien entre PKA et autophagie a été montré dans le modèle de levure *S. cerevisiae* où les protéines ATG1, ATG13 et ATG18 ont été identifiées comme de potentielles cibles de PKA.

Mavrakis et al. 2006 a montré que la sous unité régulatrice de PKA (Ri $\alpha$ ) réside dans les autophagosomes et que l'inactivation du gène dans des MEF de souris (*Prkar1a*<sup>-/-</sup>) active la voie mTOR et inhibe le processus d'autophagie. L'activation de la PKA dans ces cellules mènerait donc à l'activation du processus d'autophagie. Une autre étude (Cherra et al. 2010) a montré que la protéine LC3 peut être directement phosphorylée par PKA, ce qui inhibe le processus d'autophagie dans les neurones. Ces informations montrent que l'impact de PKA sur le processus d'autophagie est double et dépend de l'organe considéré. La protéine codée par *Sphkap* permet l'ancrage du complexe PKA à la membrane. Elle est également impliquée dans l'ancrage et la régulation de l'activité de la protéine SPHK1 (la sphingosine kinase de type 1). Il s'agit d'une kinase, exprimée dans les photorécepteurs et les cellules de Müller entre autre, qui phosphoryle la sphingosine 1P (Lacaná et al. 2002; Simón et al. 2019). Nous n'avons pas retrouvé *Sphk1* parmi les interactants de *Vps15* mais d'autres données de co-IP montrent une interaction entre *Sphkap* et *Sphk1* (Lacaná et al. 2002). Un lien entre cette protéine et l'autophagie a été rapporté par Moruno Manchon et al. 2016 : la kinase est localisée au niveau des autophagosomes en cas de stimulation du processus d'autophagie.

Cet ensemble de données montre qu'il existe un lien entre le processus d'autophagie et PKA. Il faudrait maintenant déterminer les cibles de PKA dans la rétine et le potentiel impact de la mutation R998Q de *Vps15* sur le fonctionnement de cette voie. Nous n'avons pas mis en évidence de différentiel d'interactions entre les souris WT et KI concernant les 5 sous unités identifiées. La mutation de *Vps15* peut potentiellement avoir un impact en aval de la voie PKA, au niveau des protéines cibles, en régulant le niveau d'activation du processus d'autophagie par exemple.

Nous avons également identifié puis confirmé par WB l'interaction entre *Vps15* et Bin1 (Bridging integrator 1 ou Amphiphysine 2) dans la rétine. Il s'agit d'une protéine régulatrice du processus d'endocytose qui est ancrée à la membrane plasmique grâce à un site de liaison au phosphatidylinositol (Pi). La protéine Bin-1 colocalise et interagit avec la protéine Tau : elle est impliquée dans la genèse de la maladie d'Alzheimer (Prokic, Cowling, et Laporte 2014). Il serait donc intéressant de déterminer si *Vps15* est également en interaction avec Bin-1 dans le cerveau des souris ce qui pourrait apparaître comme une première hypothèse pour relier la mutation dans *Vps15* au phénotype cérébral.



**Figure 76 : L'activation de PKA.** Représentation schématique des sous unités (ssu) du complexe PKA.

Dans le rein, nous avons identifié les 3 isoformes des récepteurs Inositol 1, 4, 5 tri-phosphate (Itpr 1, 2 et 3) en interaction avec *Vps15*. Il s'agit de récepteurs, couplés à des canaux ioniques, localisés au niveau du RE qui permettent la libération de  $\text{Ca}^{2+}$  dans le cytoplasme quand ils sont activés. Le rôle de ces récepteurs vis-à-vis de l'autophagie est double. Le  $\text{Ca}^{2+}$  relargué par l'intermédiaire de ces récepteurs est un second messager qui permet d'activer de nombreuses voies de signalisation dont le processus d'autophagie. Par exemple, *Vps34* et *Dapk1* ont été décrites comme des cibles de cette voie (Kabir, Kim, et Chae 2018). L'étude de Criollo et al. 2007 rapporte un rôle antagoniste des Itpr en montrant que ces récepteurs sont en interaction directe avec *Bcl-2*, un inhibiteur du processus d'autophagie qui séquestre *Beclin-1*. La perte des Itpr en recourant aux techniques d'ARNi ou à des antagonistes de ces récepteurs (la xestospongine B par exemple) stimule le processus d'autophagie avec un mécanisme indépendant du  $\text{Ca}^{2+}$ . Il semble donc que les récepteurs Itpr puissent réguler positivement et négativement le processus d'autophagie. Nous pourrions mettre en place des marquages du flux calcique dans les cellules des souris en culture afin de déterminer une potentielle altération du relargage calcique chez les souris KI. Nous pourrions utiliser la xestospongine B sur les cellules des souris KI pour déterminer l'effet sur le processus d'autophagie.

Grâce à l'étude interactomique, nous avons également mis en évidence, dans le rein, une interaction entre *Vps15* et *Dapk1* qui a été validée par WB. Il semble que la quantité de protéine *Dapk1* en interaction avec *Vps15* dans les reins des souris KI *Vps15* soit plus importante par rapport au WT. L'activation de *Dapk1* est dépendante de la liaison avec le  $\text{Ca}^{2+}$  qui pourrait potentiellement être relargué par les récepteurs Itpr. *Dapk1* est connue pour provoquer l'activation d'un programme de mort cellulaire dépendante du processus d'autophagie et est également impliquée dans la



pathogénèse de la maladie d'Alzheimer. Tout comme Bin-1, il serait intéressant de déterminer si *Vps15* est en interaction avec Dapk1 dans le cerveau des souris, ce qui pourrait apparaître comme une deuxième hypothèse pour relier la mutation dans *Vps15* au phénotype cérébral.

Dans l'ensemble, les données d'interactomiques nous laissent envisager de nombreuses pistes expérimentales à explorer vis-à-vis du processus d'autophagie. Les expériences mentionnées pourraient nous aider à préciser l'impact de la mutation de *Vps15* sur le processus d'autophagie. En effet, nous avons obtenu des résultats très encourageants sur les cellules des souris. Le niveau basal d'autophagie est augmenté dans les cellules KI par rapport aux WT. Ces résultats sont préliminaires et devront être reproduits. Néanmoins, ils indiquent une probable altération du processus d'autophagie. Une conclusion cohérente au regard du rôle connu de *Vps15* dans la formation des autophagosomes. La mutation R998Q n'impacte pas l'établissement des interactions entre *Vps15* et les autres protéines. Par contre, la présence de la mutation peut potentiellement perturber l'activation de voies de signalisations dans lesquels *Vps15* est acteur.

Finalement, nous n'avons pas mis en évidence les anomalies typiques de la ciliopathie observée chez l'Homme, telles que la rétinopathie pigmentaire ou la présence de kystes rénaux. Le modèle murin KI *Vps15* n'est donc pas un bon modèle d'étude pour la ciliopathie d'intérêt.

#### B) Explications potentielles à l'absence de phénotype

Le premier modèle murin KO a été étudié en 1989 (Thompson et al. 1989). Depuis des milliers d'autres publications ont rapporté des liens entre génotypes et phénotypes basés sur la caractérisation phénotypique de nouveaux modèles murins. L'objectif commun de ces travaux vise à assigner une fonction à chacun des gènes du génome murin. Par contre, le nombre de modèles murins invalidés pour un gène donné ne présentant pas de phénotype est estimé à 10-15% (Barbaric, Miller, et Dear 2007). Ce nombre est certainement sous-estimé par le manque de publications scientifiques rapportant les résultats négatifs obtenus à l'issue du phénotypage de nouveaux modèles murins. De même, la restriction des explorations à certains organes réduit la possibilité de mettre en évidence des atteintes hors du phénotype attendu. Plusieurs consortiums (comme l'IMPC par exemple) ont été créés ces dernières années afin de générer des modèles KO pour chacun des gènes de la souris et mettre en place un phénotypage standardisé pour l'ensemble des organes des animaux. Parmi ces modèles murins, certains ont montré des symptômes inattendus : le modèle KO pour *Lif* (codant pour la cytokine Leukemia Inhibitory Factor) a d'abord montré, pour les femelles gestantes, un défaut d'implantation des blastocystes (C. L. Stewart et al. 1992). Un phénotype inattendu a été mis en évidence plus tard montrant des problèmes immunitaires associés à des problèmes neurologiques (Pechnick et al. 2004). Clemen et al. 2018 a rapporté le cas d'un modèle murin porteur d'une mutation faux-sens dans *Vcp* responsable d'une sclérose latérale amyotrophique chez l'homme qui ne récapitule



pas le phénotype typique de cette pathologie. Plusieurs hypothèses peuvent être énoncées pour expliquer cette absence de phénotype ou encore la présence d'un phénotype inattendu pour ces modèles animaux. D'abord, nous ne pouvons pas écarter l'impact du fond génétique sur l'apparition d'un phénotype. Des études GWAS (Genome Wide Association Studies) appliquées au génome murin ont montré que certains loci modulent la présence/absence ou encore le niveau de sévérité d'un phénotype (Sittig et al. 2016). Par exemple, la comparaison des phénotypes de souris ayant les mêmes mutations sur les 2 fonds génétiques les plus courants (C<sub>3</sub>H/HeJ et C57BL6J) a montré des différences dans 43% des cas (Barbaric, Miller, et Dear 2007). Dans le cadre du projet portant sur la souris *Vps15*, nous pourrions envisager de changer le fond génétique des animaux par la mise en place de backcross. Nous ne pouvons écarter la possibilité que des phénotypes apparaissent chez des souris *Vps15* ayant d'autres fonds génétiques.

Même si la souris est un bon modèle d'étude pour les pathologies humaines, parmi les 29 000 gènes murins, 61% (17800) sont des orthologues des gènes humains (Schofield, Hoehndorf, et Gkoutos 2012). Il existe donc des protéines n'ayant pas forcément le même mécanisme d'action en fonction de l'espèce considérée. Certaines protéines ne seront produites que dans 1 des 2 espèces. Par exemple, une mutation du gène *Ocr1* est responsable du syndrome oculo-cérébro-rénal (OCRL) chez l'homme et le modèle murin KO pour *Ocr1* ne présente aucun phénotype. Par contre, une souris KO pour un gène inexistant dans le génome humain (*Inpp55b*, en l'occurrence, dont la protéine codante est rapportée en interaction avec la protéine OCRL1 chez la souris) va présenter un phénotype typique d'OCRL. Le rôle des protéines d'intérêts au sein des tissus peut diverger d'une espèce à l'autre. En effet, la publication El-Amraoui et Petit 2014 a permis de mettre en évidence des différences dans l'architecture des photorécepteurs murins par rapport à ceux des primates expliquant ainsi pourquoi les souris mutées dans les gènes *Ush1* et *3* ne développent pas de RP. Enfin, dans certains cas, le phénotype associé est très subtil, le nombre d'animaux inclus pour l'étude est trop limité pour permettre de déceler celui-ci. Pour contourner cette possible limitation dans le cadre du projet *Vps15*, j'ai eu l'idée de mettre en culture les fibroblastes des souris et d'étudier leurs réponses/phénotypes en fonction de la condition de culture appliquée. Les résultats ont été obtenus sur un nombre réduit de souris (n=1). Pour la suite, il sera primordial d'augmenter l'effectif à un minimum de n=3.

## 2. Est-ce un modèle pour d'autres maladies plus communes ?

### A) Refaire et valider les conclusions

A l'issue des vagues d'expériences 2 et 3, nous avons montré que le modèle murin *Vps15* présente des plaques au niveau de l'hippocampe et du pancréas à des âges avancés (12 et 16 mois). Pour le moment, nous avons réalisé des expériences préliminaires qu'il faudra reproduire telles que les IF pour déterminer la nature des plaques hippocampiques, les WB dirigés contre les sérines phosphorylées de

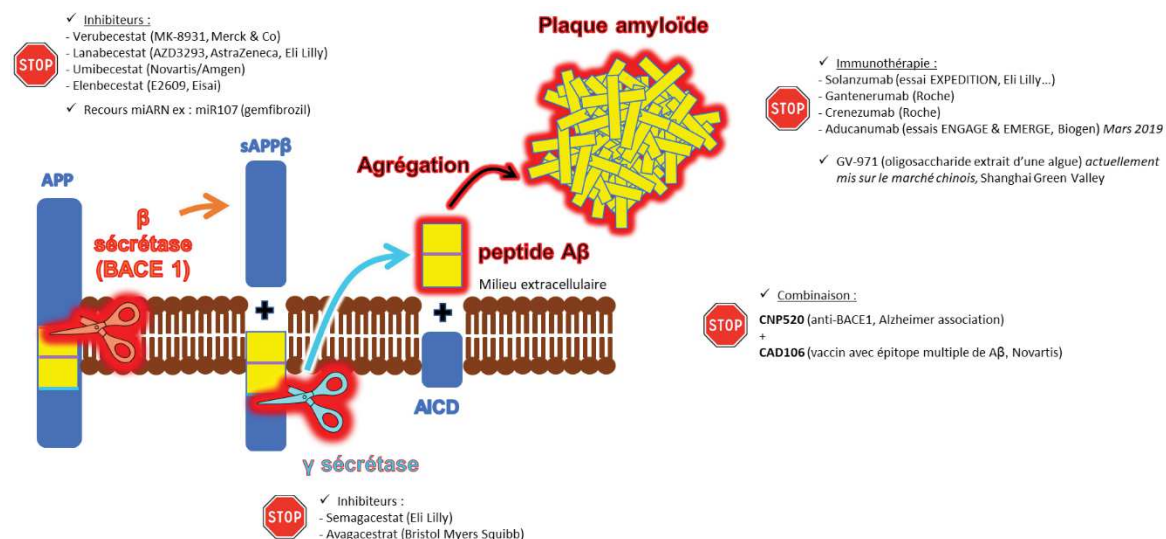
Tau ou encore les tests ELISA. Dans le cas des WB, la quantité protéique qui a été déposée dans les puits lors du dépôt est trop importante et le marquage obtenu lors de la révélation est trop fort pour permettre de distinguer des variations de quantités et de conclure. Il sera donc primordial de prendre en compte ces changements en refaisant cette expérience. Pour les ELISA, il faudra augmenter l'effectif de souris analysé et réaliser l'analyse sur plusieurs cohortes afin de pouvoir comparer les résultats selon les conditions. Globalement, pour la caractérisation portant sur le cerveau des souris KI *Vps15*, des expériences supplémentaires devront être mises en œuvre et les effectifs d'animaux inclus dans les cohortes devront être augmentés afin de déterminer si le modèle *Vps15* présente l'ensemble des critères de la maladie d'Alzheimer (mort neuronale, plaque amyloïde, hyperphosphorylation de la protéine Tau, troubles cognitifs). Pour confirmer la présence de morts cellulaires, nous pourrions envisager la mise en place d'un TUNEL (Terminal deoxynucleotidyl transferase dUTP nick end labeling) sur des cryocoupes de cerveaux afin de déterminer la proportion et la localisation des cellules en train de mourir par apoptose.

Si nous parvenons à confirmer l'ensemble de nos résultats, nous pourrions conclure que le modèle murin KI *Vps15* est un nouveau modèle d'étude pour la maladie d'Alzheimer. Ces conclusions s'avéreraient totalement inattendues au regard des premières hypothèses émises au début de mon travail de caractérisation phénotypique. Du coup, il sera également important de déterminer pourquoi la mutation dans le gène *Vps15* conduit au développement d'un tel phénotype, nous avons commencé ce travail au cours de ces derniers 6 mois en mettant en place la dernière vague d'expériences. Les résultats de Co-ip et de Full-moon devraient apporter un début de réponse et nous aiguiller pour la suite du projet.

#### B) Modèle d'étude pour la maladie d'Alzheimer

La maladie d'Alzheimer est considérée comme la première cause de démence chez les personnes âgées et concernera plus de 92 millions d'individus d'ici 2050 selon les estimations du *World Alzheimer report* édité par l'Alzheimer's disease international en 2015. Pour l'instant, les seuls traitements existants sont palliatifs. Par exemple, des inhibiteurs de la cholinestérase (AChEis) tel que le *Donepezil* ou des antagonistes aux récepteurs NMDA telle que la *Mémantine* sont prescrits aux patients afin de réduire les symptômes de la maladie. Les chercheurs travaillent sur des traitements ciblant plus précisément le processus physiopathologique qui débute bien avant l'arrivée des symptômes. Dans 22% des molécules testées, la cible appartient à la cascade de signalisation qui aboutit à la formation de la plaque amyloïde à partir du précurseur APP (P.-P. Liu et al. 2019). A l'heure actuelle, des inhibiteurs des enzymes  $\beta$  et  $\gamma$  sécrétases, des anticorps dirigés contre la plaque amyloïde ou les peptides A $\beta$  ont été développés et sont testés au cours de phases précliniques, comme le résume la Figure 77. Le test de ces molécules est réalisé en recourant à des modèles d'études pour la maladie d'Alzheimer tels que

le vers *C. elegans*, les organoïdes de mini-cerveaux, les poisson-zèbres, les primates non-humains ou encore les rongeurs. Pour ces derniers, 2 générations d'animaux ont été créées. La première génération de modèles murins générés correspond à des souris transgéniques surexprimant des protéines mutées dans les formes familiales d'Alzheimer, surexprimant le précurseur APP mutant ou surexprimant les enzymes  $\beta$ ,  $\gamma$  sécrétase ou Psen1. La deuxième génération regroupe des souris exprimant des séquences humanisées d'APP porteuses de mutations ponctuelles des formes familiales d'Alzheimer. Finalement, à l'issue des multiples croisements et des modifications de fond génétique, un total de 180 modèles murins est répertorié dans la base de données AlzForum. Chacune des lignées a été caractérisée phénotypiquement. Le succès des essais précliniques est contrebalancé par les échecs des études cliniques. Le choix du modèle pour le test des potentielles molécules thérapeutiques dépend de l'impact de la molécule d'intérêt car à l'heure actuelle, il n'existe pas de modèle disponible qui récapitule l'ensemble des critères de la maladie d'Alzheimer. Par exemple, une molécule ciblant la plaque amyloïde va être testée sur un modèle murin qui présente une plaque amyloïde mais qui ne présente pas d'hyperphosphorylation de la protéine Tau par exemple. Les modèles murins ne récapitulent pas non plus l'âge de développement tardif de la maladie observé chez les patients ni la distribution cérébrale des plaques. Il existe un véritable manque de corrélation entre les modèles murins et les patients atteints de cette maladie. L'enjeu majeur dans la recherche de traitement pour la maladie d'Alzheimer est de trouver un modèle d'étude pertinent. Les perspectives envisagées pour la suite de ce projet positionnent le modèle murin KI *Vps15* comme un potentiel candidat pouvant endosser ce rôle.



**Figure 77 : Molécules thérapeutiques anti-plaques.** Représentation schématisée repositionnant les molécules thérapeutiques testées (non exhaustif) en regard de leurs cibles dans la voie pathologique aboutissant à la formation de plaques amyloïdes à partir du précurseur APP (inspirée de Huang, Chao, et Hu 2020; Maia et Sousa 2019; Liu 2002).

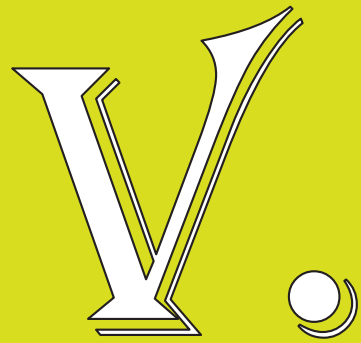
### C) Modèle d'étude pour les amyloïdoses

La maladie d'Alzheimer appartient à un groupe de maladies appelées amyloïdoses caractérisées par l'agrégation de protéines anormales sous la forme de plaques dans différents tissus (cœur, rein, pancréas, cerveau, etc.). Nous avons formulé l'hypothèse selon laquelle le modèle KI *Vps15* pourrait être un nouveau modèle d'étude de la maladie d'Alzheimer et peut être même des amyloïdoses en raison de la présence de plaques dans le pancréas. D'ailleurs, le diabète de type 2 peut être causé par une amyloïdose dans le cas où l'amyline (formée à partir d'un précurseur appelé pro-IAPP) s'agrège pour former des plaques (Lorenzo et al. 1994). Pour l'instant, nous avons une indication sur la composition des plaques présentes au niveau de l'hippocampe des souris KI grâce au marquage des peptides  $\beta$ -amyloïdes par IF. Il faudra faire un marquage similaire sur les cryocoupes de cerveaux issus de la cohorte 1 de femelles par exemple pour confirmer la nature des plaques. Pour le pancréas, nous pourrions également envisager un marquage anti-amyline afin de déterminer la composition des plaques dans cet organe. Pour confirmer l'atteinte pancréatique, nous pourrions mettre en place des tests fonctionnels pour explorer la présence d'un phénotype chez les souris tels qu'un Glucose tolerance test (GTT), des tests Elisa sur le plasma (anti-insuline, anti-glucagon ou anti-peptide C...). Nous souhaiterions, comme nous l'avons fait précédemment, réaliser une analyse de l'interactome de *Vps15* dans le pancréas.

Afin de déterminer si des plaques sont présentes dans d'autres tissus et si le modèle *Vps15* est un potentiel modèle d'étude des amyloïdoses, nous comptons élargir nos recherches à d'autres organes.

## IV. Les conclusions

Nous avons accumulé beaucoup de résultats concernant le nouveau modèle murin KI *Vps15<sup>R998Q/R998Q</sup>* au cours des 3 années de thèse. Nous pouvons conclure que les souris ne présentent pas le phénotype de ciliopathie attendu au début de ce projet. Par contre, certains résultats semblent indiquer une altération du processus d'autophagie : une conclusion similaire aux modèles KO et hypomorphe *Vps15* rapportés par d'autres équipes. L'analyse cérébrale mise en place au cours de la dernière année de thèse indique la présence de plaques dans l'hippocampe. L'exploration du pancréas a mis en évidence le même phénotype. Il faut maintenant confirmer les premiers résultats et mettre en place les expériences imaginées pour confirmer ou non nos hypothèses selon lesquelles le modèle murin KI *Vps15* serait un modèle d'étude pour la maladie d'Alzheimer ou pour les amyloïdoses plus largement. Ceci ouvrirait des perspectives importantes qui seront développées dans la discussion générale qui suit.



Discussion générale

Le travail de thèse que je vous présente dans ce manuscrit s'inscrit dans le vaste domaine de la génétique médicale appliquée aux maladies rares et aborde plus précisément certaines ciliopathies comme le BBS. Le travail que nous réalisons au laboratoire s'inscrit dans la recherche dite translationnelle qui se décrit en plusieurs étapes. Je vais les reprendre par la suite afin de montrer les principales limites et les possibles ouvertures pour chacune d'elles.

## I. Le diagnostic génétique

A ce jour, on dénombre environ 7000 maladies rares qui atteignent plus de 3 millions de personnes en France (soit 4,5% de la population française). Il est estimé que seulement 50% des malades ont un diagnostic sûr, c'est-à-dire que la mutation responsable de la maladie a été identifiée et validée. Dans le cadre du BBS, on estime que >80% des patients sont résolus. Il existe encore une grande part de patients concernés par l'errance diagnostique voire par l'impasse diagnostique. En effet, soit la cause génétique de la pathologie n'est pas encore identifiée, malgré toutes les techniques utilisées (NGS), soit une cause potentielle est identifiée, mais la preuve de la pathogénicité n'est pas clairement établie (les VUS). En France, l'errance diagnostique dure en moyenne 4 à 5 ans (CARE'19, Alliance maladies rares). L'objectif principal à atteindre à l'heure actuelle dans le cadre de ces maladies est de parvenir à 100% de taux de résolution, mais la question est la suivante : Comment y parvenir, comment pourrait-on résoudre ce qui est encore irrésolu ? Il s'agit d'ailleurs du principal objectif du Plan Prioritaire de Recherche sur les maladies rares lancé par le gouvernement pour une période de 10 ans dans la continuité du 3<sup>ème</sup> plan national maladies rares (2018-2025) et du plan France Médecine Génomique 2025. Une aide financière sera apportée dans les prochaines années aux laboratoires pour leur permettre de développer des approches innovantes afin de résoudre les impasses diagnostiques.

Un des points qu'il faut soulever à ce stade très préliminaire de la discussion est la nécessité de prendre en compte les variations des taux de résolution en fonction des cohortes considérées. En effet, il a été montré dans le cas du BBS que plus le nombre de critères cliniques majeurs est important, plus la probabilité de détecter les 2 allèles mutés dans un gène BBS connus augmente. Dans le cas où le diagnostic clinique établi est sûr, la résolution peut atteindre 100%. Par contre, si le diagnostic clinique est partagé parmi d'autres ciliopathies alors, le taux de résolution diminue (Redin et al. 2012c). Le problème c'est que de plus en plus de patients ayant des formes atypiques avec des phénotypes de plus en plus éloignés de ceux d'un BBS typique sont inclus dans les cohortes (Shamseldin et al. 2020b). La démarche scientifique pour la détection des mutations est un compromis entre le coût financier des analyses et la probabilité de détecter les variants pathogènes avec chacune des analyses. La démarche qui est habituellement mise en place se restreint au départ aux gènes connus pour être responsables des pathologies avec ce type de phénotypes (TES). Si aucun génotype compatible n'est mis en évidence

alors la recherche s'élargie au WES. Si le diagnostic génétique n'est toujours pas résolu alors une exploration du génome entier par WGS est mise en place. Ce type d'analyse va permettre la détection de milliers de mutations à grande échelle mais l'interprétation de cette masse de données s'avère complexe à l'heure actuelle. Si cela était envisageable d'un point de vue financier, l'analyse génétique serait directement faite par WGS sachant que l'analyse peut être virtuellement restreinte aux gènes connus (TES) ou à la séquence codante du génome (WES). Il s'agit d'une démarche récemment mise en place par Shamseldin et al. 2020 qui a directement fait l'analyse génétique par WGS pour des patients dont les signes cliniques n'étaient pas forcément compatibles avec le phénotype typique d'une ciliopathie. Ne serait-ce finalement pas le protocole à mettre en place à l'avenir pour répondre à la question, « comment résoudre l'irrésolu ? ». On peut d'ores et déjà répondre que oui car ce protocole correspond au plan France Médecine Génomique lancé depuis 2016 ayant pour objectif de multiplier par 5 le nombre de séquençages annuels en France par la mise en place de 12 plateformes de WGS. D'ailleurs, 2 plateformes de WGS sur les 12 prévues sont opérationnelles depuis cette année. Mais, il existe plusieurs autres hypothèses qui peuvent être avancées pour répondre à la question.

### 1. Les nouveaux gènes

D'abord, parmi les 20% des cas non résolus, il est possible qu'une partie des patients soient porteurs de mutations dans des gènes considérés comme de nouveaux gènes de ciliopathies au moment de l'analyse. Il peut s'agir de gènes candidats qui ont été mentionnés comme ayant un lien avec le cil lors d'études exploratoires expérimentales *in vitro* ou *in vivo*. Par exemple, la détermination du ciliome (Gherman, Davis, et Katsanis 2006) a permis la mise en évidence de nombreux gènes candidats. A l'époque, ce travail a donné naissance à une base de données nommée « Ciliproteome » qui avait pour objectif de partager avec la communauté scientifique les protéines ciliaires identifiées (et donc les gènes candidats potentiels) afin d'aider à la compréhension du cil primaire alors encore méconnu. De nombreux gènes ont d'ailleurs été validés par la suite comme de véritables gènes de ciliopathie. Cette première base de données a été maintenant remplacée par d'autres ayant les mêmes objectifs telles que Syscilia (<http://syscilia.org/>) ou encore CiDB (<http://cildb.cgm.cnrs-gif.fr/>). Des études d'analyses bioinformatiques « *in silico* » permettent également de déterminer des gènes candidats de ciliopathie. Par exemple, Nevers et al. 2017 est une publication du laboratoire qui rapporte une étude de génomique comparative ayant permis de déterminer une liste de 274 gènes ciliaires à partir de la conservation des gènes orthologues entre 100 espèces eucaryotes. Parmi eux, 87 gènes ont été considérés comme candidats.

Les gènes candidats de ciliopathies peuvent également être des gènes n'ayant aucun lien apparent avec une fonction ciliaire. Par exemple, la protéine LZTFL1 (Leucine Zipper Transcription Factor Like 1) a d'abord été décrite en 2010 comme un suppresseur de tumeur, considéré comme un marqueur



pronostic dans le cadre des cancers gastriques (Wei et al. 2019). Le rôle de cette protéine dans le cil a été décrit pour la première fois en 2011 (Seo et al. 2011). Ce gène a été rapporté comme étant le 17<sup>ème</sup> gène BBS par notre laboratoire en 2017 (Marion, Stutzmann, et al. 2012). Parfois, les gènes candidats peuvent être des gènes codants pour des protéines dont le rôle connu est totalement éloigné du fonctionnement du cil. Ceci est le cas pour *VPS15/PIK3R4* dont le rôle dans l'autophagie était clairement établi au préalable.

Néanmoins, il paraît relativement évident que la proportion de nouveaux gènes BBS restant à identifier soit assez faible, si on considère que depuis l'identification du gène *BBS13/MKS1*, le nombre de familles rapportées pour chacun des 11 derniers nouveaux gènes BBS est très restreint : il s'agit de contributeurs mineurs au BBS par rapport à des gènes tels que *BBS1* et *BBS10* par exemple.

## 2. Plusieurs gènes = plusieurs signes cliniques

Parmi les 20% des cas non résolus, il est possible que le phénotype de certains des patients atteints de ciliopathies syndromiques soit finalement le résultat de l'addition de plusieurs signes cliniques dans des organes différents, tous causés par des mutations dans des gènes différents. Nous avons été confrontés à ce type de conclusion dans le cadre du projet portant sur le gène *IQCE* (Estrada-Cuzcano et al. 2020b). Cette situation est loin d'être isolée comme 2 autres études le montrent. DeLuca et al. 2015 rapporte le cas d'un patient avec un phénotype proche de celui du syndrome de Usher dont la combinaison des signes cliniques est finalement causée par la mutation M390R de *BBS1* à l'état homozygote et des mutations hétérozygotes composites au niveau du gène *SLC26A4*. L'étude de Karaca et al. 2018 rapporte le cas d'une cohorte de 106 patients ayant des atteintes neurodéveloppementales dont l'analyse génétique a identifié, pour 13 d'entre eux, des variants pathogènes au niveau de 2 gènes ou plus, expliquant chacun une partie du phénotype.

## 3. Les gènes connus

### A) De nouveaux mécanismes

Parmi les 20% des cas non résolus, il est possible que les mutations soient présentes dans des gènes connus mais liées à un mécanisme qui n'est pas envisagé de manière systématique. Depuis le projet portant sur *BBS1* et l'insertion du rétrotransposon SVA-F, la PCR duplex ciblant le site d'insertion est mise en place en routine au laboratoire de diagnostic (HUS) et le protocole bioinformatique de détection des MEI a été ajouté à notre pipeline d'analyse des exomes et des génomes. Avec cette insertion dans l'exon 13 de *BBS1*, nous avons au final résolu le cas de 6 familles BBS supplémentaires et cet allèle représente à présent la 2<sup>ème</sup> mutation la plus fréquente dans le gène *BBS1*. En ajoutant la première famille publiée par Tavares et al. 2019, nous portons à 7 le nombre de familles dans le monde à avoir le même génotype.

Dans d'autres cas, la mutation peut échapper aux techniques d'identification, que ce soit en lien avec les techniques de biologies moléculaires ou bien avec les algorithmes bioinformatiques utilisés. Nous nous avons été confrontés à ce cas dans le projet portant sur *BBS5* pour lequel la délétion des 2 premiers exons n'a pu être détectée ni par le panel de gènes ni par l'exome. En effet, la délétion englobant le promoteur et les exons 1-2 du gène n'a été identifiée que par la mise en place d'un WGS. Une fois identifié, la pathogénicité de ce variant est évidente.

B) Qui s'occupe des VUS ?

Dans le cadre de l'identification et de la validation des mutations, il existe une frontière facilement franchissable entre le travail dévolu au laboratoire de diagnostic et celui dévolu au laboratoire de recherche. Théoriquement, le travail du laboratoire de recherche est d'identifier des mutations dans de nouveaux gènes et de publier ses recherches afin d'alimenter en nouveautés les connaissances de la communauté scientifique. Le travail du laboratoire de diagnostic génétique est d'identifier des mutations dans les gènes connus qui ont été rapportés précédemment par les laboratoires de recherche, en mettant en place du TES par exemple. Cependant, une part non négligeable des variants identifiés sont classés en tant que VUS (classe 3) nécessitant la mise en place d'analyses moléculaires plus poussées afin d'apporter la preuve ou pas de leur pathogénicité comme c'est le cas de la 2<sup>ème</sup> mutation identifiée dans *BBS5*. Le laboratoire de diagnostic n'a pas forcément les outils ou l'infrastructure pour mettre en œuvre ces expériences supplémentaires et transfère quand c'est possible ce travail à un laboratoire de recherche qui a plus de latitude dans la mise en place d'un large panel d'expériences afin de reclasser les variants. Bien que la coopération entre les laboratoires soit essentielle, il est évident que la validation de certains variants ne relève pas de la mission du laboratoire de recherche et devrait plutôt être à la charge du laboratoire de diagnostic.

C) La difficulté d'obtenir du matériel biologique

Dans le cadre du projet portant sur le gène *IFT27*, nous avons validé et publié en 2019 nos recherches portant sur des variants ayant déjà été rapportés par des publications de 2014 et 2018. Cette étude me permet d'illustrer la durée parfois nécessaire pour obtenir le matériel biologique nous permettant de poursuivre nos travaux de recherche. Il nous a fallu plus d'un an pour y avoir accès, ce qui a modifié notre stratégie de publication : nous sommes passés de la 2<sup>ème</sup> famille identifiée à la 4<sup>ème</sup> et notre article est devenu le 4<sup>ème</sup> à décrire des mutations dans ce gène. Bien que cela soit toujours très important de rapporter ces données dans la littérature, il est également important de considérer l'effort et la forte compétition du milieu de la recherche.

D) Le WGS

Je voudrais à présent revenir sur l'essor des techniques NGS et en particulier sur le recours du WGS pour l'identification des causes génétiques des ciliopathies. En effet, parmi les 20% des cas non résolus, on peut raisonnablement penser que de nombreuses mutations seront localisées dans des zones non codantes de notre génome comme : des variations introniques profondes, des variations des éléments régulateurs (pouvant altérer le niveau d'expression du gène) qui ne sont pas ciblées par le WES. Le recours au WGS se démocratise de plus en plus en génétique médicale en devenant plus accessible financièrement. Cette technique permet également de mettre en évidence des mutations non identifiées par le TES ou le WES. Par exemple, dans le cas du projet portant sur *BBS5*, le WGS a pu mettre en évidence une délétion des 2 premiers exons du gène *BBS5*, une zone qui était pourtant accessible par le TES et le WES. Le WGS ne nécessite pas d'étape de capture et la couverture des différentes régions (pas seulement des exons par exemple) est plus complète et homogène, ce qui offre une situation plus favorable à la détection des CNV. Le génome permet donc d'obtenir des données du WES mais de manière plus robuste.

L'un des problèmes majeurs qu'il va falloir résoudre avec l'essor du WGS est l'interprétation de l'ensemble des variants. A l'heure actuelle, il y a encore peu de données et donc d'annotations pour les zones du génome, qui n'étaient pas ou peu explorées jusqu'à présent. Ceci entraîne une plus grande difficulté pour classer les variants et laissera de nombreux variants en classe 3 (VUS). Ces lacunes devront être comblées afin de permettre le rendu d'un diagnostic pour des mutations dans des régions non codantes sans systématiquement nécessiter la mise en place de validations fonctionnelles poussées. On peut imaginer des mutations au niveau de régions régulatrices des gènes comme les « enhancers », des mutations au niveau de séquences de mi-ARN qui régulent les transcrits, des mutations introniques profondes qui perturberaient le niveau d'expression ou le processus d'épissage des gènes connus et causeraient une ciliopathie.

Une partie de ces limitations sera comblée par la généralisation du WGS et la mise à disposition des fréquences des variants dans les zones introniques et intergéniques comme le permet la base de données Genome Aggregation Database (GnomAD V3) qui référence plus de 70000 génomes d'individus indemnes de maladies pédiatriques (Karczewski et al. 2020).

L'utilisation du WGS devrait également permettre de déterminer avec précision le fond génétique de chacun des malades et d'appréhender les éventuels allèles modificateurs. Nikopoulos et al. 2019 rapporte l'étude d'une cohorte de patients japonais atteints de dégénérescence rétinienne héréditaire et aborde le cas particulier de la mutation non-sens p.Arg1933\* du gène *RP1*. Cette mutation est associée à un mode de transmission complexe : il s'agit d'une mutation fréquente au Japon (à 0,6%) qui n'est pas pathogène lorsqu'elle est présente à l'état hétérozygote ou homozygote. Par contre, elle

le devient quand elle est présente à l'état hétérozygote composite (en *trans* d'une autre mutation du même gène) ou en tant que 3<sup>ème</sup> allèle associé à de 2 mutations dans le gène *EYS*. Ces résultats montrent que la question de l'oligogénisme dans le cadre des ciliopathies est d'actualité. Cette hypothèse a été avancée par les articles Katsanis et al. 2001b et Badano et Katsanis 2002 dans le cadre du BBS par mais reste débattue au sein de notre laboratoire. Les articles Hichri et al. 2005 et Laurier et al. 2006 rapportent l'absence de preuve de triallélisme dans les cohortes étudiées. Par contre, dans le cadre du projet portant sur *BBS1*, nous avons rapporté le cas de 2 familles (issues de la cohorte américaine de N. Katsanis) pour lesquelles l'insertion du SVA-F correspond à un 3<sup>ème</sup> ou même un 4<sup>ème</sup> allèle. Le génome sera possiblement le point de départ le plus à même de répondre à ces questions.

## II. Les validations fonctionnelles

Dans le cadre de mon travail de thèse, j'ai eu la chance de pouvoir découvrir et réaliser de nombreuses techniques différentes pour répondre aux problématiques. Lorsqu'on regarde les logigrammes d'expériences pour la partie 1 du manuscrit portant sur l'étude moléculaire des gènes de ciliopathies, finalement, un schéma récurrent se dégage.

### 1. Les modèles cellulaires

#### A) Lequel choisir ?

Nous avons toujours utilisé au moins un des modèles d'études que je vous ai présenté dans l'introduction pour réaliser la validation des mutations. Les modèles cellulaires que nous utilisons sont des fibroblastes issus de biopsie de peau. Il s'agit de cellules facilement amplifiables permettant l'accès au matériel biologique essentiel à savoir l'ADNg, l'ARN et aux protéines. Elles peuvent être cultivées dans des conditions différentes en fonction de la question scientifique posée. Dans le cadre du BBS, les patients présentent des atteintes rétiniennes, rénales, cérébrales... On peut donc se demander si l'utilisation de fibroblastes de peau pour mettre en place la validation des mutations identifiées est le choix le plus adéquat. Il pourrait-être intéressant d'utiliser des modèles cellulaires de types organoïdes permettant de générer des mini-rétines, des mini-reins, des mini-cerveaux. Il s'agit de systèmes multicellulaires qui peuvent être créés par la différenciation de cellules souches induites issues la dédifférenciation de fibroblastes de patients ou de contrôles (Lou et Leung 2018). Les organoïdes sont pour l'instant ce qui se rapproche le plus des conditions *in vivo* pour un organe donné cultivé hors du corps. Ils récapitulent certains critères morphologiques, fonctionnels et transcriptomiques. Ils laissent ainsi envisager la mise en place d'expériences pouvant attester du bon fonctionnement ou d'une altération de l'organe *in vitro*. La combinaison de tels modèles avec des techniques de type « single cell transcriptomic » permettrait de comparer les profils d'expressions transcriptionnelles parmi les populations cellulaires dans des contextes physiologiques et pathologiques (Brazovskaja, Treutlein, et

Camp 2019). L'objectif serait d'identifier des signatures cellulaires pour chaque pathologie pouvant être considérées comme de futurs biomarqueurs qui faciliteraient le diagnostic moléculaire pour chacune d'elle.

Plusieurs publications rapportent des analyses fonctionnelles à partir de mini-cerveaux dans le cadre de la maladie d'Alzheimer. Par exemple, l'article de Gonzalez et al. 2018 montre que de tels organoïdes générés à partir de cellules de patients présentent des agrégats de peptides A $\beta$  et de protéines TAU phosphorylées. D'autres articles comme celui de Cowan et al. 2019 rapportent une étude transcriptomique comparative (analysant chaque type cellulaire par « single cell transcriptomic ») entre des organoïdes de rétines humaines *in vitro* et de véritables rétines humaines *in vivo* en montrant des profils d'expressions similaires. Ce sont des résultats qui viennent confirmer la pertinence des études utilisant de tels modèles et leurs applications dans l'étude des RP par exemple. Ces modèles apparaissent à l'heure actuelle comme une alternative intéressante à la culture de cellules en monocouche et aux modèles animaux. Il existe néanmoins des limites à ne pas négliger face à leur utilisation : la mise en place de tels modèles nécessite un investissement financier conséquent et il s'agit de mini-organes isolés qui n'interagissent pas avec les autres organes à l'inverse d'un organisme entier.

Dans le cadre du projet 2, j'avais débuté la culture des fibroblastes de souris pour mettre au point des IF et étudier le processus d'autophagie à une échelle plus réduite. Les premiers résultats sont prometteurs car ils montrent un niveau d'autophagie basal accru dans les cellules des souris KI. Dans ce cas, le recours aux fibroblastes permet la mise au point d'expériences pour étudier un processus ubiquitaire avant d'appliquer ces conditions à des cellules plus spécialisées par exemple. S'il est confirmé que les souris présentent effectivement des altérations au niveau cérébral et pancréatique, nous pourrions envisager d'extraire, maintenir en culture et réaliser des tests sur les neurones hippocampiques ou les îlots de Langerhans isolés par dissociation à partir des tissus murins (Beaudoin et al. 2012; Saliba et al. 2017). Nous pourrions ainsi étudier des types cellulaires spécialisés, sans que les résultats soient « dilués » par les autres types cellulaires constituant l'organe. Il s'agirait d'un pas supplémentaire pour décortiquer les mécanismes physiopathologiques qui relient la mutation R998Q aux phénotypes observés au niveau du cerveau et du pancréas.

#### B) Quelles conditions de culture ?

Dans la plupart des projets de la partie 1, les fibroblastes ont été étudiés en condition de ciliogénèse : un état où la cellule est pourvue d'un cil primaire. Dans cet état, les protéines impliquées dans le fonctionnement du cil sont présentes et une altération du cil causée par le dysfonctionnement d'une de ces protéines est visible. Dans le cadre du projet portant sur *IQCE*, le séquençage d'ARN a été réalisé sur des cellules cultivées en condition normale. Afin de mieux comprendre les relations et les

dysfonctionnements liés à des mutations du gène *IQCE*, cette analyse pourrait être réalisée en comparant les transcriptomes dans plusieurs conditions de culture telles que celles induisant la ciliogénèse seule (-SVF) et celles induisant la ciliogénèse avec l'activation de la voie SHH (-SVF+SAG). Pour aller plus loin dans cette idée, nous pourrions envisager à l'avenir de mettre en place des cinétiques en réalisant les mêmes expériences à intervalles de temps définis afin de distinguer le moment où d'éventuelles dérégulations apparaissent entre patients et contrôles. Il est possible de ne pas déceler de différences à la fin d'une expérience alors que l'évolution au cours du temps est en fait différente pour la croissance du cil par exemple (Lechtreck et al. 2017). Le cil primaire peut être finalement correctement formé bien que la mise en place de la ciliogénèse ait été retardée chez un patient atteint de ciliopathie. L'étude de Wen et al. 2010 portant sur la voie SHH montre comment évolue l'accumulation des protéines GLI2 et GLI3 à l'apex du cil après stimulation de la voie. La quantité de GLI3 accumulée atteint 80% dans les 30 à 60 minutes après l'activation, puis décline pour atteindre une phase plateau à ~50%. L'évolution est similaire pour GLI2. La mise en place de telles cinétiques nous permettrait de déterminer de potentielles variations vis-à-vis de cette courbe d'évolution telles que des décalages dans le temps (plus tôt ou plus tard) ou des quantités maximales d'accumulations moindres, dans les cellules des patients.

De la même manière, certaines lignées de fibroblastes de patients ont une croissance cellulaire plus lente ce qui peut révéler des anomalies ciliaires (Goto, Inaba, et Inagaki 2017). Il serait donc intéressant de suivre en temps réel la prolifération cellulaire afin de déterminer de manière systématique l'indice de prolifération des cellules et de le comparer en fonction des mutations, des gènes ou des ciliopathies pour voir si des correspondances se dégagent. Cette expérience simple pourrait être mise en place en routine au laboratoire (avec l'acquisition de technologies comme l'Incucyte, de Sartorius par exemple) et permettrait d'aiguiller le diagnostic moléculaire dès le début des projets qui utilisent des modèles cellulaires.

Afin de mettre en évidence un phénotype cellulaire de ciliopathie, nous marquons les cils primaires des cellules. Nous déterminons ensuite la fréquence de cellules ayant un cil et la taille de celui-ci comme cela a été réalisé dans le cadre du projet portant sur le gène *IQCE*. Dans le cas où il existe une variation significative entre le contrôle et le patient, nous pouvons seulement conclure que les cellules des patients présentent un phénotype de ciliopathie et confirmons le diagnostic clinique. Néanmoins, nous ne sommes pas capables de montrer que le phénotype cellulaire est causé par la mutation identifiée. De la même manière pour le projet portant sur le gène *BBS5*, dans l'état actuel des choses, nous avons montré une activation excessive de la voie SHH chez la patiente dans des conditions -SVF +SAG mais nous ne savons pas si c'est la mutation c.N184dup qui cause ce phénotype. Il est donc nécessaire de mettre en place des expériences de rescue/sauvetage de manière systématique au laboratoire pour établir une véritable relation de cause à effet entre la ou les mutations identifiées et

le phénotype cellulaire. Ainsi, il faudrait insérer une copie normale du gène par transfection d'un plasmide dans les cellules de patients et refaire les mêmes expériences. Pour être le plus complet possible, nous pourrions avoir la réflexion inverse en insérant la mutation au niveau du gène d'intérêt par des techniques d'édition comme CRISPR-Cas9 (Bassett 2017). Dans ce cas, si le phénotype cellulaire induit reproduit celui du patient alors nous pourrions conclure que la mutation identifiée cause effectivement le phénotype. Le problème dans ce type d'expérience reste le fond génétique qui est différent d'un individu à l'autre et provoque des modulations au niveau du phénotype. Pour déterminer à quel point le fond génétique peut moduler le phénotype, il pourrait être envisagé de réaliser en parallèle des expériences sur les cellules de patients vs contrôles, les mêmes expériences sur des lignées isogéniques dont le fond génétique est identique et ne diffère qu'au niveau du nucléotide où la mutation aura été introduite (Bassett 2017).

## 2. Le modèle du poisson-zèbre

Un autre modèle animal classiquement utilisé pour l'étude des ciliopathies est le poisson zèbre. Nous avons utilisé un tel modèle pour le projet portant sur *IQCE*. Il a été créé en recourant aux morpholinos (*lqce<sup>MO</sup>*). Il a déjà été montré que les phénotypes des poisson-zèbres créés par recours aux morpholinos en comparaison à ceux créés avec la technique CRISPR-Cas9 peuvent être assez différents alors que le même gène est ciblé (Stainier et al. 2017). Dans le premier cas, les morpholinos bloquent l'expression du gène cible de manière plus ou moins efficace en ciblant l'ARNm alors que dans le deuxième cas, l'invalidation est totale en ciblant la séquence génomique directement. Finalement, il aurait été intéressant de créer des poisson-zèbres *lqce<sup>-/-</sup>* avec la technologie CRISPR-Cas9 et de déterminer si un phénotype typique de ciliopathie syndromique était observé comme ce fût le cas avec le premier modèle *lqce<sup>MO</sup>*.

## 3. Les modèles murins

L'autre modèle animal que j'ai eu l'opportunité d'étudier au cours de ma thèse est le modèle murin KI *Vps15<sup>R998Q/R998Q</sup>*, pour lequel certaines expériences et conclusions méritent d'être discutées. J'aborderais également les expériences qui nous permettraient de poursuivre ce projet.

Pour commencer, pour plusieurs expériences nous avons analysés n=3 souris par génotype : un nombre qui est trop bas pour permettre d'établir une différence significative entre les souris mutantes et les contrôles. Pour conclure sur les résultats, il faudra refaire en incluant davantage d'animaux, au moins n=6.



A) Elargir l'exploration phénotypique des organes

Nous voulions déterminer si ce nouveau modèle murin était un bon modèle d'étude pour les ciliopathies mais nous devons malheureusement conclure négativement face aux résultats obtenus: il ne récapitule par les altérations habituellement observées dans le cadre des ciliopathies, pour les 3 organes analysés. En comparant le schéma de notre caractérisation phénotypique avec celui mis en place par les grands consortiums de générations de modèles animaux génétiquement modifiés, nous n'avons pas exploré l'ensemble des organes murins de manière systématique. Nous avons priorisé l'analyse des organes principalement atteints chez les patients porteurs de la même mutation puis nous avons cherché à élargir la caractérisation aux autres organes. L'étude histologique débutée va dans ce sens, elle porte sur une quinzaine d'organes et devrait nous aider à définir les patrons d'expressions de *Vps15* sous la forme d'un atlas histologique. Pour compléter ce travail et être le plus exhaustif possible, il faudrait poursuivre les études d'interactomiques déjà menées sur les 3 organes aux 12 autres afin de déterminer les interactants de *Vps15* et établir un différentiel entre les tissus. Il est possible qu'un différentiel, cette fois ci, d'interaction soit révélé dans certains organes en fonction du génotype des souris.

B) Elargir l'exploration phénotypique d'autres modèles murins

Comme je l'ai mentionné dans l'introduction, il existe d'autres modèles murins *Vps15* et en particulier des modèles KO. Chez l'Homme, seulement 2 familles porteuses de mutations faux-sens dans *VPS15* ont été rapportées dans le monde par Stoetzel et al. 2016 et Gstrein et al. 2018, avec des phénotypes différents. *Vps15* est une protéine ubiquitaire, essentielle au développement embryonnaire murin puisque les modèles KO total présentent une létalité embryonnaire. Ceci explique pourquoi ce sont des modèles conditionnels qui ont été étudiés jusqu'à présent. Il est fortement envisageable qu'une mutation entraînant la perte totale de *VPS15* chez l'homme soient également létale, expliquant l'absence de patient avec un tel génotype. Il serait très intéressant de générer un modèle KO conditionnel *Vps15* dans la rétine et dans le rein à partir du modèle KO de Nemazanyy et al. 2013 afin de déterminer si la perte de *Vps15* peut provoquer le développement d'une rétinopathie pigmentaire ou d'anomalies rénales. Ce sont des phénotypes normalement attendus pour un gène ayant un impact sur le fonctionnement ciliaire. Il serait également intéressant d'étudier les 3 organes d'intérêt à partir du modèle murin hypomorphe de Gstrein et al. 2018 et ainsi comparer les 3 modèles murins *Vps15* disponibles. Le principal problème auquel nous serions confrontés pour l'étude cérébrale serait l'âge d'analyse des animaux car les études ont été menées sur les modèles KO très tôt par rapport au modèle KI analysé à 16M.

Enfin, nous pourrions envisager de recourir à d'autres systèmes ciblant la dégradation de protéines d'intérêt directement telle que l'Auxin-inducible degron (AID) : une technologie transplantée des cellules de plantes à celles des mammifères comme l'Homme ou la souris. Cette technique permet une dégradation de la protéine d'intérêt en une vingtaine de minutes par l'ajout d'une phytohormone appelé auxine. Elle reconnaît un « tag » ajouté dans la séquence de la protéine qui sera polyubiquitinilée et dégradée via le protéasome. L'effet est réversible car l'activation du processus de dégradation s'arrête quand l'auxine disparaît du milieu. Natsume et al. 2016 décrit une technique couplant la technologie CRISPR-Cas 9 (pour l'ajout du tag dans la séquence du gène) au système AID et appliquée sur des cellules ES de souris. Face aux résultats concluants de cette étude, il est tout à fait envisageable qu'à l'avenir des modèles de souris KO inductibles soient générés à l'aide de cette technologie. D'ailleurs, l'hôpital Johns Hopkins de Baltimore a lancé un projet intitulé "Genetic engineering of mouse models and embryonic stem cell lines with auxin-inducible protein degradation system".

#### 4. Le modèle KI *Vps15<sup>R998Q/R998Q</sup>*

Pour les 2 paragraphes de discussion suivants, je n'ai pas souhaité trop extrapoler nos conclusions sachant qu'elles reposent sur des résultats qui doivent encore être confirmés. J'ai quand même souhaité les aborder brièvement afin de présenter les orientations possibles du projet.

En élargissant l'exploration phénotypique aux cerveaux des souris, nous nous sommes retrouvés face à des résultats de MET totalement inattendus avec la présence de plaques au niveau de l'hippocampe. Ces plaques semblent correspondre à des plaques amyloïdes ce qui nous a orienté vers la maladie d'Alzheimer. Nous avons poursuivi l'analyse cérébrale à la recherche d'autres critères de la maladie. La plupart de ces expériences exploratoires devront être reproduites. En effet, leurs résultats ne sont pas significatifs à cause d'un effectif trop réduit ou de conditions expérimentales inadéquates. Avant d'aller encore plus loin dans cette étude, nous devons assoir les conclusions en déterminant précisément la nature des agrégats protéiques : Est-ce que les plaques identifiées sont bien composées des peptides A $\beta$  1-40/1-42 ? Y a-t-il réellement une hyperphosphorylation de la protéine Tau et peut-on visualiser l'accumulation des DNF ? La neurodégénérescence mise en évidence par MET est-elle confirmée par l'application d'un TUNEL ? Si nous parvenons à répondre par l'affirmative à l'ensemble de ces questions, alors nous pourrions conclure que la souris KI *Vps15<sup>R998Q/R998Q</sup>* est un véritable modèle murin de la maladie d'Alzheimer et envisager d'autres expériences pour exploiter ce nouveau modèle. Nous pourrions par exemple réaliser une étude comparative entre ce nouveau modèle et ceux déjà utilisés dans les études de phase précliniques (X. Li, Bao, et Wang 2016). Certains des modèles murins de la maladie d'Alzheimer et les phénotypes associés ont été récapitulés dans le Tableau 9.

Type de modèle murin	Nom	Référence bibliographique	Mutation	Signes cliniques typiques de la maladie d'Alzheimer								Remarques	
				Plaque Aβ (+ âge d'apparition)	Tauopathie		Dépôt aβ autour des vaisseaux (CAA)	Neuroinflammation (gliose)	Atteintes neuronales		Perte de mémoire (+ âge d'apparition)		
					Hyperphosphorylation de Tau	DNF			Perte synaptique	Mort neuronale			
"single" transgenic mouse model	Promoteurs exogènes	PDAPP	Games et al. 1995	<i>hAPP<sup>D717F</sup></i>	☑ (6-8M)	/	/	☑	☑	/	☑	☑ (3M)	surexpression du transgène APP muté
		Pg2576	Hsiao et al. 1996	<i>hAPP<sup>K670N, M671L</sup></i>	☑ (9-11M)	/	/	☑	☑	☑	/	☑ (6M)	
		APP23	Sturchler-Pierrat et al. 1997	<i>hAPP<sup>K670N, M671L</sup></i>	☑ (6-8M)	☑	/	☑	☑	/	☑	☑ (3M)	
		J20	Mucke et al. 2000	<i>hAPP<sup>D717E, K670N, M671L</sup></i>	☑ (6-8M)	/	/	☑	☑	☑	☑	☑ (2M)	
		TgCRND8	Chishti et al. 2001	<i>hAPP<sup>D717E, K670N, M671L</sup></i>	☑ (3-5M)	/	/	☑	☑	☑	/	☑ (6M)	
		TgSweDI	Davis et al. 2004	<i>hAPP<sup>K670N, M671L, E693Q, D694N</sup></i>	☑ (3M)	/	/	☑	☑	/	/	☑ (6M)	
	APP693Δ	Tomiyama et al. 2008	<i>hAPP<sup>E693Δ</sup></i>	pas de plaque, peptides aβ seuls (8M)	☑	/	/	☑	/	☑	☑ (6M)		
	Promoteur endogène d'APP	APP N-L-G-F	Saito et al. 2014	<i>hAPP<sup>K670N, M671L, I276F, E693Q</sup></i>	☑ (2M)	/	/	/	☑	☑	/	☑ (6M)	expression normale : plus proche de la voie amyloïdogénique humaine
"double" transgenic mouse models	APPPS1-21	Radde et al. 2006	<i>hAPP<sup>K670N, M671L</sup> ; hPSEN1<sup>L166P</sup></i>	☑ (6-8M)	/	/	☑	☑	☑	/	☑ (7M)		
	APP/PS1	Jankowsky et al. 2004	<i>hAPP<sup>K670N, M671L</sup> ; hPSEN1<sup>Δex</sup></i>	☑ (6-9M)	/	/	☑	☑	☑	☑	☑ (6M)		
	5XFAD	Oakley et al. 2006	<i>hAPP<sup>K670N, M671L, I751V, V717I</sup> ; hPSEN1<sup>M146L, L286V</sup></i>	☑ (2M)	/	/	☑	☑	☑	☑	☑ (3M)		
"triple" transgenic mouse model	3XTg	Oddo et al. 2003	<i>hAPP<sup>K670N, M671L</sup> ; hPSEN1<sup>M146V</sup> ; hMAPT<sup>F302L</sup></i>	☑ (6M)	☑ (12M)	/	☑	/	/	☑ (4M)	le plus complet des modèles transgéniques d'Alzheimer, surexpression de multiples mutations		

**Tableau 9 :** Principaux modèles murins de la maladie d'Alzheimer. Annotation des âges d'apparition ☑ : présence, / : absence.

Pour établir ce tableau, je me suis basée sur la base de données « Alz-forum » qui répertorie les 180 modèles de rongeurs existants pour la maladie d'Alzheimer et sur ceux rapportés par Sasaguri et al. 2017 et Myers et McGonigle 2019. Il existe 3 notions à respecter concernant la définition d'un bon modèle animal pour une pathologie donnée. D'abord, la lignée doit présenter des signes cliniques similaires aux patients (phénotype). Ensuite, les mécanismes pathogéniques doivent être similaires entre les 2 espèces. Enfin, l'évolution des signes cliniques du modèle animal peut s'avérer prédictif et aider à la prise en charge des patients. La plupart des modèles murins existants sont porteurs de mutations rapportées pour les formes familiales de la maladie d'Alzheimer correspondant à une minorité de patients (1 à 2%) et les lignées murines sont porteuses de plusieurs mutations : une situation jamais retrouvée chez les patients. Les lignées de souris ne récapitulent que partiellement l'ensemble des critères histologiques présents chez les patients. Les phénotypes observés chez les animaux varient en fonction du fond génétique et du promoteur qui contrôle l'expression du transgène. Certaines lignées de souris (APP23, APP693Δ) présentent une hyperphosphorylation de la protéine Tau, sans DNF associés. Les animaux présentent généralement des anomalies comportementales entre 1 et 6 mois suivant l'apparition des premiers dépôts amyloïdes alors que l'augmentation des concentrations en peptides Aβ et l'apparition des premiers déficits cognitifs sont espacés de 10-15 ans chez les patients. Finalement, les 3 notions définissant un bon modèle animal d'une pathologie humaine ne sont pas respectées dans le cadre de la maladie d'Alzheimer. Ce manque de concordance explique d'ailleurs la réussite des phases pré-cliniques des molécules anti-Alzheimer testés sur des modèles pour lesquels le phénotype a été choisi en conséquence et l'échec des phases cliniques, lors de l'application aux patients. Il existe à l'heure actuelle un vrai manque concernant des modèles animaux de la maladie d'Alzheimer pertinents pouvant être utilisés comme véritable outil pour le test de molécules thérapeutiques prometteuses. Au stade actuel de nos investigations portant sur le modèle murin *Vps15*, nous pouvons envisager la possibilité que ce modèle comble ce manque si l'ensemble de nos hypothèses se confirment.

Pour le moment, les résultats de MET se sont révélés inattendus pour les 2 organes analysés : le cerveau et le pancréas. Dans ce dernier, des plaques ont également été identifiées. De la même manière que pour le cerveau, certaines questions sont restées en suspens pour le pancréas et plusieurs expériences devront être réalisées afin de préciser la composition protéique de ces plaques pancréatiques. Afin de déterminer si ce nouveau modèle murin est un modèle pour les amyloïdoses, il faudrait élargir la recherche de plaques aux autres organes de la souris et définir de la même manière la nature de ces plaques pour chacun d'eux.

Un vaste champ d'expériences pourrait donc être mis en place dans la continuité de mon travail de thèse si et seulement si les phénotypes cérébraux et pancréatiques sont confirmés.

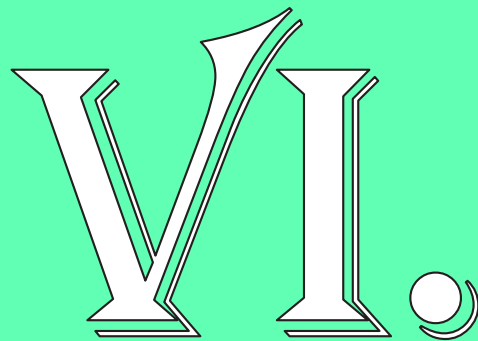
L'une des possibles explications à un tel phénotype chez un modèle murin porteur de la mutation R998Q dans *Vps15* est le lien existant entre le processus d'autophagie et la maladie d'Alzheimer. La relation de cause à effet existant entre ces 2 entités reste controversée et peut finalement être envisagée dans les 2 sens. Certains articles rapportent qu'un défaut d'autophagie conduit à l'apparition des critères de la maladie d'Alzheimer. A l'inverse, d'autres publications rapportent un défaut d'autophagie comme une conséquence de la maladie. La plupart de ces études utilisent des modèles murins porteurs de mutations dans des gènes dont les produits protéiques sont impliqués dans le processus d'autophagie (*Atg5*, *Atg9*, *Nrbf2*) ou dans la génèse de la maladie d'Alzheimer (*APP*).

L'accumulation de protéines telles que les peptides A $\beta$  ou Tau dans les modèles murins d'Alzheimer peut être causée par une dysfonction du processus de dégradation protéique autophagie-lysosome. Par exemple, les souris transgéniques *Nrbf2* KO présentent un défaut d'autophagie mis en évidence par l'accumulation de P62 et LC3II. Sur le plan comportemental, une altération de la mémoire spatiale a été relevée. Une accumulation de peptides A $\beta$ 1-42 a également été mise en évidence. La ré-expression de *Nrbf2* entraîne une réversion du phénotype observé et souligne l'impact du processus d'autophagie sur la pathogénèse de la maladie d'Alzheimer (Lachance et al. 2019). Dans la même idée, les modèles murins KO *Atg5* et *Atg7* présentent une accumulation intracellulaire de métabolites issue d'APP au niveau cérébral. D'ailleurs, une inhibition de la voie mTor entraîne une amélioration des déficits cognitifs ressemblant à ceux de la maladie d'alzheimer, une résorption de la plaque amyloïde et des DNF dans les modèles murins.

D'un autre côté, l'autophagie peut être une conséquence liée à la pathogénèse de la maladie d'Alzheimer. Les cerveaux de patients et de souris présentent des accumulations de figures d'autophagie, chez les souris, ses figures apparaissent bien avant de présenter des signes de neurodégénérences. D'ailleurs, ce type d'accumulation est maintenant reconnue comme une des potentielles marques de la maladie d'Azheimer. En cas de traitement du modèle TgCRND8 avec un inhibiteur de l'agrégation des peptides A $\beta$ 1-42, aucune anomalie du processus d'autophagie n'a été décelée. En revanche, le même modèle sans traitement présente une accumulation des figures d'autophagie (P.-P. Liu et al. 2019).

Le processus d'autophagie est impliqué dans la dégradation des peptides A $\beta$  : il est alors considéré comme non-amyloïdogénique. Il est également impliqué dans la production et la sécrétion de ces peptides et considéré, dans ce cas, comme amyloïdogénique. Finalement, la relation de cause à effet entre autophagie et Alzheimer est complexe, sûrement dépendante du contexte et possiblement régulée différemment selon le stade de la maladie considéré (précoce ou tardif) (Liu et Li 2019). En résumé, le processus physiopathologique expliquant le lien entre le génotype (mutation dans *Vps15*) et le phénotype observé (plaque hippocampique et pancréatique) doit être encore être déterminé. Les plaques sont finalement définies comme étant des agrégats de protéines dont la nature varie en

fonction du tissu d'intérêt et de la maladie considérée. Le modèle murin d'autophagie *Atg7*<sup>-/-</sup> présente des accumulations d'agrégats protéiques non dégradés, ubiquitinylés, qui mènent à une neurodégénérescence (Komatsu et al. 2006). Nous avons fait l'hypothèse que le phénotype de plaque observé est lié au rôle de *Vps15* dans le processus d'autophagie. La mutation R998Q de *Vps15* entrainerait une perturbation du processus d'autophagie conduisant ainsi à une accumulation des protéines normalement dégradés dans de multiples organes. Pour approfondir cette hypothèse, il faudrait réaliser des marquages dirigés contre P62, LC3B ou ubiquitine : des protéines liées aux produits protéiques dégradés lors du processus et dont l'accumulation est une preuve d'anomalie du processus d'autophagie. En obtenant des marquages au niveau hippocampique ou pancréatique dans les zones où les plaques ont été visualisées par MET, nous validerions l'implication du processus d'autophagie dans la formation de ces plaques. Il est possible que la souris KI *Vps15*, initialement considérée comme un nouveau modèle murin de ciliopathie soit en fait un nouveau modèle d'étude des maladies communes que sont la maladie d'Alzheimer et les amyloïdoses plus largement : des maladies caractérisées par une accumulation anormale de protéines. Nous serions donc face à un nouveau modèle murin à exploiter pour tester une multitude de molécules potentielles qui pourraient aboutir au traitement d'un grand nombre de malades.



Conclusions et

perspectives





La recherche translationnelle permet des échanges entre le patient, les professionnels de santé et les chercheurs pour améliorer la prise en charge, élargir les connaissances sur la maladie et déterminer les processus physiopathologiques sous-jacents. Au départ, les scientifiques reçoivent du patient : il autorise l'auscultation, donne des échantillons de peau et de sang. Ensuite, les cliniciens apposent un diagnostic clinique et les chercheurs mettent en place leurs logigrammes d'expériences dont les résultats viendront alimenter et étoffer le savoir de la communauté scientifique. Une fois les résultats scientifiques générés et interprétés par le chercheur, la relation s'inverse : c'est le patient qui reçoit de la part des scientifiques. Les chercheurs leur fournissent des réponses génétiques et moléculaires à leurs questions. Les cliniciens proposent une prise en charge adaptée à leur maladie. Mon travail de thèse fait partie intégrante de cette relation et a contribué à la résolution de 11 familles au total (1 pour *IFT27*, 6 pour *BBS1*, 3 pour *IQCE*, 1 pour *BBS5*).

Les travaux de biologie moléculaire et cellulaire que nous mettons en place au laboratoire peuvent aboutir (1) au développement ou au test de molécules thérapeutiques, (2) à l'identification de modèles animaux pertinents. Notre contribution intervient dans les premières étapes du long processus aboutissant à la mise sur le marché de traitements palliatifs ou curatifs. Dans le cadre de la poursuite de la partie 2, nous aurions plusieurs applications possibles. D'une part, si nous disposons effectivement d'un modèle murin de maladie d'Alzheimer, nous pourrions imaginer tester plusieurs molécules anti-Alzheimer déjà disponibles (ayant diverses cibles comme le résume Huang, Chao, et Hu 2020) et proposer notre modèle pour tester des molécules en cours de développement. D'autre part, si la souris KI *Vps15<sup>R998Q/R998Q</sup>* s'avérait être un bon modèle pour les amyloïdoses nous pourrions tester une panoplie de molécules anti-plaques dont les cibles diffèreraient pour chaque organe d'intérêt. Finalement, dans ce schéma idéal, nous montrons que la recherche sur les maladies ultra-rares peut avoir des répercussions majeures sur les maladies communes.

Les résultats des recherches réalisées sont transmis au patient par le généticien (après concertation avec le laboratoire de diagnostic génétique selon le cas). Ce retour permet au patient de pouvoir répondre à plusieurs questions au sujet de sa maladie.

D'abord, il peut répondre à « Quoi ? » : le patient connaît maintenant précisément la cause génétique responsable de la maladie dont il est atteint. Ensuite, il peut répondre à « Comment ? » : par exemple, dans le cadre du projet *BBS1*, nous savons que le rétrotransposon SVA-F s'est inséré au niveau de l'exon 13 de *BBS1* car un site d'insertion LINE-1 est présent dans la séquence. Sinon, le patient peut répondre à la question « Pourquoi ? ». Nous sommes parvenus, à l'exception du projet 4 portant sur *BBS5*, à établir une corrélation génotype-phénotype en déterminant le mécanisme physiopathologique qui explique pourquoi les mutations provoquent les signes cliniques d'une ciliopathie.

Ces résultats moléculaires ouvrent de nouvelles perspectives au patient.

Un conseil génétique peut être proposé au patient puis à sa famille (si le patient l'a demandé explicitement au généticien) afin de les informer sur la maladie et les conséquences de celle-ci. En cas de projet parental, les couples sont informés sur les probabilités de transmettre la maladie à leur descendance et la possibilité de recourir au diagnostic prénatal (DPN) ou préimplantatoire (DPI).

Le rendu d'un diagnostic univoque ouvre la porte vers de possibles thérapies ou essais cliniques auxquels le patient pourrait participer. Il s'agit de tests pour lesquels l'inclusion dépend du type de mutation, du gène concerné. Certaines thérapies géniques de l'œil ont débuté et peuvent être proposées aux patients qui remplissent les critères. Les études cliniques en cours sont récapitulées dans l'article Ducloyer et al. 2020. Par exemple une thérapie génique appelée Luxturna a obtenu une autorisation de mise sur le marché en 2018 par la Food and Drug administration (FDA) et la European Medicines Agency (EMA) en montrant son efficacité pour des patients atteints de LCA ou d'une RP causée par des mutations dans le gène *RP65*. En France, ce type de traitement est pris en charge par la sécurité sociale et certains patients ont pu en bénéficier.

La détermination du diagnostic génétique permet d'adapter la prise en charge du patient en fonction de sa ciliopathie. Par exemple, une surveillance cardiaque est mise en place pour les patients ALMS alors que les patient BBS n'en ont pas besoin. Dans le cadre de la partie 2, s'il s'avère que la mutation R998Q dans *Vps15* est responsable d'une amyloïdose telle que la maladie d'Alzheimer par exemple, il faudra mettre en place un suivi des 3 patients de la famille rapportée en 2016 pour déceler de possibles symptômes précoces, les informer des implications de tels résultats et adapter leur prise en charge en conséquence.

La recherche de demain s'oriente sans aucun doute vers la médecine personnalisée dont l'objectif est d'adapter le traitement/la prise en charge des patients en fonction de leurs profils biologiques. La médecine personnalisée n'est pas une idée nouvelle et limitée à la génétique, elle se perfectionne avec le temps et avec la quantité de données disponibles pour chaque patient. Pour l'instant, le fait de séquencer le génome complet de chaque patient de manière systématique aidera à déterminer un diagnostic génétique robuste et à mettre en place une prise en charge adaptée pour chacun. Intégrer le WGS en routine dans la pratique clinique est l'idée fondatrice du Plan Médecine génomique 2025. Une des mesures phares annonce la création de plateformes supplémentaires de séquençage génétique à très haut débit pour passer de 40 000 à 200 000 séquençages annuels en France. Les 2 premières plateformes sélectionnées en juillet 2017 (AURAGEN à Lyon et SeqOIA à Paris) sont opérationnelles depuis 2020 et les premiers génomes ont été séquencés. Les mesures sont en cours de discussions concernant la prise en charge de ces examens par la sécurité sociale. L'ancienne ministre Marisol Touraine a résumé le 22 juin 2016 les objectifs du gouvernement français qui seraient de promouvoir « une médecine personnalisée, adaptée au capital humain de chacun ».

VII.

Références

- Afzelius, B. 1959. « Electron Microscopy of the Sperm Tail; Results Obtained with a New Fixative ». *The Journal of Biophysical and Biochemical Cytology* 5 (2): 269-78. <https://doi.org/10.1083/jcb.5.2.269>.
- Afzelius, B. A. 1976. « A Human Syndrome Caused by Immotile Cilia ». *Science (New York, N.Y.)* 193 (4250): 317-19. <https://doi.org/10.1126/science.1084576>.
- Akman, Hasan O., Guido Davidzon, Kurenai Tanji, Emma J. MacDermott, Louann Larsen, Mercy M. Davidson, Ronald G. Haller, et al. 2010a. « Neutral Lipid Storage Disease with Subclinical Myopathy Due to a Retrotransposal Insertion in the PNPLA2 Gene ». *Neuromuscular Disorders* 20 (6): 397-402. <https://doi.org/10.1016/j.nmd.2010.04.004>.
- Aldahmesh, Mohammed A., Yuanyuan Li, Amal Alhashem, Shams Anazi, Hisham Alkuraya, Mais Hashem, Ali A. Awaji, et al. 2014. « IFT27, encoding a small GTPase component of IFT particles, is mutated in a consanguineous family with Bardet–Biedl syndrome ». *Human Molecular Genetics* 23 (12): 3307-15. <https://doi.org/10.1093/hmg/ddu044>.
- Al-Hamed, Mohamed H., Charles van Lennep, Ann Marie Hynes, Paul Chrystal, Lorraine Eley, Fatimah Al-Fadhly, Riham El Sayed, Roslyn J. Simms, Brian Meyer, et John A. Sayer. 2014. « Functional Modelling of a Novel Mutation in BBS5 ». *Cilia* 3 (1): 3. <https://doi.org/10.1186/2046-2530-3-3>.
- Alonso, Alejandra D., Leah S. Cohen, Christopher Corbo, Viktoriya Morozova, Abdeslem Elidrissi, Greg Phillips, et Frida E. Kleiman. 2018. « Hyperphosphorylation of Tau Associates With Changes in Its Function Beyond Microtubule Stability ». *Frontiers in Cellular Neuroscience* 12 (octobre). <https://doi.org/10.3389/fncel.2018.00338>.
- Álvarez-Satta, María, Sheila Castro-Sánchez, et Diana Valverde. 2017. « Bardet-Biedl Syndrome as a Chaperonopathy: Dissecting the Major Role of Chaperonin-Like BBS Proteins (BBS6-BBS10-BBS12) ». *Frontiers in Molecular Biosciences* 4. <https://doi.org/10.3389/fmolb.2017.00055>.
- Athanasiou, Dimitra, Monica Aguilà, Dalila Bevilacqua, Sergey S. Novoselov, David A. Parfitt, et Michael E. Cheetham. 2013. « THE CELL STRESS MACHINERY AND RETINAL DEGENERATION ». *FEBS letters* 587 (13): 2008-17. <https://doi.org/10.1016/j.febslet.2013.05.020>.
- Auton, Adam, Gonçalo R. Abecasis, David M. Altshuler, Richard M. Durbin, Gonçalo R. Abecasis, David R. Bentley, Aravinda Chakravarti, et al. 2015. « A global reference for human genetic variation ». *Nature* 526 (7571): 68-74. <https://doi.org/10.1038/nature15393>.
- Ávalos, Yennifer, Daniel Peña-Oyarzun, Mauricio Budini, Eugenia Morselli, et Alfredo Criollo. 2017. « New Roles of the Primary Cilium in Autophagy ». Research article. BioMed Research International. 2017. <https://doi.org/10.1155/2017/4367019>.
- Bachmann-Gagescu, Ruxandra. 2014. « Complexité génétique des ciliopathies et identification de nouveaux gènes ». *médecine/sciences* 30 (11): 1011-23. <https://doi.org/10.1051/medsci/20143011016>.
- Bachurin, S. O., E. V. Bovina, et A. A. Ustyugov. 2018. « Current Trends in the Development of Drugs for the Treatment of Alzheimer’s Disease and Their Clinical Trials ». *Biomedical Chemistry: Research and Methods* 1 (3): e00015-e00015. <https://doi.org/10.18097/BMCRM00015>.
- Backenroth, Daniel, Jason Homsy, Laura R. Murillo, Joe Glessner, Edwin Lin, Martina Brueckner, Richard Lifton, Elizabeth Goldmuntz, Wendy K. Chung, et Yufeng Shen. 2014. « CANOES: Detecting Rare Copy Number Variants from Whole Exome Sequencing Data ». *Nucleic Acids Research* 42 (12): e97. <https://doi.org/10.1093/nar/gku345>.

- Backer, Jonathan M. 2016. « The Intricate Regulation and Complex Functions of the Class III Phosphoinositide 3-Kinase Vps34 ». *Biochemical Journal* 473 (15): 2251-71. <https://doi.org/10.1042/BCJ20160170>.
- Badano, Jose L., et Nicholas Katsanis. 2002. « Beyond Mendel: An Evolving View of Human Genetic Disease Transmission ». *Nature Reviews Genetics* 3 (10): 779-89. <https://doi.org/10.1038/nrg910>.
- Badano, Jose L., Jun Chul Kim, Bethan E. Hoskins, Richard Alan Lewis, Stephen J. Ansley, David J. Cutler, Claudio Castellani, Philip L. Beales, Michel R. Leroux, et Nicholas Katsanis. 2003. « Heterozygous Mutations in BBS1, BBS2 and BBS6 Have a Potential Epistatic Effect on Bardet–Biedl Patients with Two Mutations at a Second BBS Locus ». *Human Molecular Genetics* 12 (14): 1651-59. <https://doi.org/10.1093/hmg/ddg188>.
- Barbaric, I., G. Miller, et T. N. Dear. 2007. « Appearances Can Be Deceiving: Phenotypes of Knockout Mice ». *Briefings in Functional Genomics and Proteomics* 6 (2): 91-103. <https://doi.org/10.1093/bfgp/elm008>.
- BARDET, G. 1920. « Sur un syndrome d'obésité congénitale avec polydactylie et rétinite pigmentaire (contribution à l'étude des formes cliniques de l'obésité hypophysaire) ». *Amedee Le Grand* 470. <https://ci.nii.ac.jp/naid/20001129230/>.
- Bassett, Andrew R. 2017. « Editing the Genome of hiPSC with CRISPR/Cas9: Disease Models ». *Mammalian Genome* 28 (7): 348-64. <https://doi.org/10.1007/s00335-017-9684-9>.
- Beales, P. L., N. Elcioglu, A. S. Woolf, D. Parker, et F. A. Flintner. 1999a. « New Criteria for Improved Diagnosis of Bardet-Biedl Syndrome: Results of a Population Survey ». *Journal of Medical Genetics* 36 (6): 437-46.
- Beales, Philip L., Jose L. Badano, Alison J. Ross, Stephen J. Ansley, Bethan E. Hoskins, Brigitta Kirsten, Charles A. Mein, et al. 2003. « Genetic Interaction of BBS1 Mutations with Alleles at Other BBS Loci Can Result in Non-Mendelian Bardet-Biedl Syndrome ». *The American Journal of Human Genetics* 72 (5): 1187-99. <https://doi.org/10.1086/375178>.
- Beaudoin, Gerard M. J., Seung-Hye Lee, Dipika Singh, Yang Yuan, Yu-Gie Ng, Louis F. Reichardt, et Jyothi Arikath. 2012. « Culturing Pyramidal Neurons from the Early Postnatal Mouse Hippocampus and Cortex ». *Nature Protocols* 7 (9): 1741-54. <https://doi.org/10.1038/nprot.2012.099>.
- BIEDL, A. 1922. « Ein Geschwister mit adiposogenitaler Dystrophie ». *Med Wochenschr* 48: 1630.
- Bird, Thomas D. 1993. « Alzheimer Disease Overview ». In *GeneReviews*<sup>®</sup>, édité par Margaret P. Adam, Holly H. Ardinger, Roberta A. Pagon, Stephanie E. Wallace, Lora JH Bean, Karen Stephens, et Anne Amemiya. Seattle (WA): University of Washington, Seattle. <http://www.ncbi.nlm.nih.gov/books/NBK1161/>.
- Blackshaw, Seth, Sanjiv Harpavat, Jeff Trimarchi, Li Cai, Haiyan Huang, Winston P. Kuo, Griffin Weber, et al. 2004. « Genomic Analysis of Mouse Retinal Development ». *PLoS Biology* 2 (9): E247. <https://doi.org/10.1371/journal.pbio.0020247>.
- Bloodgood, Robert A. 2009. « From Central to Rudimentary to Primary: The History of an Underappreciated Organelle Whose Time Has Come. The Primary Cilium ». *Methods in Cell Biology* 94: 3-52. [https://doi.org/10.1016/S0091-679X\(08\)94001-2](https://doi.org/10.1016/S0091-679X(08)94001-2).

- Bragg, D. Christopher, Kotchaphorn Mangkalaphiban, Christine A. Vaine, Nichita J. Kulkarni, David Shin, Rachita Yadav, Jyotsna Dhakal, et al. 2017. « Disease Onset in X-Linked Dystonia-Parkinsonism Correlates with Expansion of a Hexameric Repeat within an SVA Retrotransposon in *TAF1* ». *Proceedings of the National Academy of Sciences* 114 (51): E11020-28. <https://doi.org/10.1073/pnas.1712526114>.
- Brazovskaja, Agnieska, Barbara Treutlein, et J Gray Camp. 2019. « High-Throughput Single-Cell Transcriptomics on Organoids ». *Current Opinion in Biotechnology* 55 (février): 167-71. <https://doi.org/10.1016/j.cobio.2018.11.002>.
- Brett, Maggie, George Korovesis, Angeline H. M. Lai, Eileen C. P. Lim, et Ene-Choo Tan. 2017. « Intragenic Multi-Exon Deletion in the *FBN1* Gene in a Child with Mildly Dilated Aortic Sinus: A Retrotransposal Event ». *Journal of Human Genetics* 62 (7): 711-15. <https://doi.org/10.1038/jhg.2017.32>.
- Brown, Jason M., et George B. Witman. 2014. « Cilia and Diseases ». *BioScience* 64 (12): 1126-37. <https://doi.org/10.1093/biosci/biu174>.
- Brun, Agnès, Xiangxiang Yu, Cathy Obringer, Daniel Ajoy, Elodie Haser, Corinne Stoetzel, Michel J Roux, Nadia Messaddeq, Hélène Dollfus, et Vincent Marion. 2019. « In vivo phenotypic and molecular characterization of retinal degeneration in mouse models of three ciliopathies ». *Experimental Eye Research* 186 (septembre): 107721. <https://doi.org/10.1016/j.exer.2019.107721>.
- Brunner, Clément. 2016. « FUNCTIONAL ULTRASOUND IMAGING (fUSi) TO ASSESS BRAIN FUNCTION IN PHYSIOLOGICAL AND PATHOLOGICAL CONDITIONS Application to stroke ».
- Buckley, Reuben M., R. Daniel Kortschak, Joy M. Raison, et David L. Adelson. 2017. « Similar Evolutionary Trajectories for Retrotransposon Accumulation in Mammals ». *Genome Biology and Evolution* 9 (9): 2336-53. <https://doi.org/10.1093/gbe/evx179>.
- Burns, Kathleen H. 2020. « Our Conflict with Transposable Elements and Its Implications for Human Disease ». *Annual Review of Pathology: Mechanisms of Disease* 15 (1): 51-70. <https://doi.org/10.1146/annurev-pathmechdis-012419-032633>.
- Byeon, Boseon, et Igor Kovalchuk. 2016. « Non-Coding RNAs Match the Deleted Genomic Regions in Humans ». *Scientific Reports* 6 (1). <https://doi.org/10.1038/srep37452>.
- Cai, Zhiyou, Yingjun Zhou, Zhou Liu, Zunyu Ke, et Bin Zhao. 2015. « Autophagy dysfunction upregulates beta-amyloid peptides via enhancing the activity of  $\gamma$ -secretase complex ». *Neuropsychiatric Disease and Treatment* 11 (août): 2091-99. <https://doi.org/10.2147/NDT.S84755>.
- Caligioni, Claudia S. 2009. « Assessing Reproductive Status/Stages in Mice ». *Current Protocols in Neuroscience* 48 (1). <https://doi.org/10.1002/0471142301.nsa04is48>.
- Chang, Christopher C, Carson C Chow, Laurent CAM Tellier, Shashaank Vattikuti, Shaun M Purcell, et James J Lee. 2015. « Second-generation PLINK: rising to the challenge of larger and richer datasets ». *GigaScience* 4 (s13742-015-0047-8). <https://doi.org/10.1186/s13742-015-0047-8>.
- Chatron, Nicolas, Kevin Cassinari, Olivier Quenez, Stéphanie Baert-Desurmont, Claire Bardel, Marie-Pierre Buisine, Eduardo Calpena, et al. 2019. « Identification of mobile retrocopies during genetic testing: Consequences for routine diagnosis ». *Human Mutation* 40 (11): 1993-2000. <https://doi.org/10.1002/humu.23845>.



- Chen, Jian-Min, Emmanuelle Masson, Milan Macek, Odile Raguénès, Tereza Piskackova, Brigitte Fercot, Libor Fila, David N. Cooper, Marie-Pierre Audrézet, et Claude Férec. 2008. « Detection of Two Alu Insertions in the CFTR Gene ». *Journal of Cystic Fibrosis* 7 (1): 37-43. <https://doi.org/10.1016/j.jcf.2007.04.001>.
- Chen, Vivian S., James P. Morrison, Myra F. Southwell, Julie F. Foley, Brad Bolon, et Susan A. Elmore. 2017. « Histology Atlas of the Developing Prenatal and Postnatal Mouse Central Nervous System, with Emphasis on Prenatal Days E7.5 to E18.5 ». *Toxicologic pathology* 45 (6): 705-44. <https://doi.org/10.1177/0192623317728134>.
- Cherra, Salvatore J., Scott M. Kulich, Guy Uechi, Manimalha Balasubramani, John Mountzouris, Billy W. Day, et Charleen T. Chu. 2010. « Regulation of the Autophagy Protein LC3 by Phosphorylation ». *Journal of Cell Biology* 190 (4): 533-39. <https://doi.org/10.1083/jcb.201002108>.
- Chicher, Johana, Angelita Simonetti, Lauriane Kuhn, Laure Schaeffer, Philippe Hammann, Gilbert Eriani, et Franck Martin. 2015. « Purification of mRNA-Programmed Translation Initiation Complexes Suitable for Mass Spectrometry Analysis ». *PROTEOMICS* 15 (14): 2417-25. <https://doi.org/10.1002/pmic.201400628>.
- Chicois, Clara, Hélène Scheer, Shahínez Garcia, Hélène Zuber, Jérôme Mutterer, Johana Chicher, Philippe Hammann, Dominique Gagliardi, et Damien Garcia. 2018. « The UPF1 Interactome Reveals Interaction Networks between RNA Degradation and Translation Repression Factors in Arabidopsis ». *The Plant Journal* 96 (1): 119-32. <https://doi.org/10.1111/tpj.14022>.
- Chong, Jessica X., Kati J. Buckingham, Shalini N. Jhangiani, Corinne Boehm, Nara Sobreira, Joshua D. Smith, Tanya M. Harrell, et al. 2015. « The Genetic Basis of Mendelian Phenotypes: Discoveries, Challenges, and Opportunities ». *American Journal of Human Genetics* 97 (2): 199-215. <https://doi.org/10.1016/j.ajhg.2015.06.009>.
- Chou, Hui-Ting, Luise Apelt, Daniel P. Farrell, Susan Roehl White, Jonathan Woodsmith, Vladimir Svetlov, Jaclyn S. Goldstein, et al. 2019a. « The Molecular Architecture of Native BBSome Obtained by an Integrated Structural Approach ». *Structure* 27 (9): 1384-1394.e4. <https://doi.org/10.1016/j.str.2019.06.006>.
- Clemen, Christoph S., Lilli Winter, Karl-Heinz Strucksberg, Carolin Berwanger, Matthias Türk, Cornelia Kornblum, Alexandra Florin, et al. 2018. « The Heterozygous R155C VCP Mutation: Toxic in Humans! Harmless in Mice? ». *Biochemical and Biophysical Research Communications* 503 (4): 2770-77. <https://doi.org/10.1016/j.bbrc.2018.08.038>.
- Collins, Ryan L., Harrison Brand, Konrad J. Karczewski, Xuefang Zhao, Jessica Alföldi, Laurent C. Francioli, Amit V. Khera, et al. 2020. « A structural variation reference for medical and population genetics ». *Nature* 581 (7809): 444-51. <https://doi.org/10.1038/s41586-020-2287-8>.
- Collins, Stephan C., Christel Wagner, Léo Gagliardi, Perrine F. Kretz, Marie-Christine Fischer, Pascal Kessler, Meghna Kannan, et Binnaz Yalcin. 2018. « A Method for Parasagittal Sectioning for Neuroanatomical Quantification of Brain Structures in the Adult Mouse ». *Current Protocols in Mouse Biology* 8 (3): e48. <https://doi.org/10.1002/cpmo.48>.
- Conley, Mary Ellen, Julie D. Partain, Shannon M. Norland, Sheila A. Shurtleff, et Haig H. Kazazian. 2005. « Two Independent Retrotransposon Insertions at the Same Site within the Coding Region of BTK ». *Human Mutation* 25 (3): 324-25. <https://doi.org/10.1002/humu.9321>.

- Craige, Branch, Che-Chia Tsao, Dennis R. Diener, Yuqing Hou, Karl-Ferdinand Lechtreck, Joel L. Rosenbaum, et George B. Witman. 2010. « CEP290 Tethers Flagellar Transition Zone Microtubules to the Membrane and Regulates Flagellar Protein Content ». *The Journal of Cell Biology* 190 (5): 927-40. <https://doi.org/10.1083/jcb.201006105>.
- Criollo, A., M. C. Maiuri, E. Tasdemir, I. Vitale, A. A. Fiebig, D. Andrews, J. Molgó, et al. 2007. « Regulation of Autophagy by the Inositol Trisphosphate Receptor ». *Cell Death & Differentiation* 14 (5): 1029-39. <https://doi.org/10.1038/sj.cdd.4402099>.
- Crivelli, Louise, Virginie Bubien, Natalie Jones, Jennifer Chiron, Françoise Bonnet, Emmanuelle Barouk-Simonet, Patrice Couzigou, Nicolas Sevenet, Frédéric Caux, et Michel Longy. 2017. « Insertion of Alu Elements at a PTEN Hotspot in Cowden Syndrome ». *European Journal of Human Genetics : EJHG* 25 (9): 1087-91. <https://doi.org/10.1038/ejhg.2017.81>.
- Cryan, John F., et Andrew M. Holmes. 2005. « Model organisms: The ascent of mouse: advances in modelling human depression and anxiety ». *Nature Reviews Drug Discovery*. <https://doi.org/10.1038/nrd1825>.
- Cullen-McEwen, Luise, Megan R. Sutherland, et M. Jane Black. 2016. « Chapter 3 - The Human Kidney: Parallels in Structure, Spatial Development, and Timing of Nephrogenesis ». In *Kidney Development, Disease, Repair and Regeneration*, édité par Melissa H. Little, 27-40. San Diego: Academic Press. <https://doi.org/10.1016/B978-0-12-800102-8.00003-5>.
- Cummings, Beryl B., Jamie L. Marshall, Taru Tukiainen, Monkol Lek, Sandra Donkervoort, A. Reghan Foley, Veronique Bolduc, et al. 2017. « Improving genetic diagnosis in Mendelian disease with transcriptome sequencing ». *Science Translational Medicine* 9 (386): eaal5209. <https://doi.org/10.1126/scitranslmed.aal5209>.
- Davis, Roger E., Ruth E. Swiderski, Kamal Rahmouni, Darryl Y. Nishimura, Robert F. Mullins, Khristofor Agassandian, Alisdair R. Philp, et al. 2007. « A knockin mouse model of the Bardet–Biedl syndrome 1 M390R mutation has cilia defects, ventriculomegaly, retinopathy, and obesity ». *Proceedings of the National Academy of Sciences of the United States of America* 104 (49): 19422-27. <https://doi.org/10.1073/pnas.0708571104>.
- DeLuca, Adam P., Matthew C. Weed, Christine M. Haas, Jennifer A. Halder, et Edwin M. Stone. 2015. « Apparent Usher Syndrome Caused by the Combination of BBS1-Associated Retinitis Pigmentosa and SLC26A4-Associated Deafness ». *JAMA Ophthalmology* 133 (8): 967-68. <https://doi.org/10.1001/jamaophthalmol.2015.1463>.
- DePristo, Mark A, Eric Banks, Ryan Poplin, Kiran V Garimella, Jared R Maguire, Christopher Hartl, Anthony A Philippakis, et al. 2011. « A framework for variation discovery and genotyping using next-generation DNA sequencing data ». *Nature Genetics* 43 (5): 491-98. <https://doi.org/10.1038/ng.806>.
- Dickinson, Mary E., Ann M. Flenniken, Xiao Ji, Lydia Teboul, Michael D. Wong, Jacqueline K. White, Terrence F. Meehan, et al. 2016. « High-Throughput Discovery of Novel Developmental Phenotypes ». *Nature* 537 (7621): 508-14. <https://doi.org/10.1038/nature19356>.
- Dorn, Karolin V., Casey E. Hughes, et Rajat Rohatgi. 2012. « A Smoothed-Evc2 Complex Transduces the Hedgehog Signal at Primary Cilia ». *Developmental Cell* 23 (4): 823-35. <https://doi.org/10.1016/j.devcel.2012.07.004>.
- Ducloyer, Jean-Baptiste, Guylène Le Meur, Thérèse Cronin, Oumeya Adjali, et Michel Weber. 2020. « La thérapie génique des rétinites pigmentaires héréditaires ». *médecine/sciences* 36 (6-7): 607-15. <https://doi.org/10.1051/medsci/2020095>.

- Eguether, Thibaut, Jovenal T. San Agustin, Brian T. Keady, Julie A. Jonassen, Yinwen Liang, Richard Francis, Kimimasa Tobita, et al. 2014a. « IFT27 Links the BBSome to IFT for Maintenance of the Ciliary Signaling Compartment ». *Developmental Cell* 31 (3): 279-90. <https://doi.org/10.1016/j.devcel.2014.09.011>.
- El-Amraoui, Aziz, et Christine Petit. 2014. « The Retinal Phenotype of Usher Syndrome: Pathophysiological Insights from Animal Models ». *Comptes Rendus Biologies, Spotlight on vision*, 337 (3): 167-77. <https://doi.org/10.1016/j.crv.2013.12.004>.
- Estrada-Cuzcano, Alejandro, Christelle Etard, Clarisse Delvallée, Corinne Stoetzel, Elise Schaefer, Sophie Scheidecker, Véronique Geoffroy, et al. 2020a. « Novel IQCE variations confirm its role in postaxial polydactyly and cause ciliary defect phenotype in zebrafish ». *Human Mutation* 41 (1): 240-54. <https://doi.org/10.1002/humu.23924>.
- Farag, Talaat I., et Ahmad S. Teebi. 1989. « High incidence of Bardet Biedl syndrome among the Bedouin ». *Clinical Genetics* 36 (6): 463-64. <https://doi.org/10.1111/j.1399-0004.1989.tb03378.x>.
- Firth, Helen V., Shola M. Richards, A. Paul Bevan, Stephen Clayton, Manuel Corpas, Diana Rajan, Steven Van Vooren, Yves Moreau, Roger M. Pettett, et Nigel P. Carter. 2009. « DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources ». *The American Journal of Human Genetics* 84 (4): 524-33. <https://doi.org/10.1016/j.ajhg.2009.03.010>.
- Follit, John A., Jovenal T. San Agustin, Fenghui Xu, Julie A. Jonassen, Rajeev Samtani, Cecilia W. Lo, et Gregory J. Pazour. 2008. « The Golgin GMAP210/TRIP11 Anchors IFT20 to the Golgi Complex ». Édité par Susan K. Dutcher. *PLoS Genetics* 4 (12): e1000315. <https://doi.org/10.1371/journal.pgen.1000315>.
- Forsyth, Raelynn, et Meral Gunay-Aygun. 1993. « Bardet-Biedl Syndrome Overview ». In *GeneReviews®*, édité par Margaret P. Adam, Holly H. Ardinger, Roberta A. Pagon, Stephanie E. Wallace, Lora JH Bean, Karen Stephens, et Anne Amemiya. Seattle (WA): University of Washington, Seattle. <http://www.ncbi.nlm.nih.gov/books/NBK1363/>.
- Fry, Andrew M., Michelle J. Leaper, et Richard Bayliss. 2014. « The Primary Cilium: Guardian of Organ Development and Homeostasis ». *Organogenesis* 10 (1): 62-68. <https://doi.org/10.4161/org.28910>.
- Galluzzi, Lorenzo, et Douglas R. Green. 2019. « Autophagy-Independent Functions of the Autophagy Machinery ». *Cell* 177 (7): 1682-99. <https://doi.org/10.1016/j.cell.2019.05.026>.
- Gardner, Eugene J., Elena Prigmore, Giuseppe Gallone, Petr Danecek, Kaitlin E. Samocha, Juliet Handsaker, Sebastian S. Gerety, et al. 2019. « Contribution of retrotransposition to developmental disorders ». *Nature Communications* 10 (1): 4630. <https://doi.org/10.1038/s41467-019-12520-y>.
- Gasior, Stephen L., Graeme Preston, Dale J. Hedges, Nicolas Gilbert, John V. Moran, et Prescott L. Deininger. 2007. « Characterization of Pre-Insertion Loci of de Novo L1 Insertions ». *Gene* 390 (1-2): 190-98. <https://doi.org/10.1016/j.gene.2006.08.024>.
- Genin, E, A Tullio-Pelet, F Begeot, S Lyonnet, et L Abel. 2004. « Estimating the age of rare disease mutations: the example of Triple-A syndrome ». *Journal of Medical Genetics* 41 (6): 445. <https://doi.org/10.1136/jmg.2003.017962>.
- Geoffroy, Véronique, Yvan Herenger, Arnaud Kress, Corinne Stoetzel, Amélie Piton, Hélène Dollfus, et Jean Muller. 2018. « AnnotSV: An Integrated Tool for Structural Variations Annotation ». *Bioinformatics (Oxford, England)* 34 (20): 3572-74. <https://doi.org/10.1093/bioinformatics/bty304>.
- Geoffroy, Véronique, Cécile Pizot, Claire Redin, Amélie Piton, Nasim Vasli, Corinne Stoetzel, André Blavier, Jocelyn Laporte, et Jean Muller. 2015. « VaRank: A Simple and Powerful Tool for Ranking Genetic Variants ». *PeerJ* 3: e796. <https://doi.org/10.7717/peerj.796>.

- Geoffroy, Véronique, Corinne Stoetzel, Sophie Scheidecker, Elise Schaefer, Isabelle Perrault, Séverine Bär, Ariane Kröll, Marion Delbarre, Manuela Antin, Anne-Sophie Leuvrey, Charline Henry, Hélène Blanché, Eva Decker, Katja Kloth, Günter Klaus, Christoph Mache, Dominique Martin-Coignard, et al. 2018. « Whole-Genome Sequencing in Patients with Ciliopathies Uncovers a Novel Recurrent Tandem Duplication in IFT140 ». *Human Mutation* 39 (7): 983-92. <https://doi.org/10.1002/humu.23539>.
- Geoffroy, Véronique, Corinne Stoetzel, Sophie Scheidecker, Elise Schaefer, Isabelle Perrault, Séverine Bär, Ariane Kröll, Marion Delbarre, Manuela Antin, Anne-Sophie Leuvrey, Charline Henry, Hélène Blanché, Eva Decker, Katja Kloth, Günter Klaus, Christoph Mache, Dominique Martin-Coignard, et al. 2018. « Whole-genome sequencing in patients with ciliopathies uncovers a novel recurrent tandem duplication in IFT140 ». *Human Mutation* 39 (7): 983-92. <https://doi.org/10.1002/humu.23539>.
- Gherman, Adrian, Erica E. Davis, et Nicholas Katsanis. 2006. « The Ciliary Proteome Database: An Integrated Community Resource for the Genetic and Functional Dissection of Cilia ». *Nature Genetics* 38 (9): 961-63.
- Gibbons, I. R., et A. J. Rowe. 1965. « Dynein: A Protein with Adenosine Triphosphatase Activity from Cilia ». *Science (New York, N.Y.)* 149 (3682): 424-26. <https://doi.org/10.1126/science.149.3682.424>.
- Goerner-Potvin, Patricia, et Guillaume Bourque. 2018. « Computational tools to unmask transposable elements ». *Nature Reviews Genetics* 19 (11): 688-704. <https://doi.org/10.1038/s41576-018-0050-x>.
- Goetz, Sarah C., Polloneal J.R. Ocbina, et Kathryn V. Anderson. 2009. « The Primary Cilium as a Hedgehog Signal Transduction Machine ». *Methods in cell biology* 94: 199-222. [https://doi.org/10.1016/S0091-679X\(08\)94010-3](https://doi.org/10.1016/S0091-679X(08)94010-3).
- Goetz, Sarah C., Karel F. Liem, et Kathryn V. Anderson. 2012. « The Spinocerebellar Ataxia-Associated Gene Tau Tubulin Kinase 2 Controls the Initiation of Ciliogenesis ». *Cell* 151 (4): 847-58. <https://doi.org/10.1016/j.cell.2012.10.010>.
- Gonzalez, Cesar, Enrique Armijo, Javiera Bravo-Alegria, Andrea Becerra-Calixto, Charles E. Mays, et Claudio Soto. 2018. « Modeling Amyloid Beta and Tau Pathology in Human Cerebral Organoids ». *Molecular Psychiatry* 23 (12): 2363-74. <https://doi.org/10.1038/s41380-018-0229-8>.
- Goto, Hidemasa, Hironori Inaba, et Masaki Inagaki. 2017. « Mechanisms of ciliogenesis suppression in dividing cells ». *Cellular and Molecular Life Sciences* 74 (5): 881-90. <https://doi.org/10.1007/s00018-016-2369-9>.
- Gouronc, Aurélie, Vincent Zilliox, Marie-Line Jacquemont, Françoise Darcel, Anne-Sophie Leuvrey, Elsa Nourisson, Manuela Antin, et al. 2020. « High prevalence of Bardet-Biedl syndrome in La Réunion Island is due to a founder variant in ARL6/BBS3 ». *Clinical Genetics* n/a (n/a). <https://doi.org/10.1111/cge.13768>.
- Grochowsky, Angela, et Meral Gunay-Aygun. s. d. « Clinical characteristics of individual organ system disease in non-motile ciliopathies ». *Translational Science of Rare Diseases* 4 (1-2): 1-23. <https://doi.org/10.3233/TRD-190033>.
- Gstrein, Thomas, Andrew Edwards, Anna Přistoupilová, Ines Leca, Martin Breuss, Sandra Pilat-Carotta, Andi H. Hansen, et al. 2018. « Mutations in Vps15 Perturb Neuronal Migration in Mice and Are Associated with Neurodevelopmental Disease in Humans ». *Nature Neuroscience* 21 (2): 207-17. <https://doi.org/10.1038/s41593-017-0053-5>.
- Ha, Hongseok, Jui Wan Loh, et Jinchuan Xing. 2016. « Identification of Polymorphic SVA Retrotransposons Using a Mobile Element Scanning Method for SVA (ME-Scan-SVA) ». *Mobile DNA* 7 (1). <https://doi.org/10.1186/s13100-016-0072-x>.

- Hancks, Dustin C., et Haig H. Kazazian. 2010. « SVA Retrotransposons: Evolution and Genetic Instability ». *Seminars in Cancer Biology* 20 (4): 234-45. <https://doi.org/10.1016/j.semcancer.2010.04.001>.
- Hancks, Dustin C, et Haig H Kazazian. 2012. « Active Human Retrotransposons: Variation and Disease ». *Current Opinion in Genetics & Development* 22 (3): 191-203. <https://doi.org/10.1016/j.gde.2012.02.006>.
- Hancks, Dustin C., et Haig H. Kazazian. 2016. « Roles for retrotransposon insertions in human disease ». *Mobile DNA* 7 (mai). <https://doi.org/10.1186/s13100-016-0065-9>.
- Hanein, Sylvain, Isabelle Perrault, Sylvie Gerber, Gaëlle Tanguy, Fabienne Barbet, Dominique Ducroq, Patrick Calvas, et al. 2004. « Leber Congenital Amaurosis: Comprehensive Survey of the Genetic Heterogeneity, Refinement of the Clinical Definition, and Genotype-Phenotype Correlations as a Strategy for Molecular Diagnosis ». *Human Mutation* 23 (4): 306-17. <https://doi.org/10.1002/humu.20010>.
- Haslam, R. J., H. B. Koide, et B. A. Hemmings. 1993. « Pleckstrin Domain Homology ». *Nature* 363 (6427): 309-10. <https://doi.org/10.1038/363309b0>.
- Hassoun, H., T. L. Coetzer, J. N. Vassiliadis, K. E. Sahr, G. J. Maalouf, S. T. Saad, L. Catanzariti, et J. Palek. 1994. « A Novel Mobile Element Inserted in the Alpha Spectrin Gene: Spectrin Dayton. A Truncated Alpha Spectrin Associated with Hereditary Elliptocytosis ». *The Journal of Clinical Investigation* 94 (2): 643-48. <https://doi.org/10.1172/JCI117380>.
- Hayano, Takahide, Shiro Yamada, Kazuyoshi Hosomichi, Hirofumi Nakaoka, Kosuke Yoshihara, Sosuke Adachi, Katsunori Kashima, Kenichi Tanaka, Takayuki Enomoto, et Ituro Inoue. 2015. « Identification of Novel Exonic Mobile Element Insertions in Epithelial Ovarian Cancers ». *Human Genome Variation* 2 (octobre): 15030-15030. <https://doi.org/10.1038/hgv.2015.30>.
- Hichri, Haifa, Corinne Stoetzel, Virginie Laurier, Solenne Caron, Sabine Sigaudy, Pierre Sarda, Christian Hamel, et al. 2005. « Testing for Triallelism: Analysis of Six BBS Genes in a Bardet-Biedl Syndrome Family Cohort ». *European Journal of Human Genetics: EJHG* 13 (5): 607-16. <https://doi.org/10.1038/sj.ejhg.5201372>.
- Hill, Patrick, Katrin Götz, et Ulrich Rüther. 2009. « A SHH-Independent Regulation of Gli3 Is a Significant Determinant of Anteroposterior Patterning of the Limb Bud ». *Developmental Biology* 328 (2): 506-16. <https://doi.org/10.1016/j.ydbio.2009.02.017>.
- Huang, Li-Kai, Shu-Ping Chao, et Chaur-Jong Hu. 2020. « Clinical trials of new drugs for Alzheimer disease ». *Journal of Biomedical Science* 27 (1): 18. <https://doi.org/10.1186/s12929-019-0609-7>.
- Huangfu, Danwei, et Kathryn V. Anderson. 2005. « Cilia and Hedgehog Responsiveness in the Mouse ». *Proceedings of the National Academy of Sciences* 102 (32): 11325-30. <https://doi.org/10.1073/pnas.0505328102>.
- Huberman, Andrew D., et Cristopher M. Niell. 2011. « What can mice tell us about how vision works? ». *Trends in Neurosciences* 34 (9): 464-73. <https://doi.org/10.1016/j.tins.2011.07.002>.
- Hughes, Steven, Jessica Rodgers, Doron Hickey, Russell Foster, Stuart Peirson, et Mark Hankins. 2016. « Characterisation of light responses in the retina of mice lacking principle components of rod, cone and melanopsin phototransduction signalling pathways ». *Scientific Reports* 6 (juin): 28086. <https://doi.org/10.1038/srep28086>.



- Hurley, James H., et Lindsey N. Young. 2017. « Mechanisms of Autophagy Initiation ». *Annual Review of Biochemistry* 86: 225-44. <https://doi.org/10.1146/annurev-biochem-061516-044820>.
- Imani, Saber, Jingliang Cheng, Jiewen Fu, Abdolkarim Mobasher-Jannat, Chunli Wei, Saman Mohazzab-Torabi, Khosrow Jadidi, et al. 2019. « Novel Splicing Variant c. 208+2T>C in BBS5 Segregates with Bardet-Biedl Syndrome in an Iranian Family by Targeted Exome Sequencing ». *Bioscience Reports* 39 (3). <https://doi.org/10.1042/BSR20181544>.
- Jonassen, Julie A., Jovenal San Agustin, John A. Follit, et Gregory J. Pazour. 2008. « Deletion of IFT20 in the mouse kidney causes misorientation of the mitotic spindle and cystic kidney disease ». *The Journal of Cell Biology* 183 (3): 377-84. <https://doi.org/10.1083/jcb.200808137>.
- Kabir, Mohammad Fazlul, Hyung-Ryong Kim, et Han-Jung Chae. 2018. « Endoplasmic Reticulum Stress and Autophagy ». *Endoplasmic Reticulum*, novembre. <https://doi.org/10.5772/intechopen.81381>.
- Karaca, Ender, Jennifer E. Posey, Zeynep Coban Akdemir, Davut Pehlivan, Tamar Harel, Shalini N. Jhangiani, Yavuz Bayram, et al. 2018. « Phenotypic expansion illuminates multilocus pathogenic variation ». *Genetics in Medicine* 20 (12): 1528-37. <https://doi.org/10.1038/gim.2018.33>.
- Karczewski, Konrad J., Laurent C. Francioli, Grace Tiao, Beryl B. Cummings, Jessica Alföldi, Qingbo Wang, Ryan L. Collins, et al. 2020. « The Mutational Constraint Spectrum Quantified from Variation in 141,456 Humans ». *Nature* 581 (7809): 434-43. <https://doi.org/10.1038/s41586-020-2308-7>.
- Katsanis, Nicholas, Stephen J. Ansley, Jose L. Badano, Erica R. Eichers, Richard Alan Lewis, Bethan E. Hoskins, Peter J. Scambler, William S. Davidson, Philip L. Beales, et James R. Lupski. 2001a. « Triallelic Inheritance in Bardet-Biedl Syndrome, a Mendelian Recessive Disorder ». *Science* 293 (5538): 2256. <https://doi.org/10.1126/science.1063525>.
- Kazatskaya, Anna, Stefanie Kuhns, Nils J. Lambacher, Julie E. Kennedy, Andrea G. Brear, Gavin J. McManus, Piali Sengupta, et Oliver E. Blacque. 2017. « Primary Cilium Formation and Ciliary Protein Trafficking Is Regulated by the Atypical MAP Kinase MAPK15 in Caenorhabditis Elegans and Human Cells ». *Genetics* 207 (4): 1423-40. <https://doi.org/10.1534/genetics.117.300383>.
- Kazazian, Haig H., Corinne Wong, Hagop Youssoufian, Alan F. Scott, Deborah G. Phillips, et Stylianos E. Antonarakis. 1988a. « Haemophilia A Resulting from de Novo Insertion of L1 Sequences Represents a Novel Mechanism for Mutation in Man ». *Nature* 332 (6160): 164-66. <https://doi.org/10.1038/332164a0>.
- Kent, W James. 2002. « BLAT--the BLAST-like Alignment Tool ». *Genome Research* 12 (4): 656-64. <https://doi.org/10.1101/gr.229202>.
- Khan, Sher Alam, Noor Muhammad, Muzammil Ahmad Khan, Anwar Kamal, Zia Ur Rehman, et Saadullah Khan. 2016. « Genetics of Human Bardet-Biedl Syndrome, an Updates ». *Clinical Genetics*, janvier. <https://doi.org/10.1111/cge.12737>.
- Kherraf, Zine-Eddine, Amir Amiri-Yekta, Denis Dacheux, Thomas Karaouzène, Charles Coutton, Marie Christou-Kent, Guillaume Martinez, et al. 2018. « A Homozygous Ancestral SVA-Insertion-Mediated Deletion in WDR66 Induces Multiple Morphological Abnormalities of the Sperm Flagellum and Male Infertility ». *The American Journal of Human Genetics* 103 (3): 400-412. <https://doi.org/10.1016/j.ajhg.2018.07.014>.
- Klift, Heleen M. van der, Carli M. Tops, Frederik J. Hes, Peter Devilee, et Juul T. Wijnen. 2012. « Insertion of an SVA Element, a Nonautonomous Retrotransposon, in PMS2 Intron 7 as a Novel Cause of Lynch Syndrome ». *Human Mutation* 33 (7): 1051-55. <https://doi.org/10.1002/humu.22092>.

- Klink, Björn Udo, Christos Gatsogiannis, Oliver Hofnagel, Alfred Wittinghofer, et Stefan Raunser. 2020a. « Structure of the human BBSome core complex ». Édité par Andrew P Carter, John Kuriyan, et Masahide Kikkawa. *eLife* 9 (janvier): e53910. <https://doi.org/10.7554/eLife.53910>.
- Knighton, D. R., J. H. Zheng, LF Ten Eyck, N. H. Xuong, S. S. Taylor, et J. M. Sowadski. 1991. « Structure of a Peptide Inhibitor Bound to the Catalytic Subunit of Cyclic Adenosine Monophosphate-Dependent Protein Kinase ». *Science* 253 (5018): 414-20. <https://doi.org/10.1126/science.1862343>.
- Kobayashi, K., Y. Nakahori, M. Miyake, K. Matsumura, E. Kondo-Iida, Y. Nomura, M. Segawa, et al. 1998. « An Ancient Retrotransposal Insertion Causes Fukuyama-Type Congenital Muscular Dystrophy ». *Nature* 394 (6691): 388-92. <https://doi.org/10.1038/28653>.
- Kong, Jennifer H., Christian Siebold, et Rajat Rohatgi. 2019. « Biochemical Mechanisms of Vertebrate Hedgehog Signaling ». *Development* 146 (10). <https://doi.org/10.1242/dev.166892>.
- Kozminski, K. G., K. A. Johnson, P. Forscher, et J. L. Rosenbaum. 1993. « A Motility in the Eukaryotic Flagellum Unrelated to Flagellar Beating ». *Proceedings of the National Academy of Sciences* 90 (12): 5519-23. <https://doi.org/10.1073/pnas.90.12.5519>.
- Kremer, Laura S., Daniel M. Bader, Christian Mertes, Robert Kopajtich, Garwin Pichler, Arcangela Iuso, Tobias B. Haack, et al. 2017. « Genetic diagnosis of Mendelian disorders via RNA sequencing ». *Nature Communications* 8 (1): 15824. <https://doi.org/10.1038/ncomms15824>.
- Kwon, Yun-Jeong, Yuri Choi, Jungwoo Eo, Yu-Na Noh, Jeong-An Gim, Yi-Deun Jung, Ja-Rang Lee, et Heui-Soo Kim. 2013a. « Structure and Expression Analyses of SVA Elements in Relation to Functional Genes ». *Genomics & Informatics* 11 (3): 142. <https://doi.org/10.5808/GI.2013.11.3.142>.
- Lacaná, Emanuela, Michael Maceyka, Sheldon Milstien, et Sarah Spiegel. 2002. « Cloning and Characterization of a Protein Kinase A Anchoring Protein (AKAP)-Related Protein That Interacts with and Regulates Sphingosine Kinase 1 Activity ». *Journal of Biological Chemistry* 277 (36): 32947-53. <https://doi.org/10.1074/jbc.M202841200>.
- Lachance, Véronik, Qian Wang, Eric Sweet, Insup Choi, Cui-Zan Cai, Xu-Xu Zhuang, Yuanxi Zhang, et al. « Autophagy protein NRBF2 has reduced expression in Alzheimer's brains and modulates memory and amyloid-beta homeostasis in mice ». *Molecular Neurodegeneration* 14, n° 1 (27 novembre 2019): 43. <https://doi.org/10.1186/s13024-019-0342-4>.
- LAURENCE, J. 1866. « Four cases of retinitis pigmentosa occurring in the same family, and accompanied by general imperfections of development ». *Ophthalmol Rev* 2: 32-41.
- Laurier, Virginie, Corinne Stoetzel, Jean Muller, Christelle Thibault, Sandra Corbani, Nadine Jalkh, Nabihah Salem, et al. 2006. « Pitfalls of Homozygosity Mapping: An Extended Consanguineous Bardet-Biedl Syndrome Family with Two Mutant Genes (BBS2, BBS10), Three Mutations, but No Triallelism ». *European Journal of Human Genetics: EJHG* 14 (11): 1195-1203. <https://doi.org/10.1038/sj.ejhg.5201688>.
- Lehtreck, Karl F., Julie C. Van De Weghe, J. Aaron Harris, et Peiwei Liu. 2017. « Protein transport in growing and steady-state cilia ». *Traffic (Copenhagen, Denmark)* 18 (5): 277-86. <https://doi.org/10.1111/tra.12474>.
- Lee, Christopher M, Galt P Barber, Jonathan Casper, Hiram Clawson, Mark Diekhans, Jairo Navarro Gonzalez, Angie S Hinrichs, et al. 2019. « UCSC Genome Browser enters 20th year ». *Nucleic Acids Research* 48 (D1): D756-61. <https://doi.org/10.1093/nar/gkz1012>.



- Leeuwenhoek, Antoni Van. 1677. « Observations, communicated to the publisher by Mr. Antony van Leewenhoek, in a dutch letter of the 9th Octob. 1676. here English'd: concerning little animals by him observed in rain-well-sea- and snow water; as also in water wherein pepper had lain infused ». *Philosophical Transactions of the Royal Society of London* 12 (133): 821-31. <https://doi.org/10.1098/rstl.1677.0003>.
- Li, Heng. 2011. « A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data ». *Bioinformatics* 27 (21): 2987-93. <https://doi.org/10.1093/bioinformatics/btr509>.
- Li, Heng, et Richard Durbin. 2009. « Fast and accurate short read alignment with Burrows–Wheeler transform ». *Bioinformatics* 25 (14): 1754-60. <https://doi.org/10.1093/bioinformatics/btp324>.
- Li, Jin Billy, Jantje M. Gerdes, Courtney J. Haycraft, Yanli Fan, Tanya M. Teslovich, Helen May-Simera, Haitao Li, et al. 2004. « Comparative Genomics Identifies a Flagellar and Basal Body Proteome That Includes the BBS5 Human Disease Gene ». *Cell* 117 (4): 541-52. [https://doi.org/10.1016/s0092-8674\(04\)00450-7](https://doi.org/10.1016/s0092-8674(04)00450-7).
- Li, Quan, et Kai Wang. 2017. « InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines ». *The American Journal of Human Genetics* 100 (2): 267-80. <https://doi.org/10.1016/j.ajhg.2017.01.004>.
- Li, Xueyuan, Xinjie Bao, et Renzhi Wang. 2016. « Experimental models of Alzheimer's disease for deciphering the pathogenesis and therapeutic screening (Review) ». *International Journal of Molecular Medicine* 37 (2): 271-83. <https://doi.org/10.3892/ijmm.2015.2428>.
- Liu, Jie, et Lian Li. « Targeting Autophagy for the Treatment of Alzheimer's Disease: Challenges and Opportunities ». *Frontiers in Molecular Neuroscience* 12 (22 août 2019). <https://doi.org/10.3389/fnmol.2019.00203>.
- Liu, Pei-Pei, Yi Xie, Xiao-Yan Meng, et Jian-Sheng Kang. « History and progress of hypotheses and clinical trials for Alzheimer's disease ». *Signal Transduction and Targeted Therapy* 4 (23 août 2019). <https://doi.org/10.1038/s41392-019-0063-8>.
- Lindstrand, Anna, Erica E. Davis, Claudia M. B. Carvalho, Davut Pehlivan, Jason R. Willer, I.-Chun Tsai, Subhadra Ramanathan, et al. 2014. « Recurrent CNVs and SNVs at the NPHP1 Locus Contribute Pathogenic Alleles to Bardet-Biedl Syndrome. » *American Journal of Human Genetics* 94 (5): 745-54. <https://doi.org/10.1016/j.ajhg.2014.03.017>.
- Lindstrand, Anna, Stephan Frangakis, Claudia M.B. Carvalho, Ellen B. Richardson, Kelsey A. McFadden, Jason R. Willer, Davut Pehlivan, et al. 2016. « Copy-Number Variation Contributes to the Mutational Load of Bardet-Biedl Syndrome ». *The American Journal of Human Genetics* 99 (2): 318-36. <https://doi.org/10.1016/j.ajhg.2015.04.023>.
- Lindström, Nils O., Jill A. McMahon, Jinjin Guo, Tracy Tran, Qiuyu Guo, Elisabeth Rutledge, Riana K. Parvez, et al. 2018. « Conserved and Divergent Features of Human and Mouse Kidney Organogenesis ». *Journal of the American Society of Nephrology: JASN* 29 (3): 785-805. <https://doi.org/10.1681/ASN.2017080887>.
- Liu, Pei-Pei, Yi Xie, Xiao-Yan Meng, et Jian-Sheng Kang. 2019. « History and progress of hypotheses and clinical trials for Alzheimer's disease ». *Signal Transduction and Targeted Therapy* 4 (août). <https://doi.org/10.1038/s41392-019-0063-8>.
- Long, Huan, et Kaiyao Huang. 2020. « Transport of Ciliary Membrane Proteins ». *Frontiers in Cell and Developmental Biology* 7. <https://doi.org/10.3389/fcell.2019.00381>.

- Lorenzo, Alfredo, Bronwyn Razzaboni, Gordon C. Weir, et Bruce A. Yankner. 1994. « Pancreatic Islet Cell Toxicity of Amylin Associated with Type-2 Diabetes Mellitus ». *Nature* 368 (6473): 756-60. <https://doi.org/10.1038/368756a0>.
- Lou, Yan-Ru, et Alan W. Leung. 2018. « Next Generation Organoids for Biomedical Research and Applications ». *Biotechnology Advances* 36 (1): 132-49. <https://doi.org/10.1016/j.biotechadv.2017.10.005>.
- Madugula, Viswanadh, et Lei Lu. 2016. « The Ciliary Targeting of Membrane Proteins by a Ternary Complex Comprising Transportin1, Rab8 and the Ciliary Targeting Signal ». *BioRxiv*, juin, 059451. <https://doi.org/10.1101/059451>.
- Maiti, Samarpan, Susmita Mondal, Eswara M Satyavarapu, et Chitra Mandal. 2017. « mTORC2 regulates hedgehog pathway activity by promoting stability to Gli2 protein and its nuclear translocation ». *Cell Death & Disease* 8 (7): e2926. <https://doi.org/10.1038/cddis.2017.296>.
- Mancini, Gianni, Cândida Dias, Cátia Lourenço, Rui M. Barbosa, João Laranjinha, Andreza F. de Bem, et Ana Ledo. 2018. « ECS-5 - High-Fat and High Cholesterol Diet Induces Alzheimer's like Disease in Mice ». *Free Radical Biology and Medicine* 120 (mai): S25. <https://doi.org/10.1016/j.freeradbiomed.2018.04.027>.
- Marion, Vincent, Anaïs Mockel, Charlie De Melo, Cathy Obringer, Aurélie Claussmann, Alban Simon, Nadia Messaddeq, et al. 2012. « BBS-Induced Ciliary Defect Enhances Adipogenesis, Causing Paradoxical Higher-Insulin Sensitivity, Glucose Usage, and Decreased Inflammatory Response ». *Cell Metabolism* 16 (3): 363-77. <https://doi.org/10.1016/j.cmet.2012.08.005>.
- Marion, Vincent, Fanny Stutzmann, Marion Gérard, Charlie De Melo, Elise Schaefer, Aurélie Claussmann, Sophie Hellé, et al. 2012. « Exome Sequencing Identifies Mutations in LZTFL1, a BBSome and Smoothed Trafficking Regulator, in a Family with Bardet--Biedl Syndrome with Situs Inversus and Insertional Polydactyly ». *Journal of Medical Genetics* 49 (5): 317-21. <https://doi.org/10.1136/jmedgenet-2012-100737>.
- Martin, Sandra L. 2018. « On the move ». *eLife* 7 (février): e34901. <https://doi.org/10.7554/eLife.34901>.
- Mary, Laura, Kirsley Chennen, Corinne Stoetzel, Manuela Antin, Anne Leuvrey, Elsa Nourisson, Elisabeth Alanio-Detton, et al. 2019. « Bardet-Biedl syndrome: Antenatal presentation of forty-five fetuses with biallelic pathogenic variants in known Bardet-Biedl syndrome genes ». *Clinical Genetics* 95 (3): 384-97. <https://doi.org/10.1111/cge.13500>.
- Mathieson, Iain, et Gil McVean. 2014. « Demography and the Age of Rare Variants ». *PLOS Genetics* 10 (8): e1004528. <https://doi.org/10.1371/journal.pgen.1004528>.
- Mattapallil, Mary J., Eric F. Wawrousek, Chi-Chao Chan, Hui Zhao, Jayeeta Roychoudhury, Thomas A. Ferguson, et Rachel R. Caspi. 2012. « The Rd8 Mutation of the Crb1 Gene Is Present in Vendor Lines of C57BL/6N Mice and Embryonic Stem Cells, and Confounds Ocular Induced Mutant Phenotypes ». *Investigative Ophthalmology & Visual Science* 53 (6): 2921-27. <https://doi.org/10.1167/iovs.12-9662>.
- Mavrakis, Manos, Jennifer Lippincott-Schwartz, Constantine A. Stratakis, et Ioannis Bossis. 2006. « Depletion of Type IA Regulatory Subunit (RIalpha) of Protein Kinase A (PKA) in Mammalian Cells and Tissues Activates MTOR and Causes Autophagic Deficiency ». *Human Molecular Genetics* 15 (19): 2962-71. <https://doi.org/10.1093/hmg/ddl239>.
- Mayer, B. J., R. Ren, K. L. Clark, et D. Baltimore. 1993. « A Putative Modular Domain Present in Diverse Signaling Proteins ». *Cell* 73 (4): 629-30. [https://doi.org/10.1016/0092-8674\(93\)90244-k](https://doi.org/10.1016/0092-8674(93)90244-k).

- Mazzaccara, Cristina, Giuseppe Labruna, Gennaro Cito, Marzia Scarfò, Mario De Felice, Lucio Pastore, et Lucia Sacchetti. 2008. « Age-Related Reference Intervals of the Main Biochemical and Hematological Parameters in C57BL/6J, 129SV/EV and C3H/HeJ Mouse Strains ». *PLoS ONE* 3 (11). <https://doi.org/10.1371/journal.pone.0003772>.
- Medina, Julie, et Hervé Perron. 2017. « Séquences provenant d'éléments génétiques mobiles, face cachée du génome humain ». *médecine/sciences* 33 (2): 151-58. <https://doi.org/10.1051/medsci/20173302010>.
- Medstrand, Patrik, Louie N. van de Lagemaat, et Dixie L. Mager. 2002. « Retroelement Distributions in the Human Genome: Variations Associated With Age and Proximity to Genes ». *Genome Research* 12 (10): 1483-95. <https://doi.org/10.1101/gr.388902>.
- Meehan, Terrence F., Nathalie Conte, David B. West, Julius O. Jacobsen, Jeremy Mason, Jonathan Warren, Chao-Kung Chen, et al. 2017. « Disease Model Discovery from 3,328 Gene Knockouts by The International Mouse Phenotyping Consortium ». *Nature genetics* 49 (8): 1231-38. <https://doi.org/10.1038/ng.3901>.
- Meienberg, Janine, Rémy Bruggmann, Konrad Oexle, et Gabor Matyas. 2016. « Clinical sequencing: is WGS the better WES? » *Human Genetics* 135: 359-62. <https://doi.org/10.1007/s00439-015-1631-9>.
- Miao, Jin, Ruirui Shi, Longfei Li, Feng Chen, Yan Zhou, Yunn Chyn Tung, Wen Hu, Cheng-Xin Gong, Khalid Iqbal, et Fei Liu. 2019. « Pathological Tau From Alzheimer's Brain Induces Site-Specific Hyperphosphorylation and SDS- and Reducing Agent-Resistant Aggregation of Tau in Vivo ». *Frontiers in Aging Neuroscience* 11. <https://doi.org/10.3389/fnagi.2019.00034>.
- Miki, Yoshio, Toyomasa Katagiri, Fujio Kasumi, Takamasa Yoshimoto, et Yusuke Nakamura. 1996. « Mutation Analysis in the BRCA2 Gene in Primary Breast Cancers ». *Nature Genetics* 13 (2): 245-47. <https://doi.org/10.1038/ng0696-245>.
- Mills, Ryan E., E. Andrew Bennett, Rebecca C. Iskow, et Scott E. Devine. 2007. « Which Transposable Elements Are Active in the Human Genome? » *Trends in Genetics* 23 (4): 183-91. <https://doi.org/10.1016/j.tig.2007.02.006>.
- Mitchison, Hannah M., et Enza Maria Valente. 2017a. « Motile and Non-Motile Cilia in Human Pathology: From Function to Phenotypes ». *The Journal of Pathology* 241 (2): 294-309. <https://doi.org/10.1002/path.4843>.
- Mitchison, Hannah M, et Enza Maria Valente. 2017b. « Motile and non-motile cilia in human pathology: from function to phenotypes ». *The Journal of Pathology* 241 (2): 294-309. <https://doi.org/10.1002/path.4843>.
- Molinari, Elisa, et John A. Sayer. 2020. « Disease Modeling To Understand the Pathomechanisms of Human Genetic Kidney Disorders ». *Clinical Journal of the American Society of Nephrology* 15 (6): 855-72. <https://doi.org/10.2215/CJN.08890719>.
- Morleo, Manuela, et Brunella Franco. 2019. « The Autophagy-Cilia Axis: An Intricate Relationship ». *Cells* 8 (8). <https://doi.org/10.3390/cells8080905>.
- Moruno Manchon, Jose Felix, Ndidi-Ese Uzor, Steven Finkbeiner, et Andrey S. Tsvetkov. 2016. « SPHK1/sphingosine kinase 1-mediated autophagy differs between neurons and SH-SY5Y neuroblastoma cells ». *Autophagy* 12 (8): 1418-24. <https://doi.org/10.1080/15548627.2016.1183082>.
- Mourão, André, Andrew R. Nager, Maxence V. Nachury, et Esben Lorentzen. 2014. « Structural Basis for Membrane Targeting of the BBSome by ARL6 ». *Nature structural & molecular biology* 21 (12): 1035-41. <https://doi.org/10.1038/nsmb.2920>.

- Muller, Jean, C. Stoetzel, M. C. Vincent, C. C. Leitch, V. Laurier, J. M. Danse, S. Hellé, et al. 2010a. « Identification of 28 novel mutations in the Bardet–Biedl syndrome genes: the burden of private mutations in an extensively heterogeneous disease ». *Human genetics* 127 (5): 583-93. <https://doi.org/10.1007/s00439-010-0804-9>.
- Murawski, Inga J, Rita W Maina, et Indra R Gupta. 2010. « The relationship between nephron number, kidney size and body weight in two inbred mouse strains ». *Organogenesis* 6 (3): 189-94.
- Myers, Ariana, et Paul McGonigle. « Overview of Transgenic Mouse Models for Alzheimer’s Disease ». *Current Protocols in Neuroscience* 89, n° 1 (2019): e81. <https://doi.org/10.1002/cpns.81>.
- Mykytyn, Kirk, Darryl Y. Nishimura, Charles C. Searby, Gretel Beck, Kevin Bugge, Heidi L. Haines, Alberto S. Cornier, et al. 2003. « Evaluation of Complex Inheritance Involving the Most Common Bardet-Biedl Syndrome Locus (BBS1) ». *The American Journal of Human Genetics* 72 (2): 429-37. <https://doi.org/10.1086/346172>.
- Mykytyn, Kirk, Darryl Y. Nishimura, Charles C. Searby, Mythreyi Shastri, Hsan-jan Yen, John S. Beck, Terry Braun, et al. 2002a. « Identification of the Gene (BBS1) Most Commonly Involved in Bardet-Biedl Syndrome, a Complex Human Obesity Syndrome ». *Nature Genetics* 31 (4): 435-38. <https://doi.org/10.1038/ng935>.
- Nachury, Maxence V., Alexander V. Loktev, Qihong Zhang, Christopher J. Westlake, Johan Peränen, Andreas Merdes, Diane C. Slusarski, et al. 2007a. « A Core Complex of BBS Proteins Cooperates with the GTPase Rab8 to Promote Ciliary Membrane Biogenesis ». *Cell* 129 (6): 1201-13. <https://doi.org/10.1016/j.cell.2007.03.053>.
- Nakamura, Yuki, Moe Murata, Yuki Takagi, Toshihiro Kozuka, Yukiko Nakata, Ryo Hasebe, Akira Takagi, Jun-ichi Kitazawa, Midori Shima, et Tetsuhito Kojima. 2015. « SVA Retrotransposition in Exon 6 of the Coagulation Factor IX Gene Causing Severe Hemophilia B ». *International Journal of Hematology* 102 (1): 134-39. <https://doi.org/10.1007/s12185-015-1765-5>.
- Nakayama, Kazuhisa, et Yohei Katoh. 2018. « Ciliary Protein Trafficking Mediated by IFT and BBSome Complexes with the Aid of Kinesin-2 and Dynein-2 Motors ». *The Journal of Biochemistry* 163 (3): 155-64. <https://doi.org/10.1093/jb/mvx087>.
- Natsume, Toyoaki, Tomomi Kiyomitsu, Yumiko Saga, et Masato T. Kanemaki. 2016. « Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors ». *Cell Reports* 15 (1): 210-18. <https://doi.org/10.1016/j.celrep.2016.03.001>.
- Nemazanyy, Ivan, Bert Blaauw, Cecilia Paolini, Catherine Caillaud, Feliciano Protasi, Amelie Mueller, Tassula Proikas-Cezanne, et al. 2013. « Defects of Vps15 in Skeletal Muscles Lead to Autophagic Vacuolar Myopathy and Lysosomal Disease ». *EMBO Molecular Medicine* 5 (6): 870-90. <https://doi.org/10.1002/emmm.201202057>.
- Nevers, Yannis, Megana K. Prasad, Laetitia Poidevin, Kirsley Chennen, Alexis Allot, Arnaud Kress, Raymond Ripp, et al. 2017. « Insights into Ciliary Genes and Evolution from Multi-Level Phylogenetic Profiling ». *Molecular Biology and Evolution* 34 (8): 2016-34. <https://doi.org/10.1093/molbev/msx146>.
- Nicastro, Daniela, J. Richard McIntosh, et Wolfgang Baumeister. 2005. « 3D Structure of Eukaryotic Flagella in a Quiescent State Revealed by Cryo-Electron Tomography ». *Proceedings of the National Academy of Sciences* 102 (44): 15889-94. <https://doi.org/10.1073/pnas.0508274102>.
- Niederlova, Veronika, Martin Modrak, Oksana Tsyklauri, Martina Huranova, et Ondrej Stepanek. 2019. « Meta-analysis of genotype-phenotype associations in Bardet-Biedl Syndrome uncovers differences among causative genes ». *Human Mutation* 0 (ja). <https://doi.org/10.1002/humu.23862>.

- Nikopoulos, Konstantinos, Katarina Cisarova, Mathieu Quinodoz, Hanna Koskiniemi-Kuendig, Noriko Miyake, Pietro Farinelli, Atta Ur Rehman, et al. 2019. « A Frequent Variant in the Japanese Population Determines Quasi-Mendelian Inheritance of Rare Retinal Ciliopathy ». *Nature Communications* 10 (1): 2884. <https://doi.org/10.1038/s41467-019-10746-4>.
- Nilsson, Per, et Takaomi C Saido. 2014. « Dual roles for autophagy: Degradation and secretion of Alzheimer's disease A $\beta$  peptide ». *Bioessays* 36 (6): 570-78. <https://doi.org/10.1002/bies.201400002>.
- Nonaka, S., Y. Tanaka, Y. Okada, S. Takeda, A. Harada, Y. Kanai, M. Kido, et N. Hirokawa. 1998. « Randomization of Left-Right Asymmetry Due to Loss of Nodal Cilia Generating Leftward Flow of Extraembryonic Fluid in Mice Lacking KIF3B Motor Protein ». *Cell* 95 (6): 829-37. [https://doi.org/10.1016/s0092-8674\(00\)81705-5](https://doi.org/10.1016/s0092-8674(00)81705-5).
- Nozaki, Shohei, Roiner Francisco Castro Araya, Yohei Katoh, et Kazuhisa Nakayama. 2019. « Requirement of IFT-B–BBSome Complex Interaction in Export of GPR161 from Cilia ». *Biology Open* 8 (9). <https://doi.org/10.1242/bio.043786>.
- O’Gorman, S., N. A. Dagenais, M. Qian, et Y. Marchuk. 1997. « Protamine-Cre Recombinase Transgenes Efficiently Recombine Target Sequences in the Male Germ Line of Mice, but Not in Embryonic Stem Cells ». *Proceedings of the National Academy of Sciences of the United States of America* 94 (26): 14602-7. <https://doi.org/10.1073/pnas.94.26.14602>.
- Ohashi, Yohei, Shirley Tremel, et Roger L. Williams. 2019. « VPS34 complexes from a structural perspective ». *Journal of Lipid Research* 60 (2): 229-41. <https://doi.org/10.1194/jlr.R089490>.
- Okonechnikov, Konstantin, Ana Conesa, et Fernando García-Alcalde. 2015. « Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data ». *Bioinformatics* 32 (2): 292-94. <https://doi.org/10.1093/bioinformatics/btv566>.
- O’Leary, Nuala A., Mathew W. Wright, J. Rodney Brister, Stacy Ciufu, Diana Haddad, Rich McVeigh, Bhanu Rajput, et al. 2015. « Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation ». *Nucleic Acids Research* 44 (D1): D733-45. <https://doi.org/10.1093/nar/gkv1189>.
- Ostertag, Eric M., John L. Goodier, Yue Zhang, et Haig H. Kazazian. 2003a. « SVA Elements Are Nonautonomous Retrotransposons That Cause Disease in Humans ». *The American Journal of Human Genetics* 73 (6): 1444-51. <https://doi.org/10.1086/380207>.
- Ostrowski, Lawrence E., Kevin Blackburn, Kristen M. Radde, Mary B. Moyer, Daniela M. Schlatzer, Arthur Moseley, et Richard C. Boucher. 2002. « A Proteomic Analysis of Human Cilia: Identification of Novel Components ». *Molecular & Cellular Proteomics: MCP* 1 (6): 451-65. <https://doi.org/10.1074/mcp.m200037-mcp200>.
- Otto, Edgar A, Bernhard Schermer, Tomoko Obara, John F O’Toole, Karl S Hiller, Adelheid M Mueller, Rainer G Ruf, et al. 2003. « Mutations in INVS encoding inversin cause nephronophthisis type 2, linking renal cystic disease to the function of primary cilia and left-right axis determination ». *Nature genetics* 34 (4): 413-20. <https://doi.org/10.1038/ng1217>.
- Otto, Gordon P, Birgit Rathkolb, Manuela A Oestereicher, Christoph J Lengger, Corinna Moerth, Kateryna Micklich, Helmut Fuchs, Valérie Gailus-Durner, Eckhard Wolf, et Martin Hrabě de Angelis. 2016. « Clinical Chemistry Reference Intervals for C57BL/6J, C57BL/6N, and C3HeB/FeJ Mice (Mus musculus) ». *Journal of the American Association for Laboratory Animal Science : JAALAS* 55 (4): 375-86.



- Oud, Machteld M., Ideke J. C. Lamers, et Heleen H. Arts. 2017. « Ciliopathies: Genetics in Pediatric Medicine ». *Journal of Pediatric Genetics* 6 (1): 18-29. <https://doi.org/10.1055/s-0036-1593841>.
- Pampliega, Olatz, Idil Orhon, Bindi Patel, Sunandini Sridhar, Antonio Díaz-Carretero, Isabelle Beau, Patrice Codogno, Birgit Satir, Peter Satir, et Ana Maria Cuervo. 2013. « Functional interaction between autophagy and ciliogenesis ». *Nature* 502 (7470): 194-200. <https://doi.org/10.1038/nature12639>.
- Parzych, Katherine R., et Daniel J. Klionsky. 2014. « An Overview of Autophagy: Morphology, Mechanism, and Regulation ». *Antioxidants & Redox Signaling* 20 (3): 460-73. <https://doi.org/10.1089/ars.2013.5371>.
- Pazour, G. J., B. L. Dickert, Y. Vucica, E. S. Seeley, J. L. Rosenbaum, G. B. Witman, et D. G. Cole. 2000. « Chlamydomonas IFT88 and Its Mouse Homologue, Polycystic Kidney Disease Gene Tg737, Are Required for Assembly of Cilia and Flagella ». *The Journal of Cell Biology* 151 (3): 709-18. <https://doi.org/10.1083/jcb.151.3.709>.
- Pechnick, Robert N, Vera M Chesnokova, Anastasia Kariagina, Shannon Price, Catherine J Bresee, et Russell E Poland. 2004. « Reduced Immobility in the Forced Swim Test in Mice with a Targeted Deletion of the Leukemia Inhibitory Factor (LIF) Gene ». *Neuropsychopharmacology* 29 (4): 770-76. <https://doi.org/10.1038/sj.npp.1300402>.
- Petit, Florence, Anne-Sophie Jourdain, Muriel Holder-Espinasse, Boris Keren, Joris Andrieux, Martine Duterque-Coquillaud, Nicole Porchet, Sylvie Manouvrier-Hanu, et Fabienne Escande. 2016. « The Disruption of a Novel Limb Cis-Regulatory Element of SHH Is Associated with Autosomal Dominant Preaxial Polydactyly-Hypertrichosis ». *European Journal of Human Genetics* 24 (1): 37-43. <https://doi.org/10.1038/ejhg.2015.53>.
- Picken, Maria M. 2020. « The Pathology of Amyloidosis in Classification: A Review ». *Acta Haematologica* 143 (4): 322-34. <https://doi.org/10.1159/000506696>.
- Prokic, Ivana, Belinda S. Cowling, et Jocelyn Laporte. 2014. « Amphiphysin 2 (BIN1) in Physiology and Diseases ». *Journal of Molecular Medicine* 92 (5): 453-63. <https://doi.org/10.1007/s00109-014-1138-1>.
- Pusapati, Ganesh V., Casey E. Hughes, Karolin V. Dorn, Dapeng Zhang, Priscilla Sugianto, L. Aravind, et Rajat Rohatgi. 2014. « EFCAB7 and IQCE Regulate Hedgehog Signaling by Tethering the EVC-EVC2 Complex to the Base of Primary Cilia ». *Developmental Cell* 28 (5): 483-96. <https://doi.org/10.1016/j.devcel.2014.01.021>.
- Quélin, Chloé, Philippe Loget, Lucile Boutaud, Nadia Elkhartoufi, Joelle Milon, Sylvie Odent, Mélanie Fradin, et al. 2018. « Loss of Function IFT27 Variants Associated with an Unclassified Lethal Fetal Ciliopathy with Renal Agenesis ». *American Journal of Medical Genetics. Part A* 176 (7): 1610-13. <https://doi.org/10.1002/ajmg.a.38685>.
- Raffan, E., et R. K. Semple. 2011. « Next Generation Sequencing--Implications for Clinical Practice ». *British Medical Bulletin* 99 (1): 53-71. <https://doi.org/10.1093/bmb/ldr029>.
- Raiz, Julija, Annette Damert, Sergiu Chira, Ulrike Held, Sabine Klawitter, Matthias Hamdorf, Johannes Löwer, Wolf H. Strätling, Roswitha Löwer, et Gerald G. Schumann. 2012a. « The Non-Autonomous Retrotransposon SVA Is Trans -Mobilized by the Human LINE-1 Protein Machinery ». *Nucleic Acids Research* 40 (4): 1666-83. <https://doi.org/10.1093/nar/gkr863>.
- Redin, Claire, Stéphanie Le Gras, Oussema Mhamdi, Véronique Geoffroy, Corinne Stoetzel, Marie-Claire Vincent, Pietro Chiurazzi, et al. 2012a. « Targeted high-throughput sequencing for diagnosis of genetically heterogeneous diseases: efficient mutation detection in Bardet-Biedl and Alström

- Syndromes ». *Journal of Medical Genetics* 49 (8): 502-12. <https://doi.org/10.1136/jmedgenet-2012-100875>.
- Reese, M. G., F. H. Eeckman, D. Kulp, et D. Haussler. 1997. « Improved Splice Site Detection in Genie ». *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology* 4 (3): 311-23. <https://doi.org/10.1089/cmb.1997.4.311>.
- Reiter, Jeremy F., et Michel R. Leroux. 2017. « Genes and molecular pathways underpinning ciliopathies ». *Nature reviews. Molecular cell biology* 18 (9): 533-47. <https://doi.org/10.1038/nrm.2017.60>.
- Richards, Sue, Nazneen Aziz, Sherri Bale, David Bick, Soma Das, Julie Gastier-Foster, Wayne W. Grody, et al. 2015. « Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology ». *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 17 (5): 405-24. <https://doi.org/10.1038/gim.2015.30>.
- Riddle, R. D., R. L. Johnson, E. Laufer, et C. Tabin. 1993. « Sonic Hedgehog Mediates the Polarizing Activity of the ZPA ». *Cell* 75 (7): 1401-16. [https://doi.org/10.1016/0092-8674\(93\)90626-2](https://doi.org/10.1016/0092-8674(93)90626-2).
- Roboti, Peristera, Keisuke Sato, et Martin Lowe. 2015. « The Golgin GMAP-210 Is Required for Efficient Membrane Trafficking in the Early Secretory Pathway ». *J Cell Sci* 128 (8): 1595-1606. <https://doi.org/10.1242/jcs.166710>.
- Saliba, Youakim, Jules-Joel Bakhos, Tarek Itani, et Nassim Farès. 2017. « An Optimized Protocol for Purification of Functional Islets of Langerhans ». *Laboratory Investigation* 97 (1): 70-83. <https://doi.org/10.1038/labinvest.2016.123>.
- Sanchez-Navarro, Iker, Luciana R. J. da Silva, Fiona Blanco-Kelly, Olga Zurita, Noelia Sanchez-Bolivar, Cristina Villaverde, Maria Isabel Lopez-Molina, et al. 2018. « Combining Targeted Panel-Based Resequencing and Copy-Number Variation Analysis for the Diagnosis of Inherited Syndromic Retinopathies and Associated Ciliopathies ». *Scientific Reports* 8 (1): 5285. <https://doi.org/10.1038/s41598-018-23520-1>.
- Sasaguri, Hiroki, Per Nilsson, Shoko Hashimoto, Kenichi Nagata, Takashi Saito, Bart De Strooper, John Hardy, Robert Vassar, Bengt Winblad, et Takaomi C Saido. « APP mouse models for Alzheimer's disease preclinical studies ». *The EMBO Journal* 36, n° 17 (1 septembre 2017): 2473-87. <https://doi.org/10.15252/embj.201797397>.
- Sayers, Eric W., Jeff Beck, J. Rodney Brister, Evan E. Bolton, Kathi Canese, Donald C. Comeau, Kathryn Funk, et al. 2019. « Database Resources of the National Center for Biotechnology Information. » *Nucleic Acids Research*, octobre. <https://doi.org/10.1093/nar/gkz899>.
- Scheidecker, Sophie, Christelle Etard, Laurence Haren, Corinne Stoetzel, Sarah Hull, Gavin Arno, Vincent Plagnol, et al. 2015. « Mutations in TUBGCP4 Alter Microtubule Organization via the  $\gamma$ -Tubulin Ring Complex in Autosomal-Recessive Microcephaly with Chorioretinopathy ». *American Journal of Human Genetics* 96 (4): 666-74. <https://doi.org/10.1016/j.ajhg.2015.02.011>.
- Scheidecker, Sophie, Christelle Etard, Nathan W. Pierce, Véronique Geoffroy, Elise Schaefer, Jean Muller, Kirsley Chennen, et al. 2014. « Exome Sequencing of Bardet-Biedl Syndrome Patient Identifies a Null Mutation in the BBSome Subunit BBIP1 (BBS18) ». *Journal of Medical Genetics* 51 (2): 132-36. <https://doi.org/10.1136/jmedgenet-2013-101785>.



- Schofield, Paul N., Robert Hoehndorf, et Georgios V. Gkoutos. 2012. « Mouse genetic and phenotypic resources for human genetics ». *Human mutation* 33 (5): 826-36. <https://doi.org/10.1002/humu.22077>.
- Semple, Bridgette D., Klas Blomgren, Kayleen Gimlin, Donna M. Ferriero, et Linda J. Noble-Haeusslein. 2013. « Brain development in rodents and humans: Identifying benchmarks of maturation and vulnerability to injury across species ». *Progress in neurobiology* 0: 1-16. <https://doi.org/10.1016/j.pneurobio.2013.04.001>.
- Seo, Seongjin, Qihong Zhang, Kevin Bugge, David K. Breslow, Charles C. Searby, Maxence V. Nachury, et Val C. Sheffield. 2011. « A Novel Protein LZTFL1 Regulates Ciliary Trafficking of the BBSome and Smoothed ». *PLoS Genetics* 7 (11). <https://doi.org/10.1371/journal.pgen.1002358>.
- Shaheen, Ranad, Katarzyna Szymanska, Basudha Basu, Nisha Patel, Nour Ewida, Eissa Faqeih, Amal Al Hashem, et al. 2016. « Characterizing the morbid genome of ciliopathies ». *Genome Biology* 17 (1): 242. <https://doi.org/10.1186/s13059-016-1099-5>.
- Shamseldin, Hanan E., Ranad Shaheen, Nour Ewida, Dalal K. Bubshait, Hisham Alkuraya, Elham Almardawi, Ali Howaidi, et al. 2020a. « The morbid genome of ciliopathies: an update ». *Genetics in Medicine*, février. <https://doi.org/10.1038/s41436-020-0761-1>.
- Shapiro, M. B., et P. Senapathy. 1987. « RNA Splice Junctions of Different Classes of Eukaryotes: Sequence Statistics and Functional Implications in Gene Expression ». *Nucleic Acids Research* 15 (17): 7155-74. <https://doi.org/10.1093/nar/15.17.7155>.
- Simón, M. Victoria, Facundo H. Prado Spalm, Marcela S. Vera, et Nora P. Rotstein. 2019. « Sphingolipids as Emerging Mediators in Retina Degeneration ». *Frontiers in Cellular Neuroscience* 13. <https://doi.org/10.3389/fncel.2019.00246>.
- Singh, Pratibha, Palaniyandi Ramanan, et Priti Talwar. 2016. « Death Associated Protein Kinase 1 (DAPK1): A Regulator of Apoptosis and Autophagy ». *Frontiers in Molecular Neuroscience* 9 (juin). <https://doi.org/10.3389/fnmol.2016.00046>.
- Singh, Sandeep K, Miao Gui, Fujiet Koh, Matthew CJ Yip, et Alan Brown. 2020a. « Structure and activation mechanism of the BBSome membrane protein trafficking complex ». Édité par Andrew P Carter, John Kuriyan, Esben Lorentzen, et Masahide Kikkawa. *eLife* 9 (janvier): e53322. <https://doi.org/10.7554/eLife.53322>.
- Singh, Sandeep K., Miao Gui, Fujiet Koh, Matthew Cj Yip, et Alan Brown. 2020b. « Structure and Activation Mechanism of the BBSome Membrane Protein Trafficking Complex ». *ELife* 9 (janvier). <https://doi.org/10.7554/eLife.53322>.
- Sittig, Laura J., Peter Carbonetto, Kyle A. Engel, Kathleen S. Krauss, Camila M. Barrios-Camacho, et Abraham A. Palmer. 2016. « Genetic background limits generalizability of genotype-phenotype relationships ». *Neuron* 91 (6): 1253-59. <https://doi.org/10.1016/j.neuron.2016.08.013>.
- Sjöstrand, Fritiof S. 1953. « The Ultrastructure of the Inner Segments of the Retinal Rods of the Guinea Pig Eye as Revealed by Electron Microscopy ». *Journal of Cellular and Comparative Physiology* 42 (1): 45-70. <https://doi.org/10.1002/jcp.1030420104>.
- Smit, AFA, R Huble, et P Green. 2013. « RepeatMasker Open-4.0 », 2015 2013. <http://www.repeatmasker.org>.
- Stainier, Didier Y. R., Erez Raz, Nathan D. Lawson, Stephen C. Ekker, Rebecca D. Burdine, Judith S. Eisen, Philip W. Ingham, et al. 2017. « Guidelines for morpholino use in zebrafish ». *PLoS Genetics* 13 (10). <https://doi.org/10.1371/journal.pgen.1007000>.

- Stemers, Frank J, Weihua Chang, Grace Lee, David L Barker, Richard Shen, et Kevin L Gunderson. 2006. « Whole-genome genotyping with the single-base extension assay ». *Nature Methods* 3 (1): 31-33. <https://doi.org/10.1038/nmeth842>.
- Stewart, Chip, Deniz Kural, Michael P Strömberg, Jerilyn A Walker, Miriam K Konkel, Adrian M Stütz, Alexander E Urban, et al. 2011. « A Comprehensive Map of Mobile Element Insertion Polymorphisms in Humans ». *PLoS Genetics* 7 (8): e1002236-e1002236. <https://doi.org/10.1371/journal.pgen.1002236>.
- Stewart, Colin L., Petr Kaspar, Lisa J. Brunet, Harshida Bhatt, Inder Gadi, Frank Köntgen, et Susan J. Abbondanzo. 1992. « Blastocyst Implantation Depends on Maternal Expression of Leukaemia Inhibitory Factor ». *Nature* 359 (6390): 76-79. <https://doi.org/10.1038/359076a0>.
- Stoetzel, Corinne, Séverine Bär, Johan-Owen De Craene, Sophie Scheidecker, Christelle Etard, Johana Chicher, Jennifer R. Reck, et al. 2016. « A Mutation in VPS15 (PIK3R4) Causes a Ciliopathy and Affects IFT20 Release from the Cis-Golgi ». *Nature Communications* 7 (1): 13586. <https://doi.org/10.1038/ncomms13586>.
- Stoetzel, Corinne, Virginie Laurier, Erica E. Davis, Jean Muller, Suzanne Rix, José L. Badano, Carmen C. Leitch, et al. 2006. « BBS10 Encodes a Vertebrate-Specific Chaperonin-like Protein and Is a Major BBS Locus ». *Nature Genetics* 38 (5): 521-24. <https://doi.org/10.1038/ng1771>.
- Strand, Andrew D, Aaron K Aragaki, Zachary C Baquet, Angela Hodges, Philip Cunningham, Peter Holmans, Kevin R Jones, Lesley Jones, Charles Kooperberg, et James M Olson. 2007. « Conservation of Regional Gene Expression in Mouse and Human Brain ». *PLoS Genetics* 3 (4). <https://doi.org/10.1371/journal.pgen.0030059>.
- Sudmant, Peter H., Tobias Rausch, Eugene J. Gardner, Robert E. Handsaker, Alexej Abyzov, John Huddleston, Yan Zhang, et al. 2015. « An integrated map of structural variation in 2,504 human genomes ». *Nature* 526 (7571): 75-81. <https://doi.org/10.1038/nature15394>.
- Suspitsin, Evgeny, et Evgeny Imyanitov. 2016. « Bardet-Biedl Syndrome ». *Molecular Syndromology* 7 (avril). <https://doi.org/10.1159/000445491>.
- Takasu, M., R. Hayashi, E. Maruya, M. Ota, K. Imura, K. Kougo, C. Kobayashi, et al. 2007. « Deletion of Entire HLA-A Gene Accompanied by an Insertion of a Retrotransposon ». *Tissue Antigens* 70 (2): 144-50. <https://doi.org/10.1111/j.1399-0039.2007.00870.x>.
- Taniguchi-Ikeda, Mariko, Kazuhiro Kobayashi, Motoi Kanagawa, Chih-chieh Yu, Kouhei Mori, Tetsuya Oda, Atsushi Kuga, et al. 2011. « Pathogenic Exon-Trapping by SVA Retrotransposon and Rescue in Fukuyama Muscular Dystrophy ». *Nature* 478 (7367): 127-31. <https://doi.org/10.1038/nature10456>.
- Tanner, Rikki M., Todd M. Brown, et Paul Muntner. 2012. « Epidemiology of Obesity, the Metabolic Syndrome, and Chronic Kidney Disease ». *Current Hypertension Reports* 14 (2): 152-59. <https://doi.org/10.1007/s11906-012-0254-y>.
- Taşkesen, Mustafa, Gayle B Collin, Alexei V Evsikov, Ayşegül Güzel, R Köksal Özgül, Jan D Marshall, et Jürgen K Naggert. 2012. « Novel Alu Retrotransposon Insertion Leading to Alström Syndrome ». *Human Genetics* 131 (3): 407-13. <https://doi.org/10.1007/s00439-011-1083-9>.
- Tavares, Erika, Chen Yu Tang, Anjali Vig, Shuning Li, Gail Billingsley, Wilson Sung, Ajoy Vincent, Bhooma Thiruvahindrapuram, et Elise Héon. 2019a. « Retrotransposon Insertion as a Novel Mutational Event in Bardet-Biedl Syndrome ». *Molecular Genetics & Genomic Medicine* 7 (2): e00521. <https://doi.org/10.1002/mgg3.521>.

- The 1000 Genomes Project Consortium, Richard A. Gibbs, Eric Boerwinkle, Harsha Doddapaneni, Yi Han, Viktoriya Korchina, Christie Kovar, et al. 2015. « A Global Reference for Human Genetic Variation ». *Nature* 526 (7571): 68-74. <https://doi.org/10.1038/nature15393>.
- Thompson, S., A. R. Clarke, A. M. Pow, M. L. Hooper, et D. W. Melton. 1989. « Germ Line Transmission and Expression of a Corrected HPRT Gene Produced by Gene Targeting in Embryonic Stem Cells ». *Cell* 56 (2): 313-21. [https://doi.org/10.1016/0092-8674\(89\)90905-7](https://doi.org/10.1016/0092-8674(89)90905-7).
- Thorvaldsdottir, H., J. T. Robinson, et J. P. Mesirov. 2013. « Integrative Genomics Viewer (IGV): High-Performance Genomics Data Visualization and Exploration ». *Briefings in Bioinformatics* 14 (2): 178-92. <https://doi.org/10.1093/bib/bbs017>.
- Thung, Djie Tjwan, Joep de Ligt, Lisenka EM Vissers, Marloes Steehouwer, Mark Kroon, Petra de Vries, Eline P. Slagboom, Kai Ye, Joris A. Veltman, et Jayne Y. Hehir-Kwa. 2014. « Mobster: accurate detection of mobile element insertions in next generation sequencing data ». *Genome Biology* 15 (10): 488. <https://doi.org/10.1186/s13059-014-0488-x>.
- Tickle, Cheryll, et Matthew Towers. 2017. « Sonic Hedgehog Signaling in Limb Development ». *Frontiers in Cell and Developmental Biology* 5: 14. <https://doi.org/10.3389/fcell.2017.00014>.
- Torene, Rebecca I., Kevin Galens, Shuxi Liu, Kevin Arvai, Carlos Borroto, Julie Scuffins, Zhancheng Zhang, et al. 2020. « Mobile element insertion detection in 89,874 clinical exomes ». *Genetics in Medicine*, janvier. <https://doi.org/10.1038/s41436-020-0749-x>.
- Twohig, Daniel, et Henrietta M. Nielsen. 2019. «  $\alpha$ -synuclein in the pathophysiology of Alzheimer's disease ». *Molecular Neurodegeneration* 14 (1): 23. <https://doi.org/10.1186/s13024-019-0320-x>.
- Uguen, Kévin, Claire Jubin, Yannis Duffourd, Claire Bardel, Valérie Malan, Jean-Michel Dupont, Laila El Khattabi, et al. 2020. « Genome Sequencing in Cytogenetics: Comparison of Short-Read and Linked-Read Approaches for Germline Structural Variant Detection and Characterization ». *Molecular Genetics & Genomic Medicine* 8 (3): e1114-e1114. <https://doi.org/10.1002/mgg3.1114>.
- Umair, Muhammad, Khadim Shah, Bader Alhaddad, Tobias B. Haack, Elisabeth Graf, Tim M. Strom, Thomas Meitinger, et Wasim Ahmad. 2017. « Exome Sequencing Revealed a Splice Site Variant in the IQCE Gene Underlying Post-Axial Polydactyly Type A Restricted to Lower Limb ». *European Journal of Human Genetics: EJHG* 25 (8): 960-65. <https://doi.org/10.1038/ejhg.2017.83>.
- Vargas-Poussou, Rosa, Pascal Houillier, Nelly Le Pottier, Laurence Strompf, Chantal Loirat, Véronique Baudouin, Marie-Alice Macher, et al. 2006. « Genetic Investigation of Autosomal Recessive Distal Renal Tubular Acidosis: Evidence for Early Sensorineural Hearing Loss Associated with Mutations in the ATP6V0A4 Gene ». *Journal of the American Society of Nephrology: JASN* 17 (5): 1437-43. <https://doi.org/10.1681/ASN.2005121305>.
- Veleri, Shobi, Csilla H. Lazar, Bo Chang, Paul A. Sieving, Eyal Banin, et Anand Swaroop. 2015. « Biology and Therapy of Inherited Retinal Degenerative Disease: Insights from Mouse Models ». *Disease Models & Mechanisms* 8 (2): 109-29. <https://doi.org/10.1242/dmm.017913>.
- Vogt, Julia, Kathrin Bengesser, Kathleen BM Claes, Katharina Wimmer, Victor-Felix Mautner, Rick van Minkelen, Eric Legius, et al. 2014. « SVA retrotransposon insertion-associated deletion represents a novel mutational mechanism underlying large genomic copy number changes with non-recurrent breakpoints ». *Genome Biology* 15 (6): R80. <https://doi.org/10.1186/gb-2014-15-6-r80>.
- Waltz, Florent, Tan-Trung Nguyen, Mathilde Arrivé, Anthony Boehler, Johana Chicher, Philippe Hammann, Lauriane Kuhn, et al. 2019. « Small Is Big in Arabidopsis Mitochondrial Ribosome ». *Nature Plants* 5 (1): 106-17. <https://doi.org/10.1038/s41477-018-0339-y>.

- Wang, Hui, Jinchuan Xing, Deepak Grover, Dale J. Hedges, Kyudong Han, Jerilyn A. Walker, et Mark A. Batzer. 2005. « SVA Elements: A Hominid-Specific Retroposon Family ». *Journal of Molecular Biology* 354 (4): 994-1007. <https://doi.org/10.1016/j.jmb.2005.09.085>.
- Wang, Jianxin, Lei Song, Deepak Grover, Sami Azrak, Mark A Batzer, et Ping Liang. 2006. « DbRIP: A Highly Integrated Database of Retrotransposon Insertion Polymorphisms in Humans ». *Human Mutation* 27 (4): 323-29. <https://doi.org/10.1002/humu.20307>.
- Waters, Aoife M., et Philip L. Beales. 2011. « Ciliopathies: An Expanding Disease Spectrum ». *Pediatric Nephrology (Berlin, Germany)* 26 (7): 1039-56. <https://doi.org/10.1007/s00467-010-1731-7>.
- Wei, Qun, Yongxia Chen, Yi-Feng Gu, et Wenhe Zhao. 2019. « Molecular Characterization and Functional Analysis of Leucine Zipper Transcription Factor Like 1 in Zebrafish (Danio Rerio) ». *Frontiers in Physiology* 10. <https://doi.org/10.3389/fphys.2019.00801>.
- Wen, Xiaohui, Cary K. Lai, Marie Evangelista, Jo-Anne Hongo, Frederic J. de Sauvage, et Suzie J. Scales. 2010. « Kinetics of Hedgehog-Dependent Full-Length Gli3 Accumulation in Primary Cilia and Subsequent Degradation ». *Molecular and Cellular Biology* 30 (8): 1910-22. <https://doi.org/10.1128/MCB.01089-09>.
- Westerfield, Monte. *The Zebrafish Book: A Guide for the Laboratory Use of Zebrafish (Danio Rerio)*. University of Oregon Press, 2000.
- Wheway, Gabrielle, Liliya Nazlamova, et John T. Hancock. 2018. « Signaling through the Primary Cilium ». *Frontiers in Cell and Developmental Biology* 6. <https://doi.org/10.3389/fcell.2018.00008>.
- Williams, Amy L., Nick Patterson, Joseph Glessner, Hakon Hakonarson, et David Reich. 2012. « Phasing of Many Thousands of Genotyped Samples ». *The American Journal of Human Genetics* 91 (2): 238-51. <https://doi.org/10.1016/j.ajhg.2012.06.013>.
- Wilund, Kenneth R., Ming Yi, Filomena Campagna, Marcello Arca, Giovanni Zuliani, Renato Fellin, Yiu-Kee Ho, J. Victor Garcia, Helen H. Hobbs, et Jonathan C. Cohen. 2002. « Molecular Mechanisms of Autosomal Recessive Hypercholesterolemia ». *Human Molecular Genetics* 11 (24): 3019-30. <https://doi.org/10.1093/hmg/11.24.3019>.
- Wingfield, Jenna L., Karl-Ferdinand Lechtreck, et Esben Lorentzen. 2018. « Trafficking of ciliary membrane proteins by the intraflagellar transport/BBSome machinery ». *Essays in biochemistry* 62 (6): 753-63. <https://doi.org/10.1042/EBC20180030>.
- Witman, G. B., K. Carlson, J. Berliner, et J. L. Rosenbaum. 1972. « Chlamydomonas Flagella. I. Isolation and Electrophoretic Analysis of Microtubules, Matrix, Membranes, and Mastigonemes ». *The Journal of Cell Biology* 54 (3): 507-39. <https://doi.org/10.1083/jcb.54.3.507>.
- Wormser, Ohad, Libe Gradstein, Yuval Yogev, Yonatan Perez, Rotem Kadir, Inna Goliand, Yair Sadka, et al. 2019. « SCAPER localizes to primary cilia and its mutation affects cilia length, causing Bardet-Biedl syndrome ». *European Journal of Human Genetics* 27 (6): 928-40. <https://doi.org/10.1038/s41431-019-0347-z>.
- Wu, Yung-Hsuan, Ting-Fang Chou, Leslie Young, Fu-Yi Hsieh, Hsuan-Yin Pan, Shu-Ting Mo, Shani Bialik Brown, Ruey-Hwa Chen, Adi Kimchi, et Ming-Zong Lai. 2020. « Tumor Suppressor Death-Associated Protein Kinase 1 Inhibits Necroptosis by P38 MAPK Activation ». *Cell Death & Disease* 11 (5): 1-15. <https://doi.org/10.1038/s41419-020-2534-9>.
- Xu, Caixia, Xiaoming Yang, Hang Zhou, Yongyong Li, Chao Xing, Taifeng Zhou, Dongmei Zhong, et al. 2020. « A novel ZRS variant causes preaxial polydactyly type I by increased sonic hedgehog expression

- in the developing limb bud ». *Genetics in Medicine* 22 (1): 189-98. <https://doi.org/10.1038/s41436-019-0626-7>.
- Yabuki, Akira, Syusaku Suzuki, Mitsuharu Matsumoto, et Hayao Nishinakagawa. 1999. « Morphometrical Analysis of Sex and Strain Differences in the Mouse Nephron. » *Journal of Veterinary Medical Science* 61 (8): 891-96. <https://doi.org/10.1292/jvms.61.891>.
- Yardeni, Tal, Michael Eckhaus, H. Douglas Morris, Marjan Huizing, et Shelley Hoogstraten-Miller. 2011. « Retro-orbital injections in mice ». *Lab animal* 40 (5): 155-60. <https://doi.org/10.1038/labano511-155>.
- Yen, Hsan-jan. s. d. « Development of the Zebrafish as a Model for Bardet-Biedl Syndrome », 171.
- Yeo, Gene, et Christopher B. Burge. 2004. « Maximum Entropy Modeling of Short Sequence Motifs with Applications to RNA Splicing Signals ». *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology* 11 (2-3): 377-94. <https://doi.org/10.1089/1066527041410418>.
- Yıldız Bölükbaşı, Esra, Sara Mumtaz, Muhammad Afzal, Ute Woehlbier, Sajid Malik, et Aslihan Tolun. 2018. « Homozygous mutation in *CEP19*, a gene mutated in morbid obesity, in Bardet-Biedl syndrome with predominant postaxial polydactyly ». *Journal of Medical Genetics* 55 (3): 189. <https://doi.org/10.1136/jmedgenet-2017-104758>.
- Youn, Yong Ha, et Young-Goo Han. 2018. « Primary Cilia in Brain Development and Diseases ». *The American Journal of Pathology* 188 (1): 11-22. <https://doi.org/10.1016/j.ajpath.2017.08.031>.
- Zaghloul, Norann A., et Nicholas Katsanis. 2010. « Functional modules, mutational load and human genetic disease ». *Trends in Genetics* 26 (4): 168-76. <https://doi.org/10.1016/j.tig.2010.01.006>.
- Zerbino, Daniel R., et Ewan Birney. 2008. « Velvet: Algorithms for de Novo Short Read Assembly Using de Bruijn Graphs. » *Genome Research* 18 (5): 821-29. <https://doi.org/10.1101/gr.074492.107>.
- Zhang, Xiao-jie, Sheng Chen, Kai-xing Huang, et Wei-dong Le. 2013. « Why should autophagic flux be assessed? » *Acta Pharmacologica Sinica* 34 (5): 595-99. <https://doi.org/10.1038/aps.2012.184>.
- Zhao, Xinyu, et Anita Bhattacharyya. 2018. « Human Models Are Needed for Studying Human Neurodevelopmental Disorders ». *The American Journal of Human Genetics* 103 (6): 829-57. <https://doi.org/10.1016/j.ajhg.2018.10.009>.
- Zimmermann, K. W. 1898. « Beiträge zur Kenntniss einiger Drüsen und Epithelien ». *Archiv für mikroskopische Anatomie* 52 (3): 552-706. <https://doi.org/10.1007/BF02975837>.





# Du génome humain au modèle murin : Compréhension des ciliopathies

Laboratoire de Génétique médicale, INSERM UMR-S 1112, Strasbourg

## Le résumé de thèse

Les ciliopathies sont des maladies rares causées par une atteinte du cil primaire pour lesquelles au moins 20% des patients n'ont pas de diagnostic génétique résolu. Le travail présenté dans cette thèse expose le panel d'expériences mis en place au Laboratoire de Génétique Médicale afin d'identifier les gènes responsables de ciliopathies, réaliser les validations fonctionnelles des mutations identifiées et comprendre les mécanismes physiopathologiques sous-jacents en mettant en place des études *in vitro* et *in vivo*. Au cours de ces 3 années, j'ai été amenée à développer plusieurs projets de recherche tous inscrits dans la recherche translationnelle. Dans une 1<sup>ère</sup> partie, j'aborderais les projets portant sur l'identification et la validation fonctionnelle de mutations pour des gènes responsables de ciliopathies comme *BBS19/IFT27*, *BBS1*, *IQCE* et *BBS5*. Dans une 2<sup>ème</sup> partie, je rapporterais les expériences ayant permis la caractérisation phénotypique du nouveau modèle murin *Vps15<sup>R998Q</sup>*.

Les résultats obtenus au cours de cette thèse ont permis de résoudre les cas de 11 familles, de mieux comprendre la physiopathologie de ces maladies et d'identifier un potentiel nouveau modèle murin pour des maladies plus communes.

**Mot clés :** Ciliopathies, Syndrome de Bardet-Biedl (BBS), *BBS19/IFT27*, *BBS1*, *IQCE*, *BBS5*, *Vps15*

## The thesis summary

Ciliopathies are rare diseases caused by primary cilia abnormalities for which at least 20% of patients are still genetically unsolved. The work of this thesis shows all experiments implemented at the Laboratory of Human Genetics to identify ciliopathy causative genes, to achieve functional validations on identified mutations and to better understand underlying physiopathological mechanisms by setting up *in vitro* and *in vivo* experimental studies. During these 3 years, I have been implicated in several projects all included in translational research. In the first part of this manuscript, I will present projects related to identification and functional validations of mutations in known ciliopathy genes such as *BBS19/IFT27*, *BBS1*, *IQCE* and *BBS5*. In the second part, I report experiments that have enabled the phenotypic characterisation of the new mice model *Vps15<sup>R998Q</sup>*.

Overall, I contributed in the identification of the molecular causes of 11 family cases, to enhance our knowledge on some ciliopathy physiopathologies mechanism and to identify a new potential mice model for common diseases.

**Key words :** Ciliopathies, Bardet Biedl Syndrome (BBS), *BBS19/IFT27*, *BBS1*, *IQCE*, *BBS5*, *Vps15*