



**HAL**  
open science

# Étude de l'impact des variants génétiques rares sur l'obésité monogénique

Lise Folon

► **To cite this version:**

Lise Folon. Étude de l'impact des variants génétiques rares sur l'obésité monogénique. Médecine humaine et pathologie. Université de Lille, 2023. Français. NNT : 2023ULILS059 . tel-04444175

**HAL Id: tel-04444175**

**<https://theses.hal.science/tel-04444175>**

Submitted on 7 Feb 2024

**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.



**Université de Lille**

**Ecole Doctorale Biologie-Santé de Lille**

Année universitaire 2022-2023

Thèse soutenue en vue de l'obtention du titre de

**Docteur de l'Université de Lille**

Discipline : *génétique, génomique, bioinformatique*

Spécialité : *aspects moléculaires et cellulaires de la biologie*

**Etude de l'impact des variants génétiques rares sur  
l'obésité monogénique**

Par **Lise FOLON**

Sous la direction du Docteur Amélie BONNEFOND et du Docteur  
Morgane BARON

Soutenue le 19 décembre 2023 devant le jury composé de :

Professeur Karine CLEMENT

Rapporteuse

Docteur Ralf JOCKERS

Rapporteur, Président du jury

Docteur Muriel COUPAYE

Examinatrice

Docteur Morgane BARON

Co-directrice de thèse

Docteur Amélie BONNEFOND

Directrice de thèse



## Remerciements

Je tiens en premier lieu à remercier le Dr Ralf Jockers, le Pr Karine Clément et le Dr Muriel Coupaye d'avoir accepté de juger ce travail et de faire partie de mon jury de thèse.

J'adresse également mes remerciements au Pr Philippe Froguel pour m'avoir permis d'intégrer l'UMR 1283/8199 et d'y réaliser ce travail de thèse.

Un immense merci au Dr Amélie Bonnefond et au Dr Morgane Baron, qui m'ont accueillie et encadrée tout au long de ces 3 années de thèse. Merci de m'avoir donné votre confiance, ce fut un véritable plaisir de travailler avec vous. Amélie, merci pour tout ce que tu m'as appris, de m'avoir fait découvrir le monde fabuleux de la génétique et pour toutes les opportunités que tu m'as données. Morgane, merci pour ton soutien, tes précieux conseils et pour m'avoir permis d'explorer le domaine de l'enseignement.

Je souhaite exprimer ma reconnaissance envers le conseil régional des Hauts-de-France pour avoir fourni les moyens nécessaires à la réalisation de ce projet.

Je voudrais également exprimer ma gratitude envers le Pr Michel Salzet pour avoir permis la réalisation des analyses de spectrométrie de masse dans son laboratoire PRISM, ainsi qu'à Soulaïmane pour son aide précieuse dans la réalisation et les analyses de ces manipulations.

Je tiens à remercier chaleureusement toutes les équipes de l'UMR1283/8199, sans lesquelles ce travail n'aurait jamais pu voir le jour. Je tiens à remercier l'équipe NGS : Béné, Manu, Fred, Timothée, Freddie, Aurélie et Raphy. Mes remerciements vont également à Mehdi et à toute l'équipe de bioinformatique ainsi qu'aux statisticiens Mathilde, Lijiao, Aboubacar et Mickaël pour leur contribution essentielle aux analyses statistiques des tests fonctionnels. Je voudrais également remercier tous les membres de l'équipe fonctionnelle pour leur aide et leurs conseils avisés. Merci particulièrement à Hélène et Victoria pour leur aide précieuse dans le travail de mutagenèse et des analyses fonctionnelles. Merci à Audrey pour son aide sur les expériences d'immunofluorescence. Je n'oublie pas de remercier Nico et VV ainsi que Mélanie, Hélène et Constance pour leur aide et leur disponibilité au moindre petit tracas. Marlène, un merci tout particulier pour tout ce que tu m'as appris, et pour la formation d'excellence que tu m'as offerte.

Merci à toute l'équipe « Baby » : Corentin, Béné, Naomie, Audrey, Fred, Timothée et Mehdi, pour ces chouettes pauses repas et ces parties endiablées. L'entraînement va pouvoir reprendre de plus belle !

Je tiens à remercier plus particulièrement Victoria, Judith, Sarah, Justine et Claire, bien plus que des collègues, vous êtes devenues des véritables amies. Vous avez égayé mes journées, les remplissant de fous rires et de bons moments, ainsi que d'un soutien et d'une écoute inestimable dans les moments plus compliqués.

Merci à mes parents, ma sœur et mon frère de me supporter et de m'encourager depuis ces nombreuses années dans tous mes défis. Votre confiance et vos encouragements m'ont poussée à aller toujours plus loin. Merci à mes amis de m'avoir accompagné durant ces années et qui ont tous contribué d'une façon ou d'une autre à l'aboutissement de ce travail. Enfin, merci Tom, pour ta bonne humeur, ta patience et ton soutien qui m'a toujours permis de me relever quand je perdais courage.

Merci à tous ceux qui ont contribué d'une façon ou d'une autre à l'aboutissement de ce travail. Cette thèse n'aurait pas pu être achevée sans vous.

## Résumé

L'obésité est une maladie multifactorielle complexe avec une forte composante génétique. Contrairement à l'obésité commune qui est une maladie polygénique, les formes d'obésité monogénique sont causées par un seul variant génétique rare avec un effet fort et délétère. Ces formes sont rares, précoces et généralement sévères, affectant environ 5% des individus atteints d'obésité. La plupart des mutations rares liées aux formes monogéniques de l'obésité ont été identifiées dans les gènes de la voie leptine-mélanocortine, qui est essentielle pour la régulation de la prise alimentaire. L'identification de ces gènes est cruciale dans la compréhension de la physiopathologie de l'obésité et l'élaboration de nouveaux traitements.

J'ai tout d'abord étudié les variants rares hétérozygotes du gène *PCSK1* qui code l'enzyme prohormone convertase 1 (PC1/3). PC1/3 est impliquée dans la voie leptine-mélanocortine. Les mutations bialléliques de *PCSK1* causent une obésité précoce avec une endocrinopathie sévère. Les patients déficients en *PCSK1* (hétérozygotes ou homozygotes) peuvent dorénavant être traités avec des injections de setmélanotide pour perdre du poids. Or l'impact des variants rares hétérozygotes de *PCSK1* sur l'obésité, et leur pertinence en médecine de précision sont encore mal définis. Dans l'étude RaDiO incluant 9320 participants, 65 variants rares hétérozygotes de *PCSK1* ont été identifiés et analysés *in vitro*. Ces variants ont été classés en cinq groupes suivant la sévérité de leur impact sur l'activité enzymatique de PC1/3. Les résultats d'analyses d'association ont révélé que les variants rares induisant une perte totale de fonction augmentaient fortement le risque d'obésité et l'indice de masse corporelle (IMC) alors que les variants des autres groupes avec un effet partiel ou neutre sur l'activité de PC1/3 n'avaient aucun effet sur l'adiposité. Nous avons observé que les outils de prédiction *in silico* étaient peu fiables pour détecter les mutations entraînant une perte totale de fonction.

Dans un second temps, je me suis intéressée aux variants rares du gène *DYRK1B*. Bien que ce gène ne soit pas impliqué dans la voie leptine-mélanocortine, des variants pathogènes de *DYRK1B* ont été décrits chez plusieurs patients avec une obésité centrale, un diabète de type 2 (DT2) et une maladie des artères coronaires. Cependant, l'impact des variants rares de *DYRK1B* n'a pas été évalué à grande échelle. A partir de l'étude RaDiO incluant 9353 participants, 65 variants rares dans *DYRK1B* ont été détectés. Après une analyse *in vitro* de chaque variant, nous avons identifié 20 variants pathogènes ou probablement pathogènes (P/LP) d'après les critères de l'*American College of Medical Genetics and Genomics*. Parmi ces variants P/LP, six montraient un effet entraînant une perte totale de fonction de DYRK1B (P/LP-full). Les

analyses d'association ont montré que les variants P/LP-full de *DYRK1B* étaient fortement associés à une augmentation de l'IMC et de la glycémie à jeun, et à un risque accru d'obésité et de DT2, alors que les variants P/LP n'avaient qu'un effet modeste sur l'adiposité et aucun impact sur l'homéostasie glucidique.

En conclusion, l'utilisation de la génétique fonctionnelle a permis de mettre en évidence que seuls les variants hétérozygotes de *PCSK1* et *DYRK1B* avec une perte de fonction totale causaient une obésité monogénique. Pour *DYRK1B*, l'obésité est en plus associée à un DT2. Ces résultats soulignent l'importance cruciale de déterminer *in vitro* l'impact fonctionnel des mutations en vue du diagnostic génétique et de l'éventuel choix d'un traitement approprié. Nous avons démontré que les tests de prédiction *in silico* ne sont pour l'heure pas assez précis.

## Abstract

Obesity is a complex multifactorial disease with a strong genetic component. Unlike common obesity, which is a polygenic disease, monogenic forms of obesity are caused by a single rare genetic variant with a strong and deleterious effect. These monogenic forms are rare, early-onset and generally very severe, affecting around 5% of individuals with obesity. Most rare mutations associated with monogenic obesity are found in genes within the leptin-melanocortin pathway, which is crucial for the regulation of food intake. Identifying these genes is crucial for understanding the pathophysiology of obesity and developing new treatments.

I initially studied rare heterozygous variants of the *PCSK1* gene, which encodes the prohormone convertase 1 (PC1/3) enzyme. PC1/3 is involved in the leptin-melanocortin pathway. Biallelic mutations in *PCSK1* cause early-onset obesity with severe endocrinopathy. Patients with *PCSK1* deficiency (heterozygous or homozygous) can now be treated with setmelanotide injections to promote weight loss. However, the impact of rare heterozygous variants of *PCSK1* on obesity and their relevance in precision medicine are still not well-defined. In the RaDiO study, which included 9,320 participants, 65 rare heterozygous variants of *PCSK1* were identified and assessed *in vitro*. These variants were classified into five groups based on the severity of their impact on the enzymatic activity of PC1/3. Association analysis results revealed that rare variants inducing a complete loss of function significantly increased the risk of obesity and body mass index (BMI), whereas variants in other groups with partial or neutral effects on PC1/3 activity had no impact on adiposity. We observed that *in silico* prediction tools were unreliable in detecting mutations leading to a complete loss of function.

Subsequently, I focused on rare variants of the *DYRK1B* gene. Although this gene is not directly involved in the leptin-melanocortin pathway, pathogenic variants of *DYRK1B* have been described in several patients with central obesity, type 2 diabetes (T2D), and coronary artery disease. However, the impact of rare *DYRK1B* variants has not been assessed on a large scale. In the RaDiO study, which included 9,353 participants, 65 rare variants in *DYRK1B* were detected. Following *in vitro* analysis of each variant, we identified 20 pathogenic or likely pathogenic variants (P/LP) according to the criteria of the *American College of Medical Genetics and Genomics*. Among these P/LP variants, six showed an effect leading to a complete loss of function of DYRK1B (P/LP-full). Association analyses showed that P/LP-full variants of *DYRK1B* were strongly associated with increased BMI and fasting glucose levels, as well as

a heightened risk of obesity and T2D, whereas P/LP variants had only a modest effect on adiposity and no impact on glucose homeostasis.

In conclusion, the use of functional genetics has demonstrated that only heterozygous variants of *PCSK1* and *DYRK1B* with a complete loss of function cause monogenic obesity. For *DYRK1B*, obesity is additionally associated with T2D. These results underscore the critical significance of assessing the functional impact of mutations *in vitro* for genetic diagnosis and the potential selection of appropriate treatments. We have demonstrated that *in silico* prediction tests are currently not precise enough.

## Liste des abréviations

ACMG	<i>American College of Medical Genetics and Genomics</i>
ACTH	Adrénocorticotrophine
ADCY3	Adénylate cyclase 3
ADH	Hormone antidiurétique
AGRP	<i>Agouti-related peptide</i>
AMPc	AMP cyclique
AMPK	<i>AMP activated protein kinase</i>
AOMS	<i>Abdominal obesity-metabolic syndrom</i>
ARC	Noyau arqué
ASIP	<i>Agouti signaling protein</i>
BDNF	<i>Brain-derived neurotrophic factor</i>
BBS	Syndrome de Bardet-Biedl
β-EP	β-endorphine
β-LPH	β-lipotrophine
CART	<i>Cocaine-and amphetamine-regulated transcript protein</i>
CCK	Cholécystokinine
CCND1	Cycline D1
CDC	<i>Centers of disease Control and Prevention</i>
CDKN1B	Inhibiteur de kinase dépendante de la cycline 1B
C/EBPα	<i>CCAAT/enhancer-binding protein α</i>
CELA2A	<i>Chymotrypsin Like Elastase 2A</i>
CLIP	Peptide du lobe intermédiaire de type corticotrophine
CNV	<i>Copy number variants</i>
CPE	carboxypeptidase E
CREB	<i>cAMP response element binding protein</i>
CRH	Corticotrophine
ddNTP	Didésoxyribonucléotides
DESIR	<i>Data from the Epidemiological Study on the Insulin Resistance Syndrome</i>
DH	Domaine DYRK
DT2	Diabète de type 2
DYRK1B	<i>Dual specificity tyrosine phosphorylation regulated kinase 1B</i>
EPAC	<i>Exchange factor directly activated by cAMP</i>
FASP	<i>Filter-aided sample preparation</i>
FDA	<i>Food and Drug Administration</i>
FFAR4	<i>Free fatty acid receptor 4</i>
FTO	<i>Fat mass and obesity associated gene</i>
GABA	Acide γ-aminobutyrique
GHRH	Hormone de libération de l'hormone de croissance
GHSR	<i>Growth hormone secretagogue receptor</i>
GIP	<i>Glucose-dependant insulino-tropic polypeptide</i>
GLP1R	<i>Glucagon-like peptide-1 receptor</i>
gnomAD	<i>Genome Aggregation Database</i>
GNRH	Hormone de libération des gonadotrophines hypophysaires

GoF	Gain de fonction
GWAS	<i>Genome-wide association study</i>
G6PC	Glucose-6-phosphatase
GPCR	Récepteurs couplés aux protéines G
HDL	Lipoprotéine de haute densité
5-HT	Sérotonine
IC	Intervalle de confiance
IMC	Indice de masse corporelle
IOTF	<i>International Obesity Task Force</i>
IQR	Ecart interquartile
IR	Récepteur à l'insuline
JAK	<i>Janus kinase</i>
JP	Peptide jonction
kb	kilobase
KSR2	Kinase suppresseur de Ras2
LC-MS	Chromatographie en phase liquide-spectrométrie de masse
LD	<i>Linkage disequilibrium</i>
LEP	Leptine
LEPR	Récepteur à la leptine
LFQ	Quantification « sans label »
LH	Aire hypothalamique latérale
LOD	<i>Logarithm of the odds</i>
LoF	Perte de fonction
MAC	<i>Minor allele count</i>
MAF	<i>Minor allele frequency</i>
MCH	<i>Melanin-concentrating hormone</i>
MCR	Récepteur à la mélanocortine
MiST	<i>Mixed effects score test</i>
MIRK	<i>Minibrain-related kinase</i>
MRAP2	<i>Melanocortin receptor accessory protein 2</i>
MSH	Mélanocortine
NAFLD/NASH	Stéatose hépatique non alcoolique
N-AT	N-acétyltransférase
NCOA1	<i>Nuclear receptor coactivator 1</i>
NFAT	<i>Nuclear factor of activated T cells</i>
NGS	Séquençage de nouvelle génération
NRP	Neuropiline
NPY	Neuropeptide Y
NT	Non-transfecté
NTRK2	<i>Neurotrophic receptor tyrosine kinase 2</i>
OMS	Organisation mondiale de la Santé
ONPG	<i>Ortho-nitrophenyl-<math>\beta</math>-galactoside</i>
OR	<i>Odds ratio</i>
PAM	<i>Peptidyl-glycine <math>\alpha</math>-amidating monooxygenase</i>
PC	Prohormone convertase
PCSK1	Proprotéine convertase subtilisin/kexin de type 1
$\pi$	Effet moyen
PI3K	<i>Phosphatidylinositol 3-kinase</i>

P/LP	Pathogène/probablement pathogène
P/LP-full	Pathogène/probablement pathogène avec un effet totalement inhibiteur
PLXNA	Plexine A
POMC	Proopiomélanocortine
PPAR $\gamma$	<i>Peroxisome proliferator-activated receptor <math>\gamma</math></i>
PPY	Polypeptide pancréatique
PRS	Score de risque polygénique
PVN	Noyau paraventriculaire
PWS	Syndrome de Prader-Willi
PYY	Peptide tyrosine-tyrosine
RaDiO	<i>French Rare Variants Involved in Diabetes and Obesity</i>
RE	Réticulum endoplasmique
REVEL	<i>Rare exome variant ensemble learner</i>
RT-qPCR	<i>Quantitative reverse transcription PCR</i>
SD	Déviation standard
SE	Erreur standard
SEMA3	Sémaphorines de classe 3
SH2B1	<i>Src homology 2 B adaptor protein 1</i>
SHH	<i>Sonic hedgehog</i>
SIM1	<i>Single-minded homolog 1</i>
SNC	Système nerveux central
SNP	<i>Single nucleotide polymorphism</i>
SRC	<i>Steroid coreceptor activator</i>
SVF	Sérum Foetal de veau
TCF/LEF	<i>T cell factor/lymphoid enhancer factor</i>
$\tau$	Hétérogénéité
TG	Triglycéride
TRH	Thyrotrophine
TUB	<i>Tubby-like protein</i>
VV	Vecteur vide
VUS	De significativité incertaine
WES	<i>Whole exome sequencing</i>
WGS	<i>Whole genome sequencing</i>
WT	<i>Wild-type</i>



# Table des matières

<b>INTRODUCTION</b> .....	<b>15</b>
<b>1. Etat de l'art sur l'obésité</b> .....	<b>16</b>
1.1. Définition et diagnostic .....	16
1.2. Epidémiologie .....	17
1.3. Complications.....	20
1.4. Physiopathologie de la prise alimentaire.....	21
1.4.1. La régulation périphérique de la prise alimentaire .....	22
1.4.2. La régulation hypothalamique de la prise alimentaire.....	24
<b>2. La génétique de l'obésité</b> .....	<b>34</b>
2.1. Les outils génétiques .....	35
2.1.1. Les études de liaisons .....	35
2.1.2. Les approches gènes candidats.....	36
2.1.3. Les études d'association pangénomiques (GWAS).....	36
2.1.4. Le séquençage .....	38
2.2. Les formes communes d'obésité (formes polygéniques) .....	40
2.2.1. Les SNP associés à l'obésité commune.....	40
2.2.2. Les variants rares associés à l'obésité commune.....	43
2.2.3. Les autres évènements (épi)génétiques pouvant influencer l'obésité commune.....	45
2.3. Les formes rares d'obésité (formes monogéniques).....	46
2.3.1. Les obésités monogéniques liées aux gènes impliqués dans la voie leptine-mélanocortine .....	46
2.3.2. Les obésités monogéniques liées à des syndromes .....	54
2.3.3. Limites entre les formes d'obésités monogéniques dites « syndromiques » et « non syndromiques ».....	59
2.3.4. Continuum entre les formes polygéniques et monogéniques de l'obésité .....	60
<b>3. Traitements et médecine de précision</b> .....	<b>62</b>
3.1. Changement du mode de vie .....	62
3.2. Chirurgie bariatrique .....	62
3.3. Médecine de précision.....	63
3.3.1. Importance du diagnostic génétique de l'obésité monogénique.....	63
3.3.2. Leptine recombinante .....	67
3.3.3. Setmélanotide .....	67
3.3.4. Autres traitements disponibles .....	70
<b>MON TRAVAIL DE THESE</b> .....	<b>73</b>
<b>PRESENTATION DES TRAVAUX</b> .....	<b>77</b>
<b>Partie 1: Contribution des variants hétérozygotes de <i>PCSK1</i> à l'obésité et implications pour la médecine de précision</b> .....	<b>78</b>
1.1. Contexte de l'étude.....	78

1.1.1.	La biologie de <i>PCSK1</i> .....	78
1.1.2.	La génétique de <i>PCSK1</i> .....	80
1.2	Matériel et Méthode .....	84
1.2.1.	Participants de l'étude .....	84
1.2.2.	Séquençage de <i>PCSK1</i> .....	85
1.2.3.	Analyses fonctionnelles.....	86
1.2.4.	Analyses statistiques pour les études d'association.....	91
1.3	Résultats .....	92
1.3.1.	Détection de 65 variants rares codant de <i>PCSK1</i> dans RaDiO.....	92
1.3.2.	Etudes fonctionnelles <i>in vitro</i> des variants rares hétérozygotes du gène <i>PCSK1</i> .....	94
1.3.3.	Effet des variants rares de <i>PCSK1</i> sur différents traits métaboliques selon leur pathogénicité .....	100
1.3.4.	Evaluation de l'utilité des analyses <i>in silico</i> pour discriminer les variants complètement perte de fonction des autres variants de <i>PCSK1</i> .....	102
1.4	Discussion .....	105
<b>Partie 2: Les variants pathogènes avec perte de fonction totale de <i>DYRK1B</i> causent une obésité monogénique, associée à un diabète de type 2.....</b>		<b>109</b>
2.1.	Contexte de l'étude.....	109
2.1.1.	La biologie de <i>DYRK1B</i> .....	109
2.1.2.	La génétique de <i>DYRK1B</i> .....	110
2.2.	Matériel et Méthode .....	111
2.2.1.	Participants de l'étude .....	111
2.2.2.	Séquençage de <i>DYRK1B</i> .....	112
2.2.3.	Analyses fonctionnelles.....	112
2.2.4.	Analyses statistiques pour les études d'association.....	116
2.3.	Résultats .....	116
2.3.1.	Détection de 65 variants rares codant de <i>DYRK1B</i> dans RaDiO.....	116
2.3.2.	Etudes fonctionnelles <i>in vitro</i> des variants rares du gène <i>DYRK1B</i> .....	118
2.3.3.	Effet des variants de <i>DYRK1B</i> sur différents traits métaboliques selon leur pathogénicité .....	123
2.4.	Discussion .....	128
<b>DISCUSSION GENERALE ET PERSPECTIVES.....</b>		<b>131</b>
<b>BIBLIOGRAPHIE .....</b>		<b>141</b>
<b>ANNEXES.....</b>		<b>151</b>
<b>ANNEXE 1. Tableaux supplémentaires.....</b>		<b>153</b>
<b>ANNEXE 2. Figures supplémentaires .....</b>		<b>167</b>
<b>ANNEXE 3. Article “Contribution of heterozygous <i>PCSK1</i> variants to obesity and implications for precision medicine: a case-control study” .....</b>		<b>185</b>
<b>ANNEXE 4. Article “Pathogenic, total loss-of-function <i>DYRK1B</i> variants cause monogenic obesity associated with type 2 diabetes” .....</b>		<b>187</b>

# INTRODUCTION

# 1. Etat de l'art sur l'obésité

## 1.1. Définition et diagnostic

L'obésité est reconnue comme un problème majeur de santé publique depuis 1997 par l'Organisation Mondiale de la Santé (OMS)<sup>1</sup>. Elle correspond à un excès et une accumulation de masse grasse provenant d'un déséquilibre entre les apports et les dépenses énergétiques, ce qui entraîne des conséquences néfastes pour la santé.

Le diagnostic de l'obésité est principalement basé sur l'indice de masse corporelle (IMC). L'IMC est un indice simple qui correspond au poids (kg) divisé par la taille au carré (m<sup>2</sup>). Il a comme avantage d'être indépendant du sexe et de l'âge. Le surpoids chez l'adulte est défini par un IMC supérieur à 25 kg/m<sup>2</sup> et l'obésité par un IMC supérieur à 30 kg/m<sup>2</sup> par l'OMS<sup>2</sup>. De plus, selon la valeur de l'IMC, il existe plusieurs grades d'obésité :

**Tableau 1 : Les différents types d'obésité définis par l'OMS.**

Grade	Type d'obésité	IMC (kg/m <sup>2</sup> )
I	Modérée	30-34,9
II	Sévère	35-39,9
III	Sévère/Morbide	IMC ≥ 40

Chez les enfants, il est impossible de se référer à l'IMC seul car les seuils varient selon l'âge et le sexe jusqu'à l'âge adulte. C'est pourquoi l'obésité pédiatrique est définie en utilisant un IMC qui est ajusté à l'âge et au sexe suite aux changements physiologiques de l'IMC pendant la croissance. Pour cela, des outils de référence comme des courbes de croissance et des tableaux de percentiles sont utilisés. Le diagnostic de l'obésité infantile est généralement basé sur le percentile de l'IMC pour l'âge de l'enfant qui compare son IMC à une population de référence au même âge et de même sexe. Aux Etats-Unis, les références utilisées sont celles établies par les *U.S. Centers of Disease Control and Prevention* (CDC) : le surpoids est défini par un IMC ajusté à l'âge se trouvant entre le 85ème et 94ème percentile et l'obésité par un IMC ajusté à l'âge égal ou supérieur au 95ème percentile<sup>3</sup>. En Europe, les normes les plus utilisées ont été établies par l'*International Obesity Task Force* (IOTF). L'IOTF propose une classification alternative du surpoids et de l'obésité chez l'enfant selon l'âge et le sexe sur base des courbes de l'IMC établies à partir de données recueillies dans 6 grandes études de croissance transversale dans des larges populations de référence<sup>4</sup>. Ces données ont permis la constitution

de courbes de centiles spécifiques à chaque pays atteignant les valeurs de 25 kg/m<sup>2</sup> (pour le surpoids) et 30 kg/m<sup>2</sup> (pour l'obésité) à l'âge de 18 ans. La moyenne de ces courbes a ensuite été calculée en fonction de l'âge pour chaque pays afin d'obtenir des courbes spécifiques au sexe pour chaque âge<sup>5</sup>. Selon les normes de l'IOTF, le surpoids est défini par un IMC supérieur au centile IOTF25 et l'obésité par un IMC supérieur au centile IOTF30<sup>1</sup>.

Cependant, bien que l'IMC soit l'indice le plus utilisé pour évaluer le surpoids ou l'obésité au sein d'une population, il reste un guide approximatif car il ne tient pas compte de la répartition de la graisse corporelle. En effet, un même degré d'embonpoint chez différents individus peut varier selon l'âge, le sexe et l'ethnie de l'individu. De plus, une accumulation excessive de graisse abdominale, appelée obésité abdominale, est associée à un risque de complications métaboliques et cardiovasculaires plus élevé<sup>6,7</sup>. C'est pourquoi il existe d'autres méthodes qui peuvent venir en complément de l'IMC, notamment la mesure du tour de taille, l'épaisseur des plis cutanés et l'évaluation de la composition corporelle par l'impédance bioélectrique. Ces mesures peu coûteuses permettent d'estimer l'adiposité corporelle et la répartition des graisses rapidement. Il existe aussi des méthodes de critères qui mesurent la densité du corps ou qui décrivent la quantité et la distribution des différents tissus telles que la densitométrie, la tomodynamométrie, l'absorptiométrie biphotonique à rayons X (DEXA) et d'autres techniques d'imagerie<sup>8</sup>. Ces méthodes donnent des mesures plus précises de la masse grasse ; cependant, ces méthodes sont difficiles à mettre en place et sont coûteuses, elles sont plutôt utilisées lors d'études à petite échelle comme des études cliniques<sup>7</sup>.

## 1.2. Epidémiologie

Aujourd'hui, l'obésité est devenue un problème de santé mondial impactant plus de 1 milliard de personnes à travers le monde, comprenant 680 millions d'adultes, 340 millions d'adolescents et 39 millions d'enfants. Cette prévalence est encore en constante augmentation et selon l'OMS, d'ici 2025, environ 167 millions d'individus supplémentaires (adultes et enfants) seront touchés par le surpoids ou l'obésité<sup>9</sup>. Cette augmentation peut être attribuée à une alimentation hautement riche en calories et à une vie sédentaire qui diminue l'activité physique quotidienne<sup>10</sup>.

De plus, l'obésité et le surpoids sont reconnus comme la 5ème cause de mortalité par l'OMS. Ils tuent aujourd'hui davantage que la dénutrition<sup>2</sup>. En 2019, l'obésité était la maladie métabolique responsable du nombre de décès le plus élevé avec 5 millions de décès (40,4% et

41,8% pour les hommes et les femmes, respectivement). De plus, ce taux de mortalité associé à l'obésité est resté constant depuis les années 2000<sup>11</sup>.

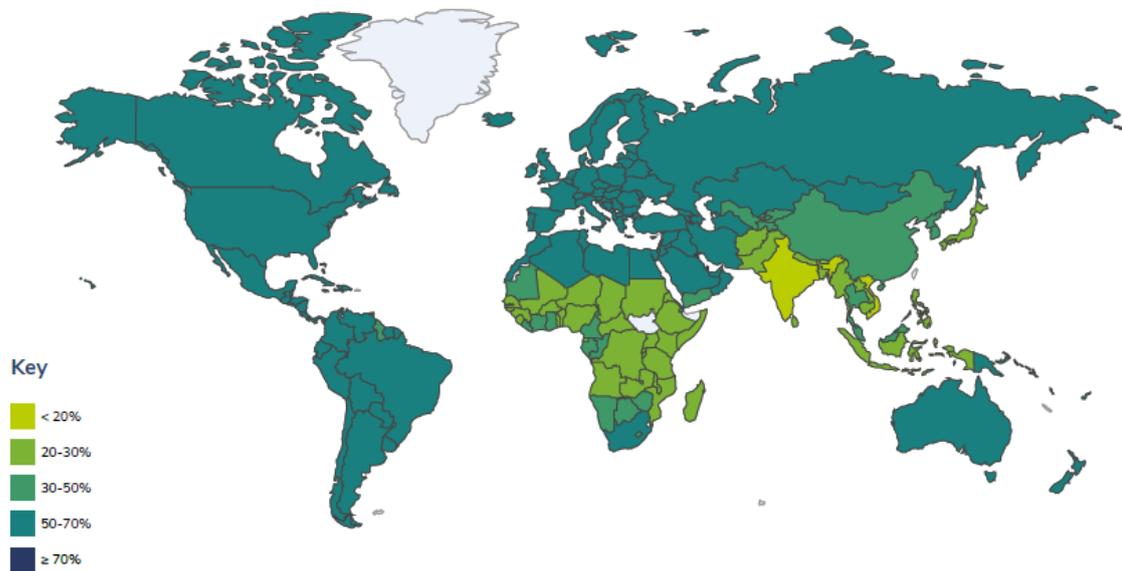
La prévalence liée à l'obésité et au surpoids varie en fonction de la distribution géographique, de l'origine ethnique, du sexe et du statut socio-économique. En général, les femmes sont plus touchées par l'obésité que les hommes<sup>12</sup>. Les régions présentant la prévalence la plus élevée d'obésité chez les adultes comprennent les îles du Pacifique, notamment Nauru (61%), les îles Cook (55,9%) et les Palaos (55,3%) ainsi que les Etats-Unis (36,2%) et certains pays du Moyen-Orient tels que le Koweït (37,9%), la Jordanie (35,5%), et l'Arabie Saoudite (35,4%) (**Figure 1**)<sup>13,14</sup>. Dans les pays à revenu élevé et dans certains pays à revenu intermédiaire notamment les Etats-Unis, la Suède, le Danemark, le Norvège, la France, l'Australie et le Japon, la prévalence de l'obésité et du surpoids augmente lentement et elle a atteint un plateau élevé depuis le début des années 2000. Cependant, elle augmente rapidement dans les pays à revenu faible ou intermédiaire, et plus particulièrement dans les zones urbaines<sup>15</sup>. De plus, une différence entre les taux de mortalité est observée : ce sont les pays à revenus moyens ou moyens supérieurs qui ont le taux de mortalité lié à l'obésité le plus important. A partir des années 2000 jusque 2019, une large diminution du taux de mortalité lié à l'obésité a été observée dans les régions européennes et américaines alors qu'elle était en augmentation dans les régions de l'Asie du Sud-Est, de la Méditerranée orientale, d'Afrique et du Pacifique occidental<sup>11</sup>. Les inégalités étant un facteur majeur de l'obésité, les personnes ayant une position socioéconomique basse présentent un risque plus élevé de développer une obésité<sup>12</sup>. Cela est d'autant plus problématique que l'obésité est associée à long terme avec des conséquences économiques négatives, influencées par le salaire moyen du pays<sup>11</sup>.

En France, en 2020, près de la moitié des adultes (47,3%) vivaient avec un surpoids et/ou une obésité, avec 17% des adultes atteints d'obésité seule (dont 2% avec une obésité de grade III). Parmi les enfants âgés de 8 à 17 ans, 21% étaient en surpoids et/ou en obésité dont 6% en obésité. Quant aux enfants âgés de 2 à 7 ans, 34% des enfants étaient en surpoids et/ou en obésité dont 18% avec une obésité. La prévalence d'obésité était plus forte dans les régions du nord avec le plus haut taux de 22,1% atteint dans la région Hauts-de-France<sup>16</sup>.

Cette augmentation de la prévalence du surpoids et de l'obésité est associée à une série d'effets négatifs sur la santé à court et moyen terme, à une baisse de la qualité de vie et à une augmentation des services de santé menant à des coûts économiques importants<sup>17</sup>.

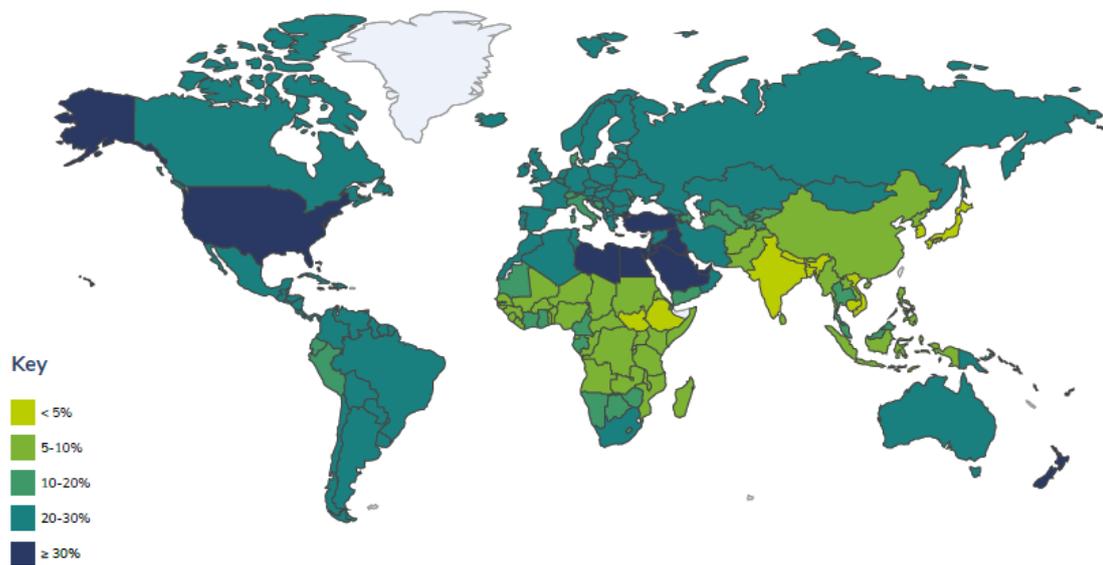
### Estimates of prevalence of overweight in adults

Overweight BMI  $\geq 25$  kg/m<sup>2</sup>



### Estimates of prevalence of obesity in adults

Obesity BMI  $\geq 30$  kg/m<sup>2</sup>



**Figure 1** : Estimation de la prévalence du surpoids (a) et de l'obésité (b) ajustée à l'âge chez les adultes dans le monde d'après la *World Obesity Federation* selon le référentiel de données d'observation mondial de l'OMS, 2020.

### 1.3. Complications

L'obésité augmente le risque de développer de nombreuses comorbidités qui peuvent avoir un impact important sur la santé : une maladie cardiovasculaire comme une maladie des artères coronaires et l'hypertension artérielle, un diabète de type 2 (DT2) lié au développement d'une résistance à l'insuline, des problèmes respiratoires tels que l'asthme et l'apnée du sommeil, certains types de cancers (cancer du sein, de l'endomètre, du côlon et du rectum, de l'utérus, de l'œsophage, du pancréas, de la prostate et du rein entre autres), des problèmes gastro-intestinaux incluant la stéatose hépatique non alcoolique (NAFLD/NASH), ou des calculs biliaires et la maladie de la vésicule biliaire, et des maladies articulaires comme l'ostéoarthrite. De plus, l'obésité peut amener à des troubles mentaux comme l'anxiété, la dépression et une basse estime de soi liée à la stigmatisation des personnes souffrant d'obésité par la société. L'obésité et ses comorbidités diminuent directement la longévité et la qualité de vie de l'individu atteint. Globalement, la présence et la sévérité de l'obésité augmentent le risque de mortalité<sup>18</sup>. Le traitement de l'obésité et de toutes ses complications ajoute une charge supplémentaire au système de santé : le coût médical des personnes atteintes d'obésité est 30 % plus élevé que celui des personnes avec un IMC normal. Le traitement des conséquences de l'obésité représente donc un défi coûteux à la fois pour l'individu et la société<sup>15</sup>.

De plus, la présence d'une obésité dès l'enfance ou l'adolescence a un impact important pour le reste de la vie, augmentant le risque de mortalité et plus particulièrement de la morbidité liée aux problèmes cardio-métaboliques<sup>17</sup>. Une revue systématique réalisée en 2019 a estimé la prévalence des comorbidités dans une population d'enfants et d'adolescents âgés de 5 à 18 ans. Les résultats de l'étude ont révélé que, par rapport à des enfants ou adolescents avec un poids normal, les enfants et adolescents avec une obésité étaient 1,4× plus susceptibles de développer du prédiabète, 4,4× plus susceptibles de souffrir d'hypertension artérielle, 26,1× plus susceptibles de développer une NAFLD et 1,7× plus susceptibles d'avoir de l'asthme<sup>19</sup>. Ces augmentations alarmantes ont une incidence directe sur les services de santé pédiatriques et adultes, mettant en évidence l'importance de la prévention, de la prise en charge précoce et du suivi de l'obésité afin de réduire le risque de complications et d'améliorer l'état de santé global<sup>17</sup>.

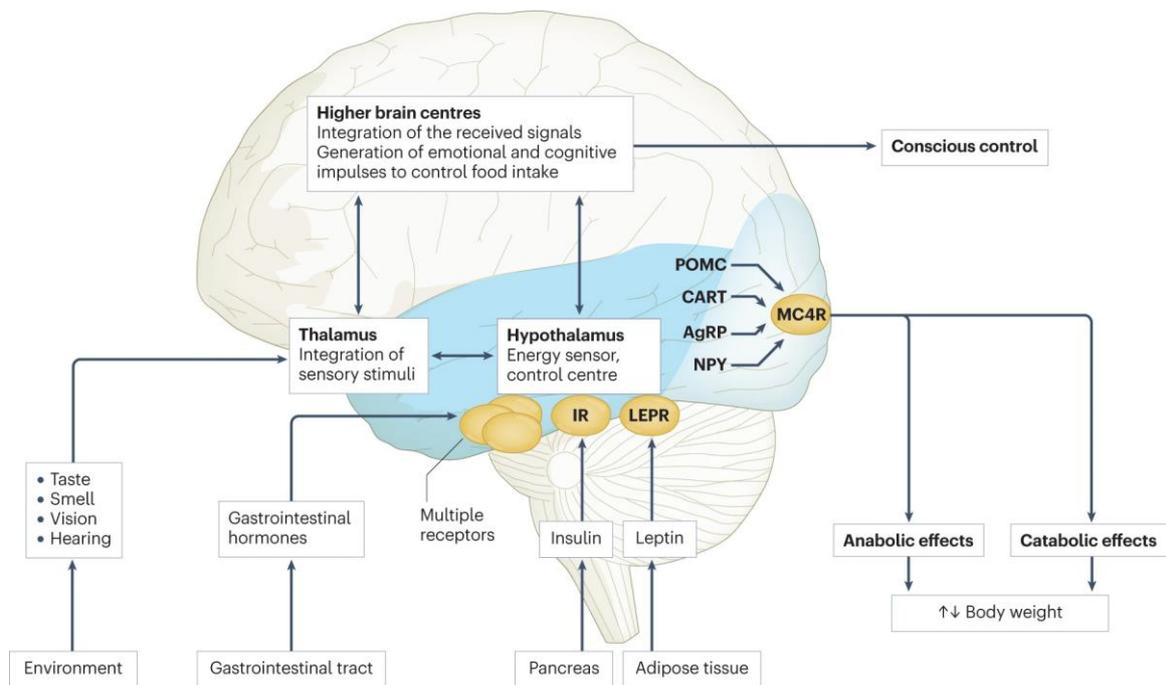
L'obésité, et plus particulièrement l'obésité abdominale, est considérée comme le facteur de risque principal du syndrome métabolique. Le syndrome métabolique correspond à un ensemble de facteurs de risques liés à la santé métabolique<sup>6</sup>. Il est caractérisé par la présence

simultanée d'au moins trois de ces cinq facteurs : une obésité abdominale (tour de taille > 102 cm pour les hommes et > 88 cm pour les femmes), une hypertriglycéridémie ( $\geq 150$  mg/dl), une cholestérolémie anormalement basse des lipoprotéines de haute densité (HDL) ( $< 40$  mg/dl pour les hommes et  $< 50$  mg/dl pour les femmes), une hypertension artérielle ( $\geq 130$  mm Hg pour la tension systolique /  $\geq 85$  mm Hg pour la tension diastolique) et une hyperglycémie à jeun ( $\geq 5,6$  mmol/l)<sup>20</sup>. Les personnes atteintes de syndrome métabolique ont un risque plus élevé de développer des maladies cardiovasculaires. L'obésité abdominale est un facteur de risque majeur pour l'inflammation systémique, l'hyperlipidémie, la résistance à l'insuline et les maladies cardiovasculaires<sup>6</sup>.

La présence d'une obésité abdominale est également liée à la résistance à l'insuline. Elle provoque le développement de cellules adipeuses plus larges et dysfonctionnelles. Ces dernières sécrètent des molécules pro-inflammatoires incluant des cytokines, des prostaglandines et de la protéine C-réactive qui peuvent interférer avec la signalisation normale de l'insuline et diminuer la sensibilité des cellules à l'action de l'insuline. Les cellules  $\beta$  pancréatiques produisent plus d'insuline en réponse à cette résistance, ce qui peut entraîner un épuisement de ces cellules  $\beta$  pancréatiques et donc un état prédiabétique voire un DT2. L'hyperinsulinémie provoquée par la résistance à l'insuline favorise le stockage excessif de graisse dans les cellules adipeuses, favorisant une prise de poids et une accumulation de graisse corporelle particulièrement dans la région abdominale. L'obésité et la résistance à l'insuline peuvent donc s'influencer mutuellement dans un cercle vicieux<sup>6</sup>.

#### 1.4. Physiopathologie de la prise alimentaire

L'obésité résulte d'un déséquilibre de la balance énergétique. Lorsque les apports énergétiques donnés par la prise alimentaire sont supérieurs aux dépenses énergétiques, l'excès de calories est stocké dans le tissu adipeux blanc sous forme de triglycérides. L'apport alimentaire est influencé par les besoins physiologiques mais également par de multiples facteurs psychologiques ou socio-culturels<sup>21</sup>. Les informations périphériques sur les réserves énergétiques et sur l'état actuel des nutriments sont transmises par les hormones au système nerveux central (SNC) afin d'induire la faim ou la satiété. Les centres cérébraux, plus particulièrement ceux situés au niveau de l'hypothalamus, intègrent ces signaux et régulent l'homéostasie énergétique par l'intermédiaire des neuropeptides. Ces signaux peuvent également être influencés par des stimuli provenant de l'environnement et des réactions émotionnelles (**Figure 2**)<sup>17</sup>.



**Figure 2 : Régulation périphérique et centrale de la prise alimentaire.** CART : *cocaine and amphetamine responsive transcript* ; IR : récepteur à l'insuline. Source : Lister, N. B. et al., « *Child and adolescent obesity* », *Nature Reviews Disease Primers*, 2023.

#### 1.4.1. La régulation périphérique de la prise alimentaire

Le stockage énergétique du tissu adipeux est régulé par des signaux hormonaux à long terme tels que la leptine et l'insuline alors que l'état nutritionnel est régulé par des signaux à court-terme provenant du tractus gastrointestinal<sup>17</sup>.

##### 1.4.1.1. La leptine

La leptine est une hormone peptidique principalement sécrétée par les adipocytes du tissu adipeux blanc mais elle est également produite en quantité moindre par d'autres organes comme le tissu adipeux brun, l'estomac, le placenta, la glande mammaire, les follicules ovariens et certains organes dans leur état fœtal comme le cœur et les os<sup>22</sup>. Sa concentration dans le sang est proportionnelle à la quantité de graisse dans le tissu adipeux et à la taille cellulaire des adipocytes<sup>23</sup>. Elle est sécrétée de façon pulsatile, selon le rythme circadien<sup>24</sup>. Le rôle principal de la leptine est de transmettre l'information sur les réserves énergétiques du corps humain au cerveau afin de contrôler l'homéostasie énergétique. Lorsque les niveaux de graisse augmentent, les cellules adipeuses sécrètent de la leptine qui est ensuite transportée par le sang jusqu'au cerveau. Une fois arrivée au niveau du cerveau, et plus particulièrement au niveau de l'hypothalamus, la leptine se lie à son récepteur spécifique (LEPR). Cette fixation provoque une inhibition de la prise alimentaire et une stimulation de la dépense énergétique par le SNC<sup>25</sup>.

LEPR est aussi exprimé au niveau des organes périphériques comme le foie, les muscles squelettiques, les cellules  $\beta$  pancréatiques et les cellules adipeuses. Ainsi, la leptine a un rôle important au niveau de la régulation énergétique globale<sup>26</sup>.

La plupart du temps, une altération de la signalisation de la leptine se manifeste par un développement d'une obésité. En effet, il a été reporté que les patients atteints d'obésité ont des concentrations de leptine significativement plus élevées dans le sang que les patients ayant un poids normal, mais sans induire une diminution de la prise alimentaire. L'hypothèse la plus probable pour expliquer cette observation est que les individus avec une obésité deviennent résistants à la signalisation de la leptine endogène, ce qui entraîne une perte de sensibilité à la leptine et une incapacité à réguler l'appétit et le métabolisme<sup>27</sup>.

#### 1.4.1.2. L'insuline

L'insuline est une hormone sécrétée par les cellules  $\beta$  du pancréas. Elle a un rôle essentiel dans la régulation de la glycémie et du métabolisme des glucides : elle stimule l'absorption du glucose et son stockage sous forme de glycogène dans le foie et les muscles. De façon similaire à la leptine, la concentration d'insuline dans le sang est proportionnelle au degré d'adiposité. Bien qu'elle ait une action moins efficace que la leptine, l'insuline joue également un rôle dans la régulation de la prise alimentaire. En effet, l'insuline se fixe sur son récepteur spécifique (IR) au niveau de l'hypothalamus, permettant de réduire la prise alimentaire et de réguler le poids du corps<sup>25</sup>.

#### 1.4.1.3. Les hormones intestinales

Le tractus intestinal joue également un rôle important dans la régulation de l'appétit *via* la sécrétion d'une grande variété d'hormones intestinales. Ces dernières sont sécrétées dans le sang par les cellules endocrines de l'intestin après un repas. Leur concentration dans le sang varie selon la prise alimentaire et la composition du repas<sup>27</sup>. Parmi ces hormones intestinales, il existe une seule hormone intestinale orexigénique, stimulant l'appétit : la ghréline. Sa concentration dans le sang est plus élevée une à deux heures avant un repas et diminue directement après la prise alimentaire. La ghréline est associée à une augmentation de la prise alimentaire, une réduction de la dépense énergétique restante et un catabolisme dans le tissu adipeux<sup>27</sup>. Toutes les autres hormones intestinales sont anorexigéniques, c'est-à-dire qu'elles diminuent l'appétit<sup>28</sup> (**Tableau 2**).

**Tableau 2 : Les différentes hormones intestinales et leur effet sur la prise alimentaire.** *Source: adapté de Perry, B. & Wang, Y., « Appetite regulation and weight control: the role of gut hormones », Nutrition & Diabetes, 2012.*

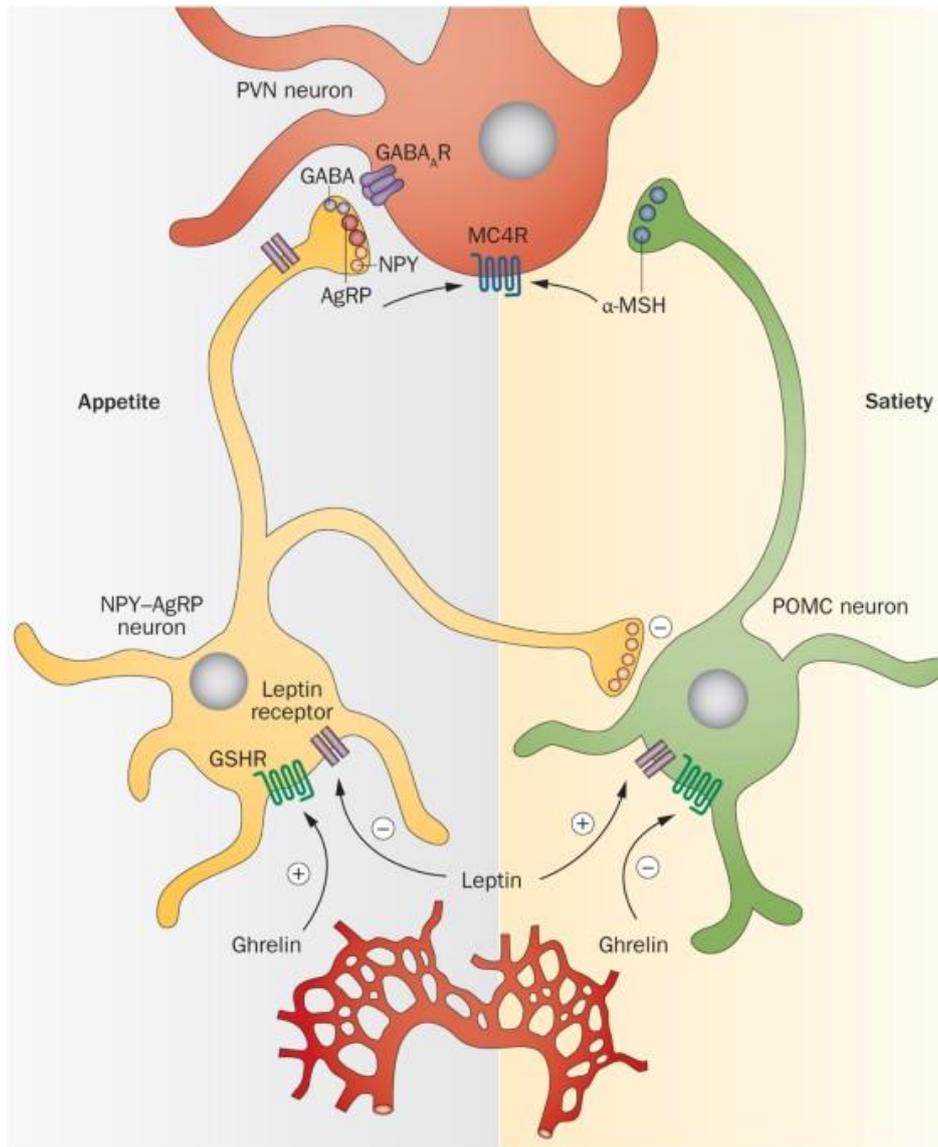
Hormones	Organe source	Récepteur	Effet
<b>Cholecystokinine (CCK)</b>	Cellules L de l'intestin	CCKAR	Anorexigénique
<b>Ghréline</b>	Cellules endocrines de l'estomac	GHSR	Orexigénique
<b>Glucagon-like peptide (GLP1)</b>	Cellules L de l'intestin	GLP1R	Anorexigénique
<b>Oxyntomoduline</b>	Cellules L de l'intestin	GLP1R ?	Anorexigénique
<b>Peptide tyrosine-tyrosine (PYY)</b>	Cellules L de l'intestin	NPY2R	Anorexigénique
<b>Polypeptide pancréatique (PPY)</b>	Pancréas et colon	NPY4R	Anorexigénique

#### 1.4.2. La régulation hypothalamique de la prise alimentaire

L'hypothalamus a un rôle central dans la régulation de la prise alimentaire, notamment *via* la signalisation de la voie leptine-mélanocortine. Les signaux provenant du système nerveux périphériques sont captés au niveau du noyau arqué (ARC) situé à la base de l'hypothalamus. La population de neurones de l'ARC est accessible aux hormones intestinales car ce noyau n'est pas complètement isolé par la barrière hémato-encéphalique, permettant l'entrée des peptides et protéines périphériques. L'ARC renferme, entre autres, deux populations de neurones responsables de la régulation de la prise alimentaire : les neurones anorexigéniques exprimant la proopiomélanocortine (POMC) et le peptide *cocaine and amphetamine-related transcript* (CART) et les neurones orexigéniques co-exprimant le neuropeptide Y (NPY) et l'*agouti-related protein* (AGRP)<sup>28</sup>. Les neurones NPY/AGRP innervent les neurones POMC et inhibent la production de POMC en libérant notamment le facteur acide  $\gamma$ -aminobutyrique (GABA)<sup>29</sup>. L'action de la leptine est double : elle stimule les neurones POMC et inhibe les neurones NPY/AGRP de l'ARC, permettant une diminution de la prise alimentaire. A l'inverse, la ghréline va se lier au *growth hormone secretagogue receptor* ou récepteur à la ghréline (GHSR) situé sur les neurones NPY/AGRP afin de stimuler la prise alimentaire (**Figure 3**)<sup>22</sup>.

Les neurones POMC et NPY/AGRP projettent leurs axones vers les noyaux de second ordre qui sont le noyau paraventriculaire (PVN) et l'aire hypothalamique latérale (LH) de l'hypothalamus. Les neurones situés dans le PVN contiennent les récepteurs à la mélanocortine (MCR) qui régulent la prise alimentaire : ils sont stimulés par les peptides provenant de POMC et inhibés par AGRP<sup>30</sup>. Alors que le PVN est le lieu de convergence des signaux provenant de

l'ARC dans l'hypothalamus, la LH abrite des neurones sensibles au glucose stimulant l'hyperphagie en cas d'hypoglycémie. La LH est responsable de la sécrétion du facteur orexigénique *melanin-concentrating hormone* (MCH) qui est impliqué dans le développement de l'obésité, de la mémoire et des émotions et des orexines qui ont un effet sur l'augmentation de la prise alimentaire<sup>22</sup>.



**Figure 3 : Régulation hypothalamique de la prise alimentaire.**  $\alpha$ -MSH: hormone mélanocortine  $\alpha$  ; AgRP : *agouti-related protein* ; GABA : acide  $\gamma$ -aminobutyrique ; GABAAR : récepteur GABA ; GHSR : *growth hormone secretagogue receptor* ; MC4R : récepteur à la mélanocortine 4 ; NPY : neuropeptide Y ; POMC : proopiomélanocortine ; PVN : noyau paraventriculaire. Source: Nasrallah, C. M. & Horvath, T. L., « Mitochondrial dynamics in the central regulation of metabolism », *Nature Reviews Endocrinology*, 2014.

#### 1.4.2.1. Les neuropeptides sécrétés par l'hypothalamus

L'hypothalamus sécrète à la fois des facteurs orexigéniques et anorexigéniques.

**NPY** est le peptide orexigène le plus abondant dans l'hypothalamus. Il est synthétisé au niveau de l'ARC et il est libéré dans le PVN. Lorsqu'il est administré au niveau du SNC à des rongeurs, il stimule une prise alimentaire puissante à court terme<sup>22</sup>.

La protéine **AGRP** est une hormone orexigène antagoniste des récepteurs MC3R et MC4R. Elle est exprimée uniquement dans l'ARC par les neurones AGRP/NPY. Contrairement à NPY, l'administration centrale d'AGRP chez les rongeurs déclenche une prise alimentaire moins intense mais prolongée, pouvant durer jusqu'à une semaine<sup>22</sup>. Chez l'homme, l'abondance d'AGRP et de NPY est corrélée à un IMC élevé<sup>31</sup>.

**CART** est un peptide anorexique puissant. CART se co-localise avec LEPR dans les neurones hypothalamiques. L'administration de CART chez le rat bloque l'initiation de la prise alimentaire induite par NPY. De plus, il active le PVN en libérant les facteurs de libération de la corticotrophine (CRH) et de la thyrotrophine (TRH) qui réduisent la prise alimentaire qu'elle soit normale ou induite chez les rongeurs<sup>22</sup>.

**POMC** est le précurseur anorexigène de nombreux neuropeptides qui ont des rôles importants dans la régulation du métabolisme : l'adrénocorticotrophine (ACTH), les mélanocortines ( $\alpha$ -MSH,  $\beta$ -MSH et  $\gamma$ -MSH), et les  $\beta$ -endorphines. POMC est clivé par les prohormones convertases (PC1/3 codée par *PCSK1* et PC2) au niveau de séquences d'acides aminés dibasiques bien définies. Les peptides spécifiques produits sont déterminés par le type de prohormone convertase exprimé spécifiquement dans les tissus. Le processus de maturation de POMC se déroule en plusieurs étapes décrites ci-dessous (**Figure 4**)<sup>32</sup>.

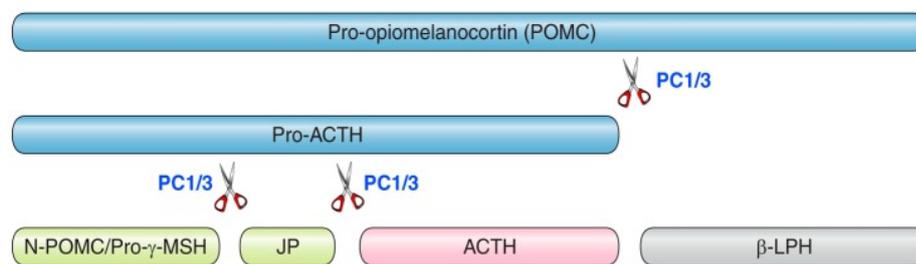
Dans l'adénohypophyse, POMC est initialement clivé par PC1/3, ce qui libère la pro-ACTH et la  $\beta$ -lipotrophine ( $\beta$ -LPH). Le second clivage de la pro-ACTH par PC1/3 libère l'ACTH, le peptide jonction (JP) et N-POMC aussi appelé la pro- $\gamma$ -MSH qui contient un groupement terminal NH<sub>2</sub>. La maturation de POMC se poursuit ensuite dans l'hypothalamus et la *pars intermedia* de l'adénohypophyse<sup>32</sup> (une zone présente chez les rongeurs et à l'état fœtal chez l'homme, mais assez rudimentaire chez l'adulte) :

- L'ACTH est clivé par PC2 ce qui libère l'ACTH (1-17) et le peptide du lobe intermédiaire de type corticotrophine (CLIP), qui correspond à l'ACTH (18-39). Les groupements terminaux d'acides aminés COOH de l'ACTH (1-17) sont éliminés par la

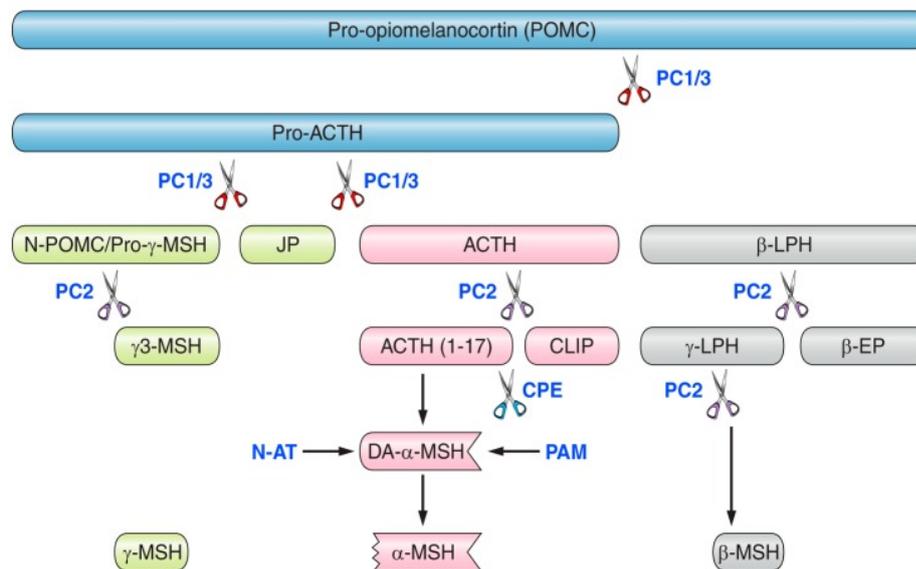
carboxypeptidase E (CPE). L'ACTH (1-13) est ensuite amidée au niveau de l'extrémité COOH par la peptidyl-glycine  $\alpha$ -amidating monoxygénase (PAM) ce qui forme l'acétyl- $\alpha$ -MSH (DA- $\alpha$ -MSH). La DA- $\alpha$ -MSH est finalement acétylée au niveau de son domaine terminal NH<sub>2</sub> par la N-acétyltransférase (N-AT) libérant le peptide  $\alpha$ -MSH.

- La  $\beta$ -LPH est clivée en  $\gamma$ -LPH et en  $\beta$ -endorphine ( $\beta$ -EP) par PC2. La  $\gamma$ -LPH est ensuite clivée au niveau de son site lysine-lysine par PC2 et libère la  $\beta$ -MSH.
- La pro- $\gamma$ -MSH est clivée par PC2 pour donner la  $\gamma$ -3-MSH. Elle est suivie par deux clivages supplémentaires pour libérer la  $\gamma$ -MSH<sup>32</sup>.

### A Processing in the anterior lobe of the pituitary in humans



### B Processing in the hypothalamus, skin, pars intermedia of pituitary



### Figure 4 : Processus de maturation de la pro-opiomélanocortine (POMC).

PC1/3 : prohormone convertase 1/3 ; ACTH : adrénocorticotrophine ; JP : peptide de jonction ;  $\beta$ -LPH :  $\beta$ -lipotrophine ; MSH : mélanocortine ; PC2 : prohormone convertase 2 ; CLIP : peptide du lobe intermédiaire de type corticotrophine ; CPE : carboxypeptidase E ; DA- $\alpha$ -MSH : acétyl- $\alpha$ -MSH ; N-AT : N-acétyltransférase ; PAM : peptidyl-glycine  $\alpha$ -amidating monoxygénase ;  $\beta$ -EP :  $\beta$ -endorphine. Source: Harno, E., Gali Ramamoorthy, T., Coll, A. P. & White, A., « POMC: The Physiological Power of Hormone Processing », *Physiological Reviews*, 2018.

**Les mélanocortines** libérées par POMC se fixent sur les récepteurs à la mélanocortine (MC1R à MC5R), qui sont des récepteurs couplés aux protéines G (GPCR). Ces récepteurs sont impliqués dans différents processus physiologiques. Ils sont différenciés par leur fonction, leur localisation tissulaire et leur affinité pour leur ligand.

- Le rôle principal de MC1R est la régulation de la pigmentation de la peau et des follicules pileux où il est exprimé majoritairement. Son ligand principal est l' $\alpha$ -MSH qui est produit de façon endogène par les kératinocytes de la peau et des cheveux après une exposition aux ultra-violets. Il peut aussi être activé par l'ACTH et dans une moindre mesure par la  $\beta$ -MSH et la  $\gamma$ -MSH.
- MC2R est exprimé au niveau du cortex surrénal où il est connu comme le récepteur à ACTH. Il lie uniquement l'ACTH qui active la cascade de synthèse des glucocorticoïdes par l'axe hypothalamo-hypophysaire.
- MC3R a un rôle moins important dans l'homéostasie énergétique, il agit principalement comme un « auto-récepteur » sur les neurones POMC de l'ARC. Il lie les ligands  $\alpha$ -MSH,  $\beta$ -MSH,  $\gamma$ -MSH et ACTH de façon égale. Il est exprimé au niveau de l'hypothalamus, des régions limbiques et dans des tissus périphériques comme l'estomac, le pancréas, le duodénum, le cœur, les testicules, les ovaires, le muscle squelettique et le rein où son rôle n'est pas encore bien établi.
- MC4R est fortement exprimé dans le SNC mais plus particulièrement au niveau du PVN de l'hypothalamus. Le principal agoniste de MC4R est l' $\alpha$ -MSH qui est libéré des neurones POMC dans le PVN afin d'inhiber la prise alimentaire. MC4R peut également lier la  $\beta$ -MSH et l'ACTH, ainsi que la  $\gamma$ -MSH dans une moindre mesure. Une molécule antagoniste de MC4R est le ligand AGRP qui est libéré dans le PVN par les neurones AGRP/NPY. L'AGRP, par sa fixation sur MC4R, inhibe la signalisation de MC4R, ce qui provoque une augmentation de la prise alimentaire<sup>33</sup>. La fixation de l'AGRP et de l' $\alpha$ -MSH fonctionne comme un système dynamique afin de réguler finement la prise alimentaire *via* la signalisation de MC4R<sup>22</sup>.
- MC5R est hautement exprimé pendant l'embryogenèse et il est impliqué dans la fonction des glandes exocrines. Sa différence principale avec les autres MCR est qu'il est exprimé dans plusieurs tissus périphériques mais pas dans le SNC. Son ligand principal est l' $\alpha$ -MSH mais il peut également lier l'ACTH, la  $\beta$ -MSH et la  $\gamma$ -MSH<sup>32</sup>.

Par ailleurs, les opioïdes endogènes incluant la  **$\beta$ -endorphine**, la dynorphine, l'enképhaline stimulent la prise alimentaire par l'intermédiaire de leurs récepteurs  $\mu$ -opioïde,  $\kappa$ -opioïde et  $\delta$ -opioïde<sup>22</sup>.

Il existe d'autres neuropeptides qui peuvent exercer un rôle dans la prise alimentaire : les **endocannabinoïdes** stimulent la prise alimentaire par leur fixation sur les récepteurs CB1 et CB2<sup>22,34</sup> et les **neurotransmetteurs monoaminergiques** tels que la noradrénaline, la dopamine et la sérotonine (5-HT) qui modulent la prise alimentaire. La fixation de la sérotonine sur son récepteur *5-hydroxytryptamine receptor 1B* (5HT1B) inhibe la prise alimentaire<sup>22</sup>. La dopamine régule la faim par l'intermédiaire des récepteurs D1 et D2 en fonction de la région hypothalamique spécifique. Quant à la noradrénaline, elle a un effet anorexigène quand elle se fixe sur ses récepteurs adrénergiques  $\alpha 1$  et  $\beta$ -2 et un effet orexigène par son récepteur  $\alpha 2$ -adrénergique<sup>35</sup>.

#### 1.4.2.2. La voie leptine-mélanocortine

La régulation de l'appétit par l'hypothalamus est possible grâce à la voie leptine-mélanocortine. En résumé, cette voie est stimulée par la fixation spécifique de la leptine sur son récepteur, ce qui active les neurones POMC et inhibe les neurones NPY/AGRP au niveau de l'ARC. Cette stimulation déclenche la production de la prohormone POMC qui ensuite est clivée en  $\alpha$ -MSH par PC1/3 et PC2. Le ligand  $\alpha$ -MSH active la signalisation de MC4R, provoquant la libération d'AMP cyclique (AMPc) qui, par l'activation d'autres voies métaboliques, entraîne un sentiment de satiété et diminue la prise alimentaire<sup>36</sup>.

##### a) Activation de LEPR

Après un repas, la leptine est sécrétée par le tissu adipeux dans la circulation sanguine. Elle se fixe sur le récepteur LEPR qui est exprimé à la surface des neurones POMC et AGRP au niveau de l'ARC de l'hypothalamus. LEPR fait partie de la famille des récepteurs à cytokine. Il est exprimé sous forme de dimère à la surface cellulaire. Bien qu'il existe sous différentes formes d'isoformes, c'est le long isoforme LEPRb (NM\_002303) qui est exprimé dans plusieurs régions du SNC et qui transmet le signal de la leptine. En effet, LEPR contient un long domaine intracellulaire qui est essentiel à l'initiation de la cascade de signalisation activée par la leptine<sup>31</sup>.

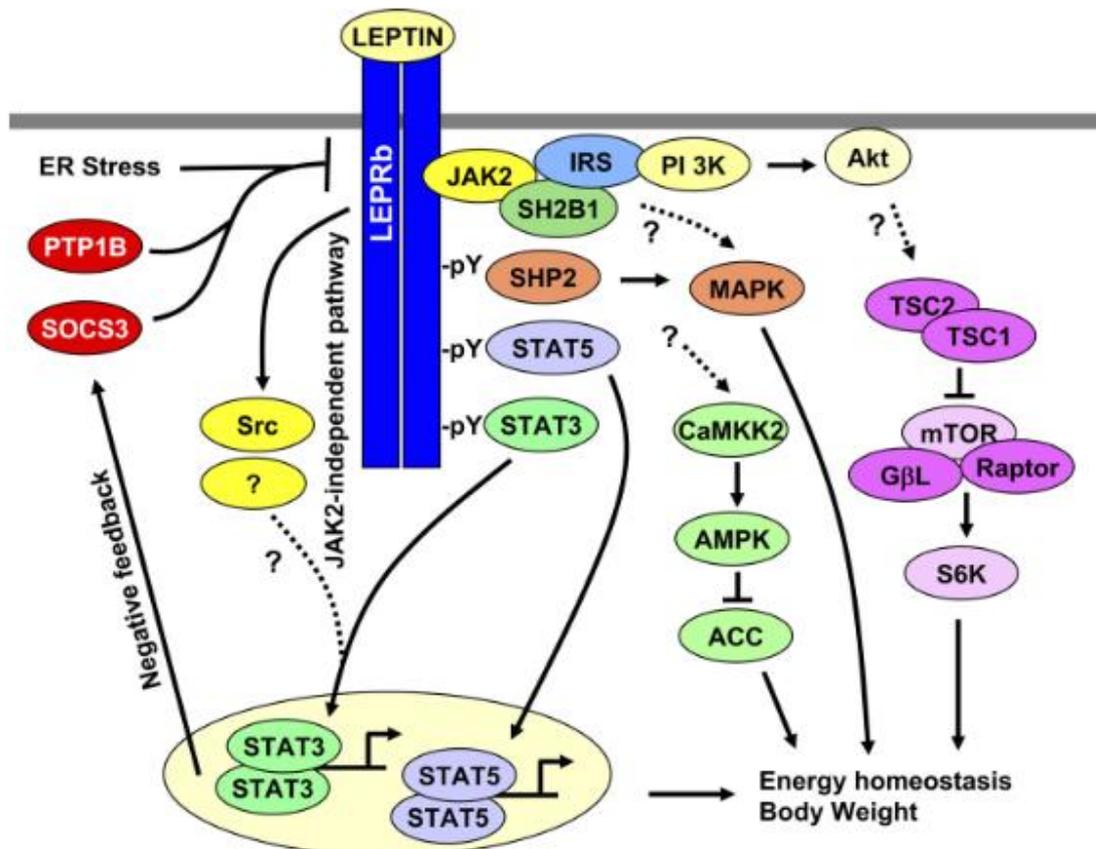
La fixation de la leptine sur LEPRb déclenche le recrutement de plusieurs voies de signalisation comprenant la voie *janus kinase-signal transducer and activator of transcription* (JAK2/STAT3/STAT5)<sup>25</sup>, la voie *insulin receptor substrate/phosphoinositide 3-kinase*

(IRS/PI3K), la voie *SH2-containing protein tyrosine phosphatase 2/MAPK* (SHP2/MAPK), la voie *AMP activated protein kinase/acétyl-CoA-carboxylase* (AMPK/ACC), la voie *mammalian target of rapamycin/S6 kinase* (mTOR/S6K) et la voie indépendante de JAK2 (**Figure 5**)<sup>24,37</sup>.

LEPR est lié constitutivement à JAK2, une tyrosine kinase de la famille des JAK. La fixation de la leptine stimule la dimérisation de LEPR qui active JAK2. JAK2 phosphoryle LEPR au niveau de plusieurs résidus tyrosines (p.Y985, p.Y1077 et p.Y1138 au niveau du transcrit NM\_002303) de son domaine intracellulaire.

- Les facteurs de transcription STAT3 et STAT5 présents dans le cytoplasme cellulaire se lient aux résidus tyrosines phosphorylés de LEPR (p.Y1138 et p.Y1077, respectivement), ce qui stimule la phosphorylation de JAK2 et active STAT3 et STAT5. Ces derniers transloquent du cytoplasme vers le noyau où ils se lient aux promoteurs du gène *POMC* (actifs dans les neurones POMC) et aux promoteurs du gène *AGRP* (actifs dans les neurones AGRP/NPY), ce qui stimule l'expression de *POMC* et inhibe celle d'*AGRP*. Cette action est principalement réalisée par STAT3 et dans une moindre mesure par STAT5<sup>24,37</sup>.
- En parallèle, JAK2 phosphorylé interagit avec sa protéine accessoire *Src homology 2 B adaptor protein 1* (SH2B1) qui se lie aux récepteurs *insulin receptor substrate* (IRS), ce qui provoque la phosphorylation de ces derniers. IRS phosphorylé active la voie de la *phosphatidylinositol 3-kinase* (PI3K) et de l'*AKT serine/threonine kinase 1* (AKT)<sup>24</sup>. AKT inhibe l'activité et l'expression du facteur de transcription orexigène *forkhead boîte 01* (FOXO1) qui stimule l'expression d'*AGRP* et de NPY et qui inhibe l'expression de *POMC*. L'inactivation de FOXO1 entraîne la liaison de STAT3 aux promoteurs de *POMC* et d'*AGRP*. De plus, l'activation de la voie PI3K est essentielle dans l'induction de la dépolarisation des neurones POMC, qui est nécessaire à leur activation. Parallèlement à l'action de la leptine, la fixation de l'insuline sur les neurones POMC active également IRS, ce qui entraîne l'activation de PI3K et la dépolarisation des neurones POMC<sup>31,37</sup>.
- Des analyses *in vitro* ont montré que la leptine peut stimuler STAT3 et la voie des MAPK également de façon indépendante de JAK2. Cette stimulation serait possible grâce aux membres de la famille des tyrosines kinases SRC. Ainsi, les voies dépendantes et indépendantes de JAK2 agissent de façon coordonnée dans l'activation de STAT3<sup>37,38</sup>.

- La fixation de la leptine sur LEPR régule également d'autres voies de signalisation telles que la voie mTOR/S6K par l'intermédiaire de la voie PI3K/AKT, la voie AMPK/ACC et la voie des MAPK de façon dépendante ou non de la protéine phosphatase SHP2<sup>24,37</sup>.
- Enfin, la voie de signalisation de la leptine LEPRb est inhibée par l'induction du *suppressor of cytokine signaling 3* (SOCS3) et de la protéine tyrosine phosphatase 1B (PTP1B), grâce à un processus de rétroaction négative<sup>24,37</sup>.



**Figure 5 : La voie de signalisation de la leptine.** LEPRb : récepteur à la leptine ; JAK2 : *janus kinase 2* ; STAT : *signal transducer and activator of transcription* ; SHP2 : *SH2-containing protein tyrosine phosphatase 2* ; SOCS3 : *Suppressor of cytokine signaling 3* ; PTP1B : *protein tyrosine phosphatase 1B* ; SH2B1 : *Src homology 2 B adaptor protein 1* ; IRS : *insulin receptor substrate* ; PI3K : *phosphoinositide 3-kinase* ; CaMKK2 : *Ca<sup>2+</sup>/calmodulin-dependent protein kinase kinase* ; AMPK : *AMP-activated protein kinase* ; ACC : *acetyl-CoA carboxylase* ; mTOR : *mammalian target of rapamycin* ; TSC : *tuberous sclerosis complex* ; S6K : *S6 kinase*. Source: Morris, D. L. & Rui, L., « *Recent advances in understanding leptin signaling and leptin resistance* », *American Journal of Physiology Endocrinology and Metabolism*, 2009.

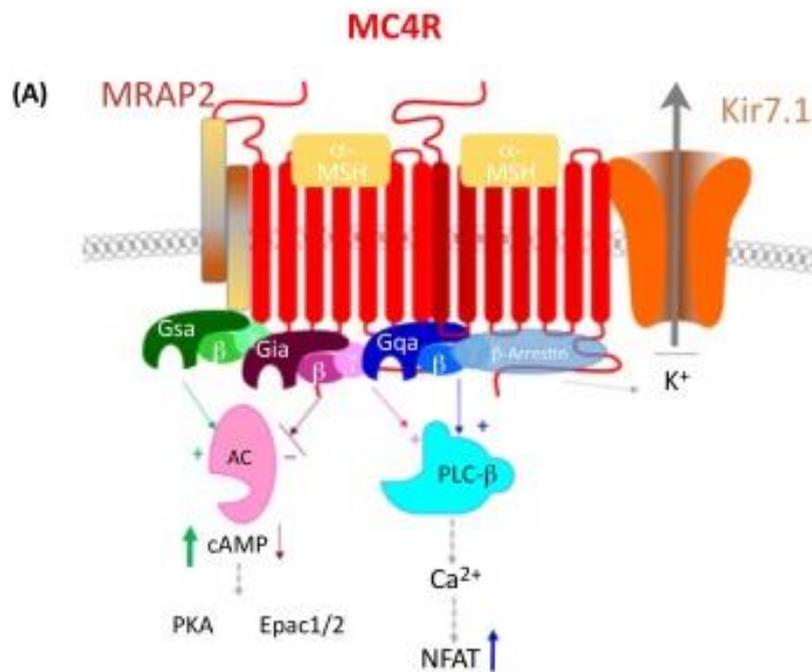
### b) Production de POMC

Suite à l'activation de la voie de signalisation par la leptine à la surface des neurones POMC, la prohormone précurseur POMC est produite puis maturée par des modifications post-

traductionnelles réalisées par les enzymes PC1/3 et PC2 afin de produire plusieurs neuropeptides comme décrits au point 4.2.1, notamment les mélanocortines  $\alpha$ -MSH et  $\beta$ -MSH<sup>39</sup>.

c) *Activation du récepteur MC4R*

Les neurones POMC de l'ARC projettent leurs axones sur les neurones MC4R au niveau du PVN<sup>39</sup>. Les MSH libérés vont se fixer sur MC4R, activant la voie de signalisation ce qui réduit la prise alimentaire (Figure 6)<sup>33</sup>.



**Figure 6 : La voie de signalisation du récepteur MC4R.** AC : adénylate cyclase ; Kir7.1 : *inwardly rectifying potassium channels*; NFAT : *nuclear factor of activated T cells*; Epac1/2 : *exchange factor directly activated by cAMP*; MC4R : récepteur 4 à la mélanocortine; PLC-β : phospholipase C-β; Gs: protéine G stimulatrice; Gi: protéine G inhibitrice; MRAP2 : *melanocortin receptor-associated protein 2*; PKA : protéine kinase A. Source : Kühnen, P., Krude, H. & Biebermann, H., « *Melanocortin-4 Receptor Signalling: Importance for Weight Regulation and Obesity Treatment* », *Trends in Molecular Medicine*, 2019.

Le récepteur MC4R forme des homodimères et interagit avec la protéine accessoire MRAP2 qui est orientée en antiparallèle. En tant que GPCR, MC4R peut se coupler aux trois classes principales des protéines G (G<sub>S</sub>, G<sub>i/o</sub> et G<sub>q</sub>)<sup>40</sup>.

La fixation des hormones MSH sur MC4R active la sous-unité de la protéine G stimulatrice (G<sub>S</sub>) qui recrute à son tour l'adénylate cyclase qui convertit l'ATP en AMPc<sup>31</sup>. L'AMPc active la protéine kinase A (PKA) ou la signalisation *exchange factor directly activated by cAMP* (EPAC). La PKA activée affecte de nombreux processus physiologiques en activant des

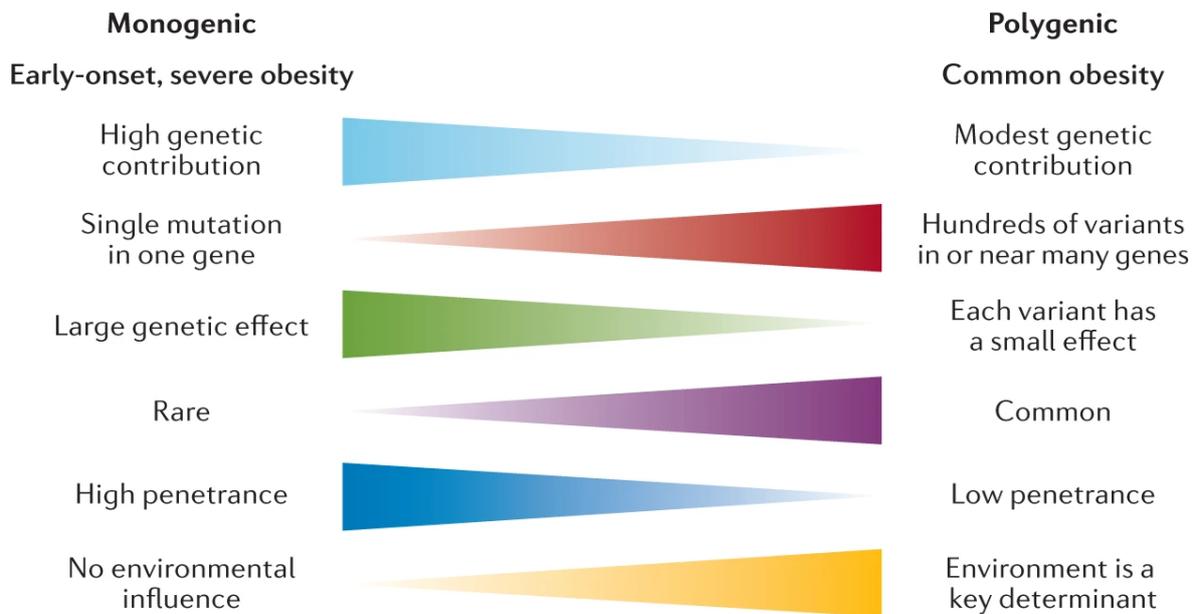
protéines effectrices ultérieures, notamment des kinases, des canaux ioniques et des protéines de signalisation. La PKA active également les *cAMP response element binding protein* (CREB), ce qui phosphoryle la PKA et augmente l'expression des gènes ciblés<sup>41</sup>. De plus, il a été également reporté que l'activation de Gq/11 par les MSH stimule la phospholipase C (PLC) qui successivement provoque une libération de calcium et active la signalisation de *nuclear factor of activated T cells* (NFAT). En outre, la fixation de  $\alpha$ -MSH sur MC4R déclenche la fermeture du canal potassique KIR7.1, ce qui dépolarise la cellule<sup>33</sup>. A l'inverse, les molécules antagonistes de MC4R, comme l'AGRP, activent la sous-unité  $G_{i/o}$  qui inhibe l'adénylate cyclase, diminuant les niveaux d'AMPc et désactivant la voie de MC4R<sup>41</sup>.

#### 1.4.2.3. La voie BDNF/NTRK2

La signalisation induite par le *brain-derived neurotrophic factor* (BDNF) et son récepteur *neurotrophic receptor tyrosine kinase 2* (NTRK2) interagit avec la signalisation MC4R dans l'hypothalamus. BDNF qui est un facteur de croissance du système nerveux, est largement exprimé dans le SNC où il joue un rôle important dans le développement neuronal et la plasticité synaptique<sup>39</sup>. Des études ont montré que l'activation de MC4R stimule l'expression de BDNF dans le noyau ventromédial de l'hypothalamus, suggérant que cette voie agit en aval de MC4R, diminuant la prise alimentaire<sup>42</sup>. De plus, il a été mis en évidence que l'injection de BDNF dans différentes régions cérébrales atténue l'hyperphagie et la prise de poids chez des souris déficientes en *Mc4r*<sup>43</sup>. Cependant, le lien exact entre ces deux voies n'a pas encore été établi.

## 2. La génétique de l'obésité

L'obésité est une maladie multifactorielle dépendante à la fois de facteurs environnementaux et génétiques. L'héritabilité de l'obésité est estimée entre 40 et 75% selon les études<sup>44</sup>. Aujourd'hui, on considère qu'il existe deux grandes formes d'obésité selon l'implication génétique : monogénique et polygénique (**Figure 7**)<sup>39</sup>.



**Figure 7 : Les caractéristiques principales des formes monogéniques et polygéniques de l'obésité.** Source: Loos, R. J. F. & Yeo, G. S. H., « *The genetics of obesity: from discovery to biology* », *Nature Reviews Genetics*, 2022.

Les formes d'obésité monogénique, relativement rares, sont caractérisées par un seul évènement génétique avec un effet fort qui est suffisant pour déclencher l'apparition de la maladie. Elles sont provoquées par des variants rares appelés plus souvent mutations (avec une fréquence allélique mineure [MAF *minor allele frequency*] < 1%). Ces formes, contrairement aux formes polygéniques, surviennent plus tôt dans la vie du patient et sont généralement plus sévères. Les formes polygéniques, aussi plus couramment appelées obésité commune, sont connues comme complexes et multifactorielles. En effet, elles se caractérisent par une composante environnementale et une composante génétique. Elles sont le résultat de l'effet combinatoire de multiples *single nucleotide polymorphisms* (SNP) fréquents (MAF > 1%) présents dans plusieurs gènes qui, couplés à des facteurs environnementaux, provoquent le développement de la maladie. Ces SNP ont des effets faibles car aucun d'entre eux n'est assez pénétrant pour provoquer seul l'apparition de la maladie<sup>39,45</sup>.

## 2.1. Les outils génétiques

L'identification des régions génétiques impliquées dans les différentes formes d'obésité a été rendue possible grâce à plusieurs stratégies basées sur des outils génétiques. Au départ, cette recherche était limitée aux formes monogéniques en raison de contraintes technologiques. Cependant, le progrès et le développement d'outils plus performants ont ouvert la voie à une caractérisation plus détaillée à la fois des formes monogéniques et polygéniques, permettant une grande avancée dans la compréhension de la génétique de l'obésité<sup>44</sup>.

### 2.1.1. Les études de liaisons

Au cours des années 1990, les études de liaison familiale ont été largement utilisées pour identifier des *loci* génétiques spécifiques associés à une maladie ou un trait phénotypique partagés entre des individus porteurs apparentés. Cette méthode se base sur le principe que la distance génétique entre deux *loci* correspond à la probabilité que ces *loci* puissent être séparés naturellement par recombinaison lors de la méiose. Au plus cette distance est faible, au plus il est probable que ces deux *loci* soient situés sur le même bloc de recombinaison, et donc qu'ils soient transmis ensemble. De ce fait, l'existence d'une coségrégation entre des marqueurs génétiques et la maladie à travers les générations suggère que les marqueurs génétiques et le locus responsable de la maladie soient situés à proximité dans le génome<sup>44</sup>. La mesure utilisée pour calculer la probabilité de coségrégation entre un phénotype et les blocs de recombinaison des membres d'une même famille est le *logarithm of the odds* (LOD) Score Z. Celui-ci correspond au ratio de la liaison observée sur la probabilité qu'il s'agisse d'un événement aléatoire. Une liaison est établie généralement quand le LOD score est supérieur à 3. Dans ce cas, les limites des blocs de recombinaison de chaque individu sont alors combinées afin de déterminer quelle région présente un pic de liaison et dans laquelle il est vraisemblable de retrouver un gène associé à la maladie. A l'inverse, on va considérer qu'il n'y a pas de liaison entre deux *loci* quand le LOD score est inférieur à  $-2^{46}$ . Les marqueurs génétiques utilisés sont des marqueurs polymorphiques, notamment les microsatellites qui étaient employés dans le passé, et des puces à ADN (incluant des centaines de milliers de SNP). Ces marqueurs pouvaient être utilisés sur une région d'intérêt ou sur la totalité du génome<sup>44</sup>. Au plus le nombre de patients étudiés était important, au plus la région retenue était précise et permettait alors de s'approcher de l'identification du gène responsable de la maladie qui peut alors être séquencé<sup>39</sup>.

Les études de liaison présentaient comme avantage qu'elles ne nécessitaient pas de connaissances préalables sur la physiologie de la maladie étudiée, sauf si elles étaient réalisées

sur une zone ciblée dans le cadre d'étude gène candidat (cf infra). Elles ont été particulièrement utiles pour identifier des variants génétiques rares associés à des maladies monogéniques où une seule variation génétique peut être responsable du phénotype. Cependant, en raison des difficultés liées au recrutement de larges familles, cette technique n'est plus couramment utilisée aujourd'hui, car d'autres outils génétiques plus performants sont désormais disponibles, comme décrits ci-dessous.

### 2.1.2. Les approches gènes candidats

Les approches gènes candidats se concentrent sur des gènes spécifiques soupçonnés d'être impliqués dans la pathogenèse de la maladie. Ils sont sélectionnés en se basant sur le rôle physiologique des protéines qu'ils codent et qui pourraient contribuer à l'obésité ou sur des résultats issus d'études sur des modèles animaux. Par exemple, les gènes connus pour réguler l'homéostasie énergétique ou exprimés dans le tissu adipeux chez la souris sont considérés comme des candidats potentiels au développement de l'obésité chez l'humain. Pour démontrer l'implication d'un gène candidat dans la maladie, plusieurs approches peuvent être utilisées, notamment la recherche de liaisons génétiques dans la région chromosomique du gène d'intérêt dans des familles souffrant d'obésité ou chez des individus non apparentés présentant le même phénotype. Une autre méthode consiste à séquencer les exons de ces gènes candidats chez les individus cas (personnes atteintes d'obésité) et témoins (personnes non atteintes d'obésité) de l'étude afin de détecter de potentielles mutations responsables de la maladie<sup>44</sup>. On compare ensuite la fréquence des variants entre les deux populations afin de déterminer s'il existe une surreprésentation de mutations dans ce gène parmi les patients souffrant d'obésité.

Les approches gènes candidats ont comme avantage de se concentrer uniquement sur des régions spécifiques du génome et de pouvoir être utilisées chez des individus non apparentés, ce qui facilite le recrutement par rapport aux larges familles<sup>47</sup>. Elles ont également été très efficaces, en regard des limitations technologiques de l'époque, pour identifier des gènes dans le contexte d'obésité monogénique où l'effet génétique attendu est fort et spécifique. Le premier exemple notable est, en 1994, la découverte du gène de la leptine (*LEP*) sur base de modèle murin (cf infra). Cependant, elles ont pour inconvénient de nécessiter des connaissances préalables sur l'implication de gènes spécifiques dans l'obésité et de ne pas explorer de manière exhaustive l'ensemble du génome<sup>48</sup>.

### 2.1.3. Les études d'association pangénomiques (GWAS)

Le développement des études d'association pangénomiques ou *genome-wide association studies* (GWAS) en anglais, a été possible grâce à l'avènement des techniques de séquençage à haut débit à partir des années 2000. Les GWAS ont permis une avancée majeure dans la compréhension de la génétique de l'obésité polygénique et d'autres maladies métaboliques complexes. Ils se basent sur le génotypage de plusieurs milliers de SNP chez des dizaines de milliers d'individus via l'utilisation de puces à ADN. Ces puces abritent plusieurs milliers de petites sondes d'ADN correspondant à des SNP particuliers<sup>25</sup>.

En général, les GWAS liées à l'obésité se divisent en deux catégories : les études cas-témoins et les analyses d'association de traits quantitatifs représentatifs de l'obésité. Dans la première catégorie, les individus sont classés selon leur phénotype : les "cas" sont atteints d'obésité ( $IMC \geq 30 \text{ kg/m}^2$ ) et les "témoins" ont un poids normal ( $IMC < 25 \text{ kg/m}^2$ ). Les fréquences de leurs SNP sont alors comparées entre les deux groupes. La seconde catégorie vise à identifier les variants génétiques associés à des caractéristiques quantitatives de l'obésité comme l'IMC ou la taille dans une cohorte représentative de la population générale. Dans ce cas, les analyses statistiques sont effectuées afin d'évaluer les associations entre ces variants génétiques et le trait quantitatif étudié sans sélectionner les individus<sup>44</sup>.

Les GWAS doivent tester les fréquences des allèles de chaque SNP afin de déterminer son association avec le phénotype en question. Les GWAS reposent sur le principe que des SNP génotypés montrant une association avec le phénotype d'intérêt peuvent être hérités ensemble suite à un déséquilibre de liaison (*linkage disequilibrium* [LD]) avec les variants responsables du phénotype. Cela permet de génotyper environ 80% de tous les variants fréquents en une seule fois avec uniquement 500 000 SNP sélectionnés<sup>49</sup>. Des analyses statistiques sont ensuite réalisées afin d'attribuer une p-valeur à chaque SNP, indiquant sa probabilité qu'il soit associé à l'obésité ou au trait métabolique quantitatif étudié. Ces études nécessitent de larges cohortes afin de disposer d'une puissance statistique suffisante pour détecter des associations dont l'effet est généralement faible pour les variants fréquents<sup>44</sup>. C'est pourquoi des collaborations internationales à large échelle ont été formées comme le *Genetic Investigation for Anthropometric Traits* (GIANT) consortium qui combinent des résumés des statistiques de GWAS individuels pour générer des ensembles de données comprenant des centaines de milliers d'individus<sup>39</sup>. La taille des échantillons est d'autant plus augmentée par la disponibilité croissante d'importantes cohortes et biobanques, telles que la *UK BioBank*, le *Million Veterans Project*, *All of Us*, la *Biobank Japan* et *23andMe*, dépassant facilement les 4 millions de

participants dans les méta-analyses afin d'accélérer la découverte de nouveaux *loci* associés à l'obésité<sup>39</sup>.

L'avantage des GWAS est qu'elles permettent d'explorer l'ensemble du génome pour identifier de nouveaux variants fréquents associés à un phénotype d'intérêt sans avoir besoin d'hypothèse ou de preuves d'implication sur la maladie<sup>26</sup>. De cette façon, les GWAS ont mis en évidence près de 1000 *loci* associés à l'obésité et/ou à une augmentation de l'IMC, permettant la découverte de gènes candidats potentiels pour l'obésité<sup>50</sup>. En revanche, leur principale limite est leur interprétation : le résultat est une corrélation statistique entre un SNP et un trait phénotypique mais la relation de cause à effet sur la pathologie reste inconnue. Les GWAS ne permettent pas de déterminer la biologie qui lie des variants génétiques à une maladie d'intérêt<sup>51</sup>. C'est pourquoi, le principal défi aujourd'hui est de déterminer quel(s) gène(s) se cachent derrière un locus associé à l'obésité ou à l'IMC. De plus, la majorité des SNP associés à l'obésité identifiés par GWAS (> 95%) sont situés dans des parties non codantes des gènes de l'ADN, abritant des éléments de régulation qui influencent l'expression des gènes à proximité ou parfois sur des longues distances. En outre, les *loci* identifiés ont un effet modéré sur le risque d'obésité ; en effet, ils expliquent moins de 6% de la variance de l'IMC<sup>49,51</sup>. Enfin, la majorité des GWAS sur l'obésité souffrent du biais qu'elles ont été réalisées principalement dans des populations européennes. Quelques *loci* ont également été identifiés dans des populations d'Asie, d'Afrique, Hispanique ou autres mais les échantillons de taille de population sont beaucoup plus petits<sup>39</sup>.

#### 2.1.4. Le séquençage

La technique de séquençage permet le décryptage de tous les variants, dont les variants rares (MAF < 1%) ou de faibles fréquences (MAF < 5%). La plupart des gènes de l'obésité monogénique ont été découverts sur base d'études cas-contrôles et de séquençage : les patients avec une obésité sévère ainsi que les membres de leur famille affectés et non affectés ont été diagnostiqués pour des mutations causales au sein de gènes potentiels par séquençage<sup>39</sup>.

##### 2.1.4.1. Le séquençage Sanger

Le séquençage Sanger, développé dans les années 1970, est une technique de séquençage qui est capable de séquencer des courtes régions d'ADN (d'environ 400-800 paires de bases en moyenne). Il a permis le séquençage de l'être humain pour la première fois qui a pris plus de 10 ans et a coûté 3 milliards de dollars. Aujourd'hui, le séquençage Sanger est utilisé classiquement dans le cas de petit séquençage ponctuel. Il a permis dans le passé de séquencer

l'ADN génomique des pics de liaison mis en évidence dans les études de liaison familiale, permettant de déterminer le gène ou la mutation lié au développement de la maladie. Le séquençage Sanger consiste en une PCR où sont utilisés des didésoxyribonucléotides (ddNTP) fluorescents en plus des désoxyribonucléotides classiques (dNTP) ce qui génère plusieurs amplicons de différentes longueurs. Lorsque la polymérase incorpore un des ddNTP, la synthèse du fragment complémentaire s'arrête. Un mélange de plusieurs amplicons est ainsi généré. Ils sont séparés en fonction de leur taille par électrophorèse sur gel. La séquence de l'ADN complémentaire est déterminée par la lecture de la fluorescence émise par chaque ddNTP incorporé à l'extrémité des différents amplicons. En utilisant un alignement sur une séquence de référence, il est possible d'identifier les mutations présentes dans la région d'ADN d'intérêt.

#### 2.1.4.2. Le séquençage à haut débit (NGS)

Le séquençage à haut débit ou le *next generation sequencing* (NGS) permet de séquencer l'ensemble du génome de milliers d'individus. Depuis la première technologie lancée en 2005, le NGS n'a cessé de progresser, entraînant une véritable révolution dans le domaine de la génomique. Il a comme avantage d'offrir une meilleure efficacité que le séquençage Sanger avec plus grande rapidité et à faible coût. Aujourd'hui, un génome humain entier peut être séquencé en quelques jours pour moins de 1000 dollars. Il est possible de séquencer le génome entier (*whole genome sequencing*, [WGS]), l'exome uniquement (*whole exome sequencing* [WES]) ou encore un panel de gènes sélectionnés pour leur implication dans la maladie. C'est pourquoi le NGS est maintenant utilisé comme outil de choix pour identifier de nouveaux variants et chercher des gènes candidats d'obésité en milieu clinique, sans nécessiter de connaissances préexistantes sur le gène en question<sup>26</sup>. L'approche séquençage NGS la plus couramment utilisée est celle développée par Illumina, qui se déroule en 4 étapes :

- La préparation des bibliothèques d'ADN : l'ADN est fragmenté aléatoirement mécaniquement ou par l'utilisation d'enzymes de restriction. Ces fragments s'apparient ensuite sur des adaptateurs, des petites molécules de séquences connues, qui leur permettent de s'accrocher à une cellule de flux (*flow-cell*). Dans le cas du WES ou de séquençage de panel de gènes, les fragments d'ADN correspondant aux régions d'intérêt sont sélectionnés par hybridation à des sondes spécifiques.
- Les fragments d'ADN sont amplifiés par PCR sur la cellule de flux pour obtenir un signal suffisamment important au moment du séquençage.
- Les fragments amplifiés sont ensuite séquencés grâce à des amorces de séquençage, de l'ADN polymérase et des nucléotides fluorescents portant une molécule empêchant

l'extension de la synthèse du brin d'ADN. Une fois que le signal fluorescent du nucléotide complémentaire du fragment d'ADN est lu par le séquenceur, la molécule qui bloque l'allongement du brin d'ADN est retirée du nucléotide, permettant la fixation d'un nouveau nucléotide. La séquence est lue grâce à ces ajouts et lecture successifs des différents nucléotides complémentaires au fragment d'ADN.

- La séquence de tous les fragments (*reads*) est alors alignée sur la séquence d'un génome utilisé comme référence afin de reconstituer la séquence entière du génome ou des régions d'intérêt d'un individu.

Grâce à cette technique, on peut retrouver les variants de certains gènes qui sont présents chez les individus atteints mais absents du génome de référence. Des études statistiques d'association peuvent être ensuite réalisées entre les variants rares et la maladie afin de déterminer s'ils sont associés au risque de développer la maladie. Cependant, étant donné que très peu d'individus sont porteurs de ces variants, la puissance statistique de ces études est considérablement réduite, rendant difficile l'établissement d'une association significative entre ces variants et la maladie. C'est pourquoi ces études sont souvent réalisées en regroupant plusieurs variants rares en un seul cluster pour augmenter le nombre de patients inclus dans l'analyse, ce qui permet d'augmenter la puissance statistique afin de déterminer l'association entre ces variants et l'obésité<sup>39</sup>. De plus, cette analyse peut être étendue à différentes catégories de clusters classés en fonction de la pathogénicité des variants, permettant de spécifier si les variants les plus pathogènes ont un effet sur la maladie et si cet impact est plus prononcé que celui des variants peu ou non pathogènes.

## 2.2. Les formes communes d'obésité (formes polygéniques)

Les formes polygéniques d'obésité sont des types d'obésité communes qui sont influencés par de multiples variants génétiques contribuant à une prédisposition à l'obésité. Ces variants interagissent également avec des facteurs environnementaux pour contribuer à l'étiologie de l'obésité.

### 2.2.1. Les SNP associés à l'obésité commune

Les premières études sur la génétique de l'obésité commune étaient basées sur des études de liaison et des approches gènes candidats. Cependant, bien que ces deux techniques soient très utiles pour la découverte des gènes d'obésité monogénique, elles n'ont eu qu'un impact marginal sur la progression de la découverte des gènes des formes d'obésité communes. En

effet, elles présentent plusieurs contraintes : un petit nombre d'échantillons, une couverture éparse des variations génétiques du génome et un manque de réplification des résultats<sup>39</sup>. Ainsi, à la fin de l'année 2005, bien que plus de 250 *loci* liés à l'obésité ont été identifiés à partir de 61 scans d'analyses de liaison du génome, ces résultats n'ont pas pu être répliqués dans d'autres cohortes<sup>52</sup>. De même après une méta-analyse de 37 études de liaison réalisées sur le génome entier (portant sur environ 31 000 individus), aucun *loci* spécifique pour l'IMC ou l'obésité n'a pu être mis en évidence<sup>53</sup>. Cela peut s'expliquer par le fait que les variants fréquents ont un effet trop faible pour être détectés efficacement par les analyses de liaison.

Les approches gène candidat ont également été utilisées afin de valider le rôle de gènes associés à l'obésité extrême chez l'humain et la souris dans le développement de l'obésité commune. Des variants communs de ces gènes candidats ont été testés pour évaluer leur association avec le risque d'obésité, l'IMC et d'autres traits liés à l'obésité. Bien qu'une centaine de gènes candidats aient été testés, seulement 6 gènes ont démontré une association solide et répliquable entre leurs variants fréquents et l'obésité commune dans différentes populations<sup>39</sup> : *ADRB3*<sup>54</sup>, *BDNF*<sup>55</sup>, *CNRI*<sup>56</sup>, *MC4R*<sup>57</sup>, *PCSK1*<sup>58</sup> et *PPARG*<sup>59</sup>.

L'investigation des variants fréquents concernant l'obésité commune s'est accélérée avec le développement des GWAS. Les premiers GWAS sur l'obésité ont été publiés en 2007. Ils ont identifié un premier cluster de variants communs significativement associé avec l'IMC dans le premier intron du locus *fat mass and obesity associated gene (FTO)*<sup>60,61</sup>. *FTO* est le locus qui a été le plus étudié et le plus répliqué avec les différents traits d'obésité parmi plusieurs populations d'origine différente<sup>62</sup>. En outre, ce cluster de SNP situé dans le premier intron de *FTO* a montré l'effet le plus important parmi tous les *loci* identifiés, augmentant l'IMC de 0,39 kg/m<sup>2</sup> par allèle et multipliant le risque d'obésité par 1,2<sup>62</sup>. Les variants dans *FTO* sont particulièrement communs dans la population européenne (MAF = 40-45%), représentant 1% de la variance de l'IMC dans la population générale<sup>63</sup>. Cependant, malgré cette association avec l'obésité largement répliquée, l'identification exacte des variants et/ou du gène causal autour de *FTO* ainsi que le mécanisme par lequel ils influencent l'obésité n'ont pas été élucidés. En effet, certaines études proposent que le variant causal de *FTO* régulerait en fait l'expression du gène *RPGRIP1L* situé à proximité tandis que d'autres études suggèrent que le variant régulerait les gènes *IRX3* et *IRX5* qui sont situés à distance dans des types de cellules spécifiques<sup>64</sup>. Le gène *FTO* en tant que tel, a été prédit par des analyses bioinformatiques pour coder une *Fe(II)- and-2-oxoglutarate-dependent dioxygenase* avec un rôle potentiel dans la réparation et la

modification des acides nucléiques. *FTO* est par ailleurs fortement exprimée dans les parties du cerveau régulant l'équilibre énergétique et le comportement alimentaire<sup>49</sup>.

Le deuxième locus de variants fréquents le plus répliqué comme étant associé à des traits de l'obésité est situé dans *MC4R*. Ces variants ont d'abord été reportés comme associés à l'IMC dans une première méta-analyse de GIANT<sup>65</sup> et ensuite d'autres SNP dans *MC4R* ont été reportés comme associés tour de taille et à la résistance à l'insuline<sup>66</sup>. Ces associations avec l'obésité et l'IMC ont depuis été largement répliquées.

Depuis lors, de nombreuses études GWAS ont été menées, et jusqu'à présent, plus de 1000 *loci* indépendants associés avec l'IMC ont été identifiés dans de larges cohortes (environ 700 000 individus). Cependant, collectivement, tous ces *loci* n'expliquent que 6% de la variance de l'IMC suite à leur effet modeste sur le risque de la maladie<sup>50</sup>. De plus, l'identité du gène causal de la plupart de ces *loci* et sa contribution dans la physiopathologie de l'obésité est inconnue. Toutefois, il est possible d'identifier le gène candidat le plus probable d'un point de vue physiologique à l'intérieur ou à proximité du SNP associé à l'obésité.

Ainsi, des variants fréquents situés dans les gènes *FTO*, *MC4R*, *SH2B1*, *BDNF*, neurexine 3 (*NRXN3*), neuronal growth regulator 1 (*NEGR1*), protéine transmembranaire 18 (*TMEM18*), potassium channel tetramerization domain-containing 15 (*KCTD15*), glucosamine-6-phosphate deaminase-2 (*GNPDA2*), mitochondrial carrier homolog 2 (*MTCH2*), endosomal/lysosomal Niemann-Pick C1 gene (*NPC1*), ets variant 5 (*ETV5*), serologically defined colon cancer antigen 8 (*SDCCAG8*), Fas apoptotic inhibitory molecule 2 (*FAIM2*) et prolactine (*PRL*) sont associés avec l'IMC ou l'obésité par GWAS et ils sont tous exprimés ou connus pour jouer un rôle au niveau du SNC<sup>49,67,68</sup>. De plus, deux autres *loci* identifiés par GWAS sont localisés en aval des gènes des molécules d'adhésion cellulaire *CADMI* et *CADM2* qui codent des protéines d'adhésion cellulaire de la famille des immunoglobulines et qui régulent l'assemblage synaptique du SNC<sup>39</sup>.

D'autres variants fréquents associés à l'obésité ou à ces traits phénotypiques sont situés dans des gènes qui peuvent avoir un effet périphérique sur la répartition des graisses et l'obésité centrale : transcription factor activating enhancer-binding protein 2  $\beta$  (*TFAP2B*), natural cytotoxicity triggering receptor 3 (*NCR3*), phosphotriesterase related gene (*PTER*), *MAF* (codant le facteur de transcription c-MAF)<sup>68</sup> et lysophospholipase-like-1 (*LYPLAL1*), qui n'est associé à l'obésité que chez les femmes<sup>69</sup>.

Ces résultats montrent que dans l'ensemble, l'expression des gènes situés à proximité de *loci* associé à l'obésité est augmentée dans des régions qui peuvent être divisées en deux grandes catégories : les régions du cerveau avec un rôle clé dans l'addiction et la récompense (particulièrement la substance noire et le cortex insulaire<sup>70</sup>) et les régions en périphérie agissant sur le stockage de graisse et l'obésité centrale par l'intermédiaire du tissu adipeux<sup>49</sup>. Néanmoins, de nombreux *loci* sont en attente d'études afin de trouver de nouveaux indices sur des gènes sous-tendant des nouvelles voies biologiques<sup>39</sup>.

### 2.2.2. Les variants rares associés à l'obésité commune

Les variants rares (MAF < 1%) et de faible fréquence (MAF < 5%) ont un effet plus fort sur l'étiologie des maladies que les variants fréquents, ils sont des candidats potentiels pour expliquer une part de l'héritabilité manquante non déterminée par les variants fréquents<sup>71-73</sup>.

En 2012, c'est par une approche gène candidat qu'un variant de faible fréquence dans le gène *GPR120* a été mis en évidence pour augmenter le risque d'obésité commune. Le gène *GPR120*, aussi appelé *free fatty acid receptor 4 (FFAR4)*, code le récepteur des acides gras insaturés à longue chaîne, un GPCR qui a un rôle essentiel dans la régulation du métabolisme des lipides, la régulation de l'appétit et des préférences alimentaires. Un dysfonctionnement de ce récepteur lipidique peut provoquer une mauvaise régulation du métabolisme des lipides qui peut contribuer à l'accumulation de graisse corporelle qui peut entraîner une obésité chez la souris et l'humain. En effet, dans cette étude, il a été montré que les souris déficientes en *Gpr120* nourries avec un régime hyperlipidique développent une obésité et une intolérance au glucose plus importante que les souris normales. Chez l'homme, il a été montré qu'une mutation perte de fonction (p.R270H) dans *GPR120* (NM\_181745) chez 3% de patients avec une obésité était associée à une altération de la régulation des lipides, et augmentait le risque d'obésité commune de 1,62<sup>72</sup>.

Néanmoins, aujourd'hui, ce sont principalement les études de séquençage à grande échelle qui permettent d'identifier les variants rares ou de faible fréquence associés à l'obésité commune ou à l'IMC. Par exemple, une méta-analyse combinant les données de génotypage d'exome de plus de 700 000 individus a permis d'identifier 14 variants rares et de faible fréquence (comprenant 1 variant non-sens et 13 variants non synonymes prédits comme pathogènes par les algorithmes de prédiction *in silico*) associés à l'IMC dans 13 gènes différents. Parmi ceux-ci, on compte 2 gènes déjà identifiés dans l'obésité monogénique (*MC4R* et *KSR2*), 8 nouveaux gènes impliqués dans l'obésité humaine (*ZBTB7B*, *ACHE*, *RAPGEF3*, *RAB21*, *ZFHX3*,

*ENTPD6*, *ZFR2*, *ZNF169*), et 3 gènes déjà identifiés par GWAS (*PRKAG1*, *HIP1R*, *MC4R* et *GIPR*). Cette étude a mis notamment en évidence l'association avec un IMC plus bas pour 2 variants rares non synonymes prédits comme pathogènes dans le gène *gastric inhibitory polypeptide receptor (GIPR)* qui joue un rôle dans la sécrétion d'insuline et le dépôt de graisse. Les analyses de voies de signalisation ont confirmé l'enrichissement des gènes avec un rôle clé dans les processus neuronaux mais également ceux avec un rôle dans la biologie des adipocytes et de la dépense énergétique. Ces variants rares ont un effet plus important sur l'IMC que les variants communs mais avec une pénétrance variable<sup>64</sup>.

De même, une large étude d'association réalisée sur plus de 400 000 participants de la UK Biobank s'est intéressée à l'effet des variants de faible fréquence et rares nuls, qui sont par conséquent considérés comme induisant une perte de fonction de la protéine associée d'après les règles de génétique, sur différents traits métaboliques. Ils ont identifié un variant rare non-sens dans le gène *GPR151* (qui code GPCR151) et un autre non-sens dans le gène *polycystic kidney and hepatic disease 1* aussi appelé *fibrocystine-L (PKHD1L1)*. *PKHD1L1* est un gène lié à des maladies comme la dyskinésie ciliaire primaire et la polykystose rénale. La fonction exacte de *GPR151* est inconnue mais il est exprimé dans le SNC, notamment au niveau des neurones hypothalamiques. Il serait impliqué dans le comportement addictif. Cette étude a démontré que le variant rare non-sens (p.R95\*) identifié dans le gène *GPR151* (NM\_194251) a un effet protecteur sur l'obésité et le DT2. En effet, cette mutation était associée à une réduction de l'IMC, et les porteurs avaient 12% de chance de moins de souffrir d'obésité (14% de moins pour le DT2 et 9% pour les maladies des artères coronaires)<sup>74</sup>.

Par ailleurs, une étude récente exploitant des données de WES sur plus de 600 000 personnes a identifié 9 gènes pour lesquels des variants rares nuls et non synonymes prédits comme pathogènes par les logiciels de prédiction *in silico* étaient associés significativement à l'IMC : *MC4R*, *GPR151*, *GIPR*, *UHMK1*, *SPARC*, *CALCR*, *KIAA0586*, *DPP9* et *ANKDR27*. En outre, cette même analyse a également mis en évidence l'association entre l'IMC et des variants rares uniquement nuls dans 7 autres gènes : *PCSK1*, *GPR75*, *ROBO1*, *KIAA1109*, *UBR2*, *PDE3B* et *ANO4*. Parmi tous ces gènes identifiés, 5 codent des GPCR exprimés dans le cerveau, spécifiquement au niveau de l'hypothalamus (*CALCR* codant le récepteur à la calcitonine, *MC4R*, *GIPR*, *GPR151* et *GPR75*). Les chercheurs de cette étude se sont davantage intéressés à *GPR75*. Ils ont mis en évidence que des variants nuls, tronquant la protéine *GPR75*, étaient associés à une adiposité plus basse et à un effet de protection contre l'obésité chez l'homme. De plus, ils ont validé cette observation chez un modèle de souris porteur de la délétion

génétique *Gpr75*. Ils ont montré qu'un régime riche en graisses dans ce modèle menait à une résistance au gain de poids, une plus grande sensibilité à l'insuline et une amélioration du contrôle glycémique<sup>75</sup>.

### 2.2.3. Les autres événements (épi)génétiques pouvant influencer l'obésité commune

Les *copy number variants* (CNV) sont des variations structurales de l'ADN impliquant des changements dans le nombre de copies d'un segment spécifique de l'ADN. Contrairement aux SNP, les CNV peuvent impliquer des gains ou des pertes de portions d'ADN. Ils peuvent se produire de différentes manières selon les événements génétiques : duplication, délétion, inversion, translocation, amplification<sup>49</sup>. Bien que les GWAS étudient principalement les variants fréquents, ils ont également été utilisés pour déterminer l'impact de certains CNV sur l'obésité commune. De cette façon quelques CNV fréquents ont montré une association significative avec l'obésité : une délétion de 1p31.1 proche du gène *neuronal growth regulator 1* (*NEGR1*) codant une molécule d'adhésion cellulaire exprimée dans le cerveau<sup>67</sup>, une délétion 16p12.3 en amont de *G-protein coupled receptor family C group 5 member B* (*GPRC5B*) qui a un rôle dans la régulation de la sécrétion d'insuline<sup>76</sup>, un CNV en 10q11.2 dans le gène *pancreatic polypeptide receptor 1* (*PPYR1*) qui régule l'homéostasie énergétique<sup>77</sup>, un CNV en 11q11 qui comprend les récepteurs olfactifs *OR4P4*, *OR4S2* et *OR4C6*<sup>77</sup>, et un CNV multi-allélique en 1p21.1 qui englobe le gène *amylase alpha 1 A* (*AMY1A*) qui produit l' $\alpha$ -amylase salivaire importante pour la digestion de l'amidon<sup>78</sup>.

Une meilleure compréhension du rôle des **ARN non codants**, tels que les microARN (miARN), pourrait contribuer à expliquer une part plus importante de l'héritabilité de l'obésité. Les miARN sont de petites molécules d'ARN non codantes qui régulent l'expression des gènes après la transcription en se liant à la région 3'UTR des gènes cibles pour limiter leur traduction en protéines. Les miARN jouent un rôle reconnu dans la différenciation des adipocytes, l'obésité, la résistance à l'insuline et la régulation de l'appétit<sup>79</sup>. Ainsi, des perturbations dans l'expression des miARN, qu'elles soient causées par des facteurs génétiques ou environnementaux, pourraient potentiellement influencer la susceptibilité à l'obésité<sup>49</sup>.

Les **modifications épigénétiques** sont des modifications qui affectent l'expression des gènes sans modifier la séquence de l'ADN. Parmi celles-ci, la méthylation de l'ADN et les modifications des histones, en affectant l'expression des gènes associés, sont susceptibles de jouer un rôle clé dans l'hérédité et la susceptibilité à l'obésité<sup>49,80</sup>.

### 2.3. Les formes rares d'obésité (formes monogéniques)

Les formes d'obésité monogénique sont des formes rares d'obésité (<5% de la population atteinte d'obésité dans les pays occidentaux), qui sont causées par une seule mutation génétique spécifique et identifiable avec un effet fort qui est suffisant pour déclencher l'apparition de la maladie. Ces formes apparaissent dès l'enfance et sont généralement très sévères accompagnées la plupart du temps d'hyperphagie avec parfois des troubles métaboliques, des anomalies hormonales ou d'autres syndromes variés.

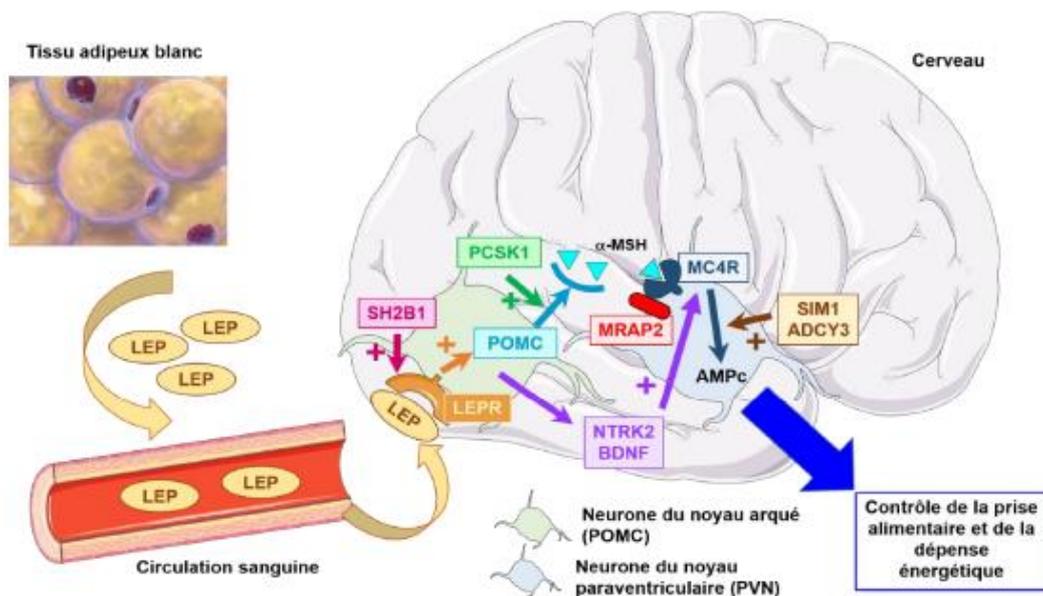
Les approches gène candidat ont été les premières techniques utilisées pour définir les formes rares d'obésité. Elles ont permis, sur base de données cliniques d'un petit nombre de sujets présentant une obésité sévère précoce de définir les formes rares d'obésité. L'identification des mutations responsables des formes d'obésité monogénique s'est accélérée avec le développement d'outils de séquençage génétique plus précis et plus efficace comme l'utilisation du WES dans des cohortes de patients atteints d'obésité sévère et précoce<sup>81</sup>. De plus, cette identification et caractérisation a été facilitée par les études des populations consanguines car ces populations ont plus de probabilité de présenter des mutations délétères homozygotes. En effet, alors que les mutations rares concernent 2 à 5% de la population d'individus atteints d'obésité en Europe, des études réalisées à partir d'une population consanguine au Pakistan ont montré que des mutations dans les gènes codant *LEP*, *LEPR* et *MC4R* expliquent plus de 30% des cas d'obésité sévère observés chez les enfants dans cette population<sup>82,83</sup>.

#### 2.3.1. Les obésités monogéniques liées aux gènes impliqués dans la voie leptine-mélanocortine

Plus de 10 gènes responsables de formes d'obésité monogénique ont été identifiés comme décisifs dans la voie leptine-mélanocortine<sup>36</sup> (**Figure 8**). Les mutations pathogènes dans un ou plusieurs gènes inclus dans cette voie peuvent altérer son bon fonctionnement biologique, dérégulant le contrôle de la prise alimentaire, menant à une obésité précoce accompagnée d'hyperphagie<sup>84</sup>. Suite au rôle primordial de l'hypothalamus dans la régulation hormonale, les phénotypes des patients présentant ce type d'obésité sont souvent associés à des anomalies endocriniennes et/ou neurologiques centrales<sup>10</sup> (**Tableau 3**).

**Tableau 3 : Les obésités monogéniques liées aux gènes de la voie leptine-mélanocortine.** *Source adaptée de : Baron, M., Froguel, P. & Bonnefond, A., « Du nouveau dans la génétique des formes monogéniques d'obésité et son impact pour mieux en comprendre la physiopathologie », médecine/sciences, 2020.*

<b>Gène</b>	<b>Transmission</b>	<b>Phénotype associé</b>
<b>Adénylate cyclase 3 (ADCY3)</b>	Autosomal récessif	Obésité sévère, hyperphagie, anosmie
<b>Agouti signaling protein (ASIP)</b>	Autosomal dominant (duplication en tandem)	Obésité sévère, hyperphagie, surcroissance, hypopigmentation (cheveux roux et peau pâle)
<b>Brain-derived neurotrophic factor (BDNF)</b>	Autosomal dominant	Obésité sévère, hyperphagie, retard de développement, troubles de la nociception
<b>Leptine (LEP)</b>	Autosomal récessif	Obésité sévère et précoce, hyperphagie, réduction du tonus nerveux sympathique, hypogonadisme, hypothyroïdie, altération de la fonction immunitaire
<b>Récepteur à la leptine (LEPR)</b>	Autosomal récessif Autosomal dominant ?	Obésité sévère et précoce, hyperphagie, hypogonadisme, hypothyroïdie, altération de la fonction immunitaire, anomalies de croissance
<b>Récepteur 4 à la mélanocortine (MC4R)</b>	Autosomal dominant (AD) et récessif (AR)	Obésité sévère (AR) ou inexistante, modérée ou sévère (AD), hyperphagie, grande taille
<b>Melanocortin 2 receptor accessory protein 2 (MRAP2)</b>	Autosomal dominant	Obésité ou surpoids, hyperglycémie, hypertension
<b>Neurotrophic tyrosine kinase receptor 2 (NTRK2)</b>	Autosomal dominant	Obésité sévère, hyperphagie, retard du développement, troubles de la nociception
<b>Proprotein convertase subtilisin/kexin type 1 (PCSK1)</b>	Autosomal récessif Autosomal dominant ?	Obésité sévère et précoce, hyperphagie, malabsorption intestinale, hypoglycémie postprandiale, déficience en glucocorticoïdes, hypogonadisme
<b>Proopiomélanocortine (POMC)</b>	Autosomal récessif	Obésité sévère et précoce, hyperphagie, hypopigmentation (cheveux roux et peau pâle), insuffisance en ACTH
<b>Src homology 2 [SH2] B adaptor protein 1 (SH2B1)</b>	Autosomal dominant	Obésité sévère, hyperphagie, petite taille
<b>SIM bHLH transcription factor 1 (SIM1)</b>	Autosomal dominant	Obésité sévère et précoce, hyperphagie, syndrome de Prader-Willi, retard de développement



**Figure 8 : Les gènes impliqués dans la voie leptine-mélanocortine dont les mutations pathogènes entraînent des formes d'obésité monogénique.** Source adaptée de : Baron, M., Froguel, P & Bonnefond, A., « Du nouveau dans la génétique des formes monogéniques d'obésité et son impact pour mieux en comprendre la physiopathologie », médecine/sciences, 2020.

### 2.3.1.1. Les gènes *LEP* et *LEPR*

Le gène *LEP* code la leptine dont la fixation sur son récepteur (codé par le gène *LEPR*) est le point de départ de la voie leptine-mélanocortine, qui entraîne une diminution de la prise alimentaire *via* un sentiment de satiété. Le rôle de la leptine a été mis en évidence en 1994, grâce aux études sur le modèle murin de la déficience congénitale en leptine (modèle *ob/ob*). Ce modèle présente une obésité sévère, une hyperphagie et une diminution de la dépense énergétique<sup>85</sup>. De façon similaire, le modèle de la déficience en *LEPR* (le modèle murin *db/db*) ne répond pas à la leptine endogène ou exogène, et présente un phénotype comparable à celui des souris *ob/ob*<sup>86,87</sup>.

C'est à la fin des années 1990, que des mutations homozygotes situées dans le gène de la leptine ou dans le gène de son récepteur *LEPR* ont été mises en évidence chez des patients atteints d'obésité sévère principalement dans des populations consanguines<sup>88,89</sup>. Depuis 1997, moins de 20 individus portant des mutations dans le gène *LEP* et moins de 50 dans le gène *LEPR* ont été identifiés<sup>86</sup>. Les patients déficients en *LEP* ou *LEPR* présentent des traits phénotypiques très semblables : une hyperphagie, une réduction du tonus nerveux sympathique, un hypogonadisme, une légère hypothyroïdie, une immunité lymphocytaire altérée et un retard de

développement à la puberté<sup>36</sup>. De plus, un retard de croissance peut également être observé chez des patients déficients en *LEPR*<sup>84</sup>.

#### 2.3.1.2. Le gène *POMC*

La prohormone POMC, codée par le gène *POMC*, peut être clivée en plusieurs agonistes ( $\alpha$ -, $\beta$ -, $\gamma$ -MSH) des récepteurs de la mélanocortine (MCR) par des modifications post-traductionnelles, permettant le fonctionnement de nombreuses fonctions biologiques (cf supra). Des mutations homozygotes dans le gène *POMC* ont été décrites pour la première fois chez deux patients en 1998. Une déficience en *POMC* cause une obésité sévère avec une hyperphagie, une insuffisance surrénalienne, et le plus souvent, des cheveux roux et une peau pâle<sup>90,91</sup>. Ces traits sont dus à l'absence d' $\alpha$ -MSH qui active normalement le récepteur MC1R, impliqué dans la pigmentation de la peau. Les patients peuvent parfois présenter aussi d'autres déficiences endocriniennes comme une insuffisance thyrotropique, somatotropique et gonadotropique<sup>84</sup>. En revanche, les mutations hétérozygotes de *POMC* n'ont qu'un effet modeste sur l'IMC et ne contribuent pas au développement de l'obésité monogénique<sup>92</sup>.

#### 2.3.1.3. Le gène *PCSK1*

PC1/3, codée par le gène *proprotein convertase subtilisin/kexin type 1 (PCSK1)*, est une endoprotéase à sérine dépendante du calcium. Elle est exprimée principalement au niveau de l'hypophyse, du cerveau (hypothalamus, cervelet, cortex cérébral, hippocampe), du pancréas endocrine et dans les cellules entéro-endocrines de l'intestin. Cette protéine est caractérisée par une activité enzymatique responsable du clivage d'un nombre important de prohormones et de proneuropeptides en leur forme mature. Le site de clivage des substrats est généralement un site dibasique Lys-Arg ou Arg-Arg en C-terminal mais c'est également possible pour les sites Arg-Lys et Arg-X-X-Arg (ou X représente n'importe quel acide aminé). Ses substrats jouent un rôle indispensable dans la régulation de l'appétit et de l'homéostasie énergétique : la proghréline, le proglucagon, la proinsuline, le précurseur de la protéine AGRP, du peptide CCK, du *glucose-dependant insulino-tropic polypeptide* (GIP). Pour la maturation de certains substrats tels que le peptide neuroendocrinien CART, la progastrine et en particulier les hormones peptidiques dérivées de POMC, PC1/3 agit de concert avec la protéine PC2<sup>93</sup>.

Comme PC1/3 est responsable du clivage de POMC en  $\alpha$ -MSH qui va se fixer au récepteur MC4R, le gène *PCSK1* a un rôle important dans la voie leptine/mélanocortine (cf supra)<sup>94</sup>. Les mutations pathogènes homozygotes de *PCSK1* causent une obésité sévère et précoce, associée à un ensemble d'endocrinopathies regroupant une malabsorption intestinale, une hypoglycémie

post-prandiale, un déficit en hormone de croissance et en glucocorticoïdes, un hypogonadisme hypogonadotrope et un hypothyroïdisme<sup>94</sup>. Le phénotype des enfants atteints de ces mutations est une malabsorption intestinale sévère pendant les premières années de vie avant l'apparition d'une hyperphagie associée à une obésité sévère. Les conséquences de ces complications gastro-intestinales sont importantes (diarrhée chronique, déshydratation, perte de poids et acidose métabolique), ce qui augmente fortement le risque de mortalité pendant la période néonatale<sup>93</sup>. En revanche, bien que quelques études se soient intéressées aux mutations hétérozygotes rares de *PCSK1*<sup>95,96</sup>, leur impact réel sur l'obésité n'a pas été clairement défini.

#### 2.3.1.4. Le gène *MC4R*

La signalisation induite par l'activation du récepteur MC4R, codé par le gène *MC4R*, a un rôle décisif dans la balance énergétique. Le gène *MC4R* est un des premiers gènes de la voie leptine-mélanocortine qui a été étudié car il n'est constitué que d'un seul exon codant, facilitant son séquençage. En 1998, deux équipes de recherche ont décrit en parallèle deux familles différentes dont les individus avec une obésité étaient porteurs de mutations *frameshift* dans le gène *MC4R* à l'état hétérozygote<sup>97,98</sup>. Depuis cette découverte, plus d'une centaine de mutations situées dans le gène *MC4R* ont été identifiées au sein de plusieurs cohortes d'individus avec une obésité : la majorité était des mutations hétérozygotes faux-sens et peu de mutations homozygotes étaient observées<sup>33</sup>. Les porteurs de *MC4R* à l'état homozygote sont plus rares en Europe ; cependant, on les retrouve plus fréquemment dans les populations consanguines comme au Pakistan<sup>83</sup>. La prévalence de mutations dans *MC4R* est d'environ 5% dans les cohortes d'enfants avec une obésité et 2% dans les cohortes d'adultes avec une obésité, cette fréquence va jusqu'à 0,3% dans la population générale ce qui fait de *MC4R* la première cause connue d'obésité monogénique<sup>39</sup>.

Les patients porteurs de mutations dans *MC4R* présentent une obésité précoce due à une hyperphagie sévère mais sans désordre endocrinien associé. Les enfants porteurs de ces mutations montrent une hyperphagie marquée mais qui diminue avec l'âge<sup>81</sup>. Ce phénotype est caractérisé par un mode de transmission dominant avec une pénétrance incomplète. La sévérité des phénotypes varie en fonction de l'impact fonctionnel de la mutation sur le récepteur<sup>99</sup> et de facteurs environnementaux<sup>100</sup>. Les porteurs de variants homozygotes ou hétérozygotes composites présentent un phénotype clinique plus sévère que les hétérozygotes en terme d'hyperphagie, et souffrent de problème de croissance et d'hyperinsulinémie<sup>33,99</sup>.

#### 2.3.1.5. Le gène *MRAP2*

Le gène *MRAP2* code une protéine accessoire des GPCR qui permet le transport à la membrane et/ou les voies de signalisation induites par leurs récepteurs. *MRAP2* est notamment la protéine accessoire du récepteur MC4R. Récemment, sept mutations pathogènes perte de fonction de *MRAP2* identifiées chez des patients adultes et enfant en surpoids ou atteints d'obésité ont été décrites<sup>73</sup>. Les porteurs de ces mutations présentaient également un syndrome métabolique lié à une hyperglycémie et une hypertension<sup>73</sup>. Ce syndrome métabolique peut être expliqué par le fait que *MRAP2* est exprimé dans d'autres tissus métaboliques que l'hypothalamus, y compris dans les îlots de Langerhans et les cellules  $\beta$  pancréatiques<sup>36</sup>.

#### 2.3.1.6. Le gène *SIMI*

Le gène *single-minded family basic helix-loop-helix [bHLH] transcription factor 1 (SIM1)* code le facteur de transcription SIM1 qui est impliqué dans le développement et le fonctionnement du PVN de l'hypothalamus, permettant la régulation de l'expression de MC4R. Les mutations hétérozygotes du gène *SIMI* induisant une perte de fonction de la protéine causent une obésité sévère parfois accompagnée d'un spectre syndromique (incluant troubles neurodéveloppementaux et hypotonie) similaire à celle observée dans les cas de syndrome de Prader-Willi<sup>101,102</sup>.

#### 2.3.1.7. Le gène *BDNF*

Une haploinsuffisance dans le gène *BDNF*, provoquée par une inversion chromosomique *de novo*, a été décrite chez une patiente atteinte d'obésité sévère et précoce, d'hyperactivité et de troubles cognitifs avec une altération de la nociception<sup>103</sup>. De façon similaire, des micro-délétions dans le gène *BDNF* ont été identifiées chez un cas d'héritage maternel, touchant une mère et son enfant souffrant d'obésité et d'un retard de développement mais sans hyperactivité<sup>104</sup>. De plus, une méta-analyse a montré une association significative entre l'obésité et des niveaux réduits de BDNF dans le sang à la fois chez les adultes, les enfants et les adolescents. Les personnes avec une obésité présentent des niveaux plus bas de BDNF par rapport aux personnes avec un poids normal<sup>105</sup>. La délétion du gène *BDNF* contribue parfois au retard mental observé dans certains cas de syndrome WAGR qui est un trouble génétique rare caractérisé par la présence de plusieurs anomalies congénitales incluant une tumeur de Wilms, une aniridie, des anomalies génito-urinaires et un retard mental. Cependant, la présence de délétions du gène *BDNF* n'est pas systématique dans tous les cas de syndrome WAGR, et d'autres facteurs génétiques sont également impliqués<sup>106</sup>.

#### 2.3.1.8. Le gène *NTRK2*

Le gène *neurotrophic tyrosine kinase receptor 2 (NTRK2)* code le récepteur de facteur neurotrophique TrkB qui est activé par des neurotrophines incluant le BDNF. Lorsqu'il est activé, TrkB déclenche une cascade de signalisation intracellulaire qui régule la croissance, la différenciation et la survie des neurones qui est située en aval de MC4R. Une mutation perte de fonction hétérozygote dans *NTRK2* a été décrite chez un enfant atteint d'une obésité précoce et sévère associée avec un retard mental, des troubles du développement et des anomalies dans certaines fonctions neurologiques comme une altération de la mémoire, de l'apprentissage et de la nociception<sup>107</sup>.

#### 2.3.1.9. Le gène *SH2B1*

Le gène *Src homology 2 B adaptor protein 1 (SH2B1)* code la protéine adaptatrice SH2B1 de la voie intracellulaire de plusieurs hormones dont la leptine et l'insuline. Des mutations hétérozygotes et perte de fonction dans *SH2B1* ont été décrites chez des patients présentant une obésité sévère et précoce, une résistance à l'insuline avec une hyperinsulinémie ainsi qu'une petite taille. Ces patients présentaient également un retard de développement de la parole et du langage avec une tendance à l'isolement social accompagné d'un comportement agressif<sup>108</sup>. De plus, des porteurs de délétions situées dans le chromosome 16p11.2, couvrant notamment *SH2B1* présentaient une obésité sévère avec de l'hyperphagie et une résistance sévère à l'insuline<sup>109</sup>.

#### 2.3.1.10. Le gène *ADCY3*

Alors que la plupart des gènes impliqués dans l'obésité monogénique ont été identifiés dans des formes extrêmes d'obésité, *ADCY3* a été d'abord découvert par GWAS abritant des variants fréquents associés à l'obésité commune, puis a été ultérieurement confirmé pour jouer un rôle dans des formes rares d'obésité monogénique<sup>39</sup>. *ADCY3* code l'adénylate cyclase, située dans la membrane plasmique, qui catalyse la synthèse d'AMPc à partir d'ATP. Son absence provoque un défaut de la signalisation de l'AMPc dans la voie initiée par MC4R. De plus, *ADCY3* se localise conjointement avec MC4R au niveau des cils primaires des neurones du PVN qui sont requis spécifiquement pour le contrôle de l'homéostasie énergétique par MC4R<sup>110</sup>. Il a été démontré que les porteurs de mutations homozygotes ou hétérozygotes composites de ce gène présentent une obésité sévère ainsi qu'une anosmie (perte de l'odorat)<sup>111</sup>.

#### 2.3.1.11. Le gène *ASIP*

Le gène *agouti signaling protein (ASIP)* code la protéine ASIP partageant une structure similaire avec AGRP et qui agit également en tant qu'antagoniste des MCR. La protéine ASIP

est normalement exprimée dans les cellules de la peau afin de réguler la pigmentation par son interaction avec MC1R. Lorsque l'on induit l'expression de son homologue *Agouti* chez la souris, cela génère le modèle murin bien connu *agouti*, qui est caractérisé par un pelage jaune, une obésité liée à l'augmentation de la prise alimentaire, au stockage des lipides dans le tissu adipeux et à une réduction de la dépense énergétique<sup>112</sup>. Récemment, une mutation causée par un réarrangement chromosomique dans ce gène provoquant un phénotype similaire a été identifiée chez l'humain<sup>113</sup>. Cette mutation implique une duplication en tandem dans le gène *ASIP*, conduisant à une expression ubiquitaire et ectopique de la protéine. Lorsque son expression est dérégulée, la protéine ASIP peut stimuler la prise alimentaire et provoquer une diminution de la pigmentation de la peau suite à sa fixation sur les MCR. Le phénotype des patients porteurs de cette mutation est caractérisé par une obésité sévère, de l'hyperphagie, une grande taille, une hyperinsulinémie et une hypopigmentation similaire à celle observée chez les patients déficients en *POMC*<sup>113</sup>.

#### 2.3.1.12. Autres gènes liés à la voie leptine-mélanocortine

Récemment, des mutations pathogènes associées à une obésité précoce et sévère ont été identifiées chez plusieurs autres gènes qui interviennent dans la régulation de la voie leptine-mélanocortine.

La kinase suppresseur de Ras2 (**KSR2**) est impliquée dans plusieurs voies de signalisation notamment la voie Raf-MEK-ERK. Plusieurs variants rares dans *KSR2* ont été identifiés chez 44 individus présentant une obésité sévère avec une basse fréquence cardiaque, une réduction du taux basal métabolique et une sévère résistance à l'insuline<sup>114</sup>.

Les sémaphorines de classe 3 (**SEMA3**), par l'intermédiaire de leurs liaisons aux récepteurs plexinesA et co-récepteurs neuropiline participent au développement des circuits de certains neurones hypothalamiques incluant ceux exprimant POMC. 40 variants rares hétérozygotes, identifiés par WES et situés dans les gènes *SEMA3A*, *SEMA3B*, *SEMA3C*, *SEMA3D*, *SEMA3E*, *SEMA3F* et *SEMA3G* (codant les sémaphorines 3A, 3B, 3C, 3D, 3E, 3F et 3G), ainsi que dans les gènes *PLXNA1*, *PLXNA2*, *PLXNA3*, *PLXNA4* (codant les plexines A1, A2, A3 et A4) et des gènes des *NRP1* et *NRP2* (codant les neuropilines 1 et 2), étaient significativement enrichis dans 982 individus atteints d'obésité sévère comparativement à 4449 contrôles<sup>115</sup>.

Le *nuclear receptor coactivator 1* (**NCOA1**), également appelé *steroid coreceptor activator-1* (SRC-1), interagit avec STAT3 afin d'activer la transcription de *POMC* dans les neurones POMC. La délétion de *Ncoa1* dans les neurones POMC chez la souris suivant un régime riche

en graisses provoque une diminution de la dépolarisation induite par LEPR ainsi que de l'expression de *POMC* et une augmentation de la prise alimentaire, menant à une obésité. Chez l'humain, 15 variants hétérozygotes dans *NCOA1* ont été identifiés chez des patients souffrant d'obésité sévère<sup>116</sup>.

A partir d'un WES réalisé sur 3665 individus, 13 variants rares hétérozygotes ont été identifiés dans le gène *serotonin 2C receptor (5-HTR2C)* chez 19 individus non apparentés. Ces 19 individus étaient atteints d'obésité sévère, d'hyperphagie et présentaient des troubles du comportement à des degrés divers. Parmi ces 13 variants rares, 11 entraînaient une perte de fonction par analyse *in vitro*. Les chercheurs ont également développé un modèle de souris portant un variant très perte de fonction de *HTR2C* (Knock-in *Htr2c*<sup>F327L</sup>). Ces souris développaient une obésité et un comportement social d'exploration réduit. De plus, elles présentaient une altération de la dépolarisation de leurs neurones POMC lorsqu'ils étaient stimulés avec un agoniste sérotoninergétique<sup>117</sup>.

De plus, une mutation homozygote dans le gène *tubby-like protein (TUB)* a aussi été identifiée par WES chez un sujet présentant une obésité précoce, une cécité nocturne, une diminution de l'acuité visuelle et des traits électrophysiologiques d'une dystrophie des cônes en des bâtonnets<sup>118</sup>. Enfin, une mutation homozygote a également été identifiée dans le gène *CPE* chez une femme en obésité sévère avec une déficience intellectuelle, un DT2 et un hypogonadisme hypogonadotrope<sup>119</sup>.

### 2.3.2. Les obésités monogéniques liées à des syndromes

Dans certains cas, les formes d'obésité monogénique sont associées à des traits phénotypiques particuliers tels que des troubles neurodéveloppementaux et/ou dysmorphiques. Ces types d'obésité sont couramment regroupés sous le nom d'obésité syndromique. Aujourd'hui plus de 79 syndromes différents associés avec une obésité ont été décrits. Parmi les syndromes les plus connus, on retrouve celui de Prader-Willi (PWS) et de Bardet-Biedl (BBS)<sup>120</sup>. Les défauts génétiques sous-jacents à ces syndromes sont souvent des anomalies chromosomiques qui englobent plusieurs gènes, augmentant les difficultés pour déterminer les mécanismes biologiques précis liés à la régulation du poids du corps<sup>39</sup>. De plus, la prévalence extrêmement basse de la plupart de ces syndromes rend difficile l'identification de leur cause génétique. C'est pourquoi, sur 79 syndromes reportés, 19 ont été complètement élucidés génétiquement, 11 sont partiellement élucidés, 27 ont été cartographiés dans une région chromosomique et aucun *loci* ou aucun gène n'a été identifié pour les 22 syndromes restants<sup>120</sup>.

### 2.3.2.1. Le syndrome de Prader-Willi (PWS)

Le PWS est la forme d'obésité syndromique observée le plus couramment : il affecte 1 naissance sur 20000 à 25000 dans toutes les ethnies. Le PWS est caractérisé par une évolution en deux phases : d'abord une hypotonie néonatale avec des difficultés pour s'alimenter liées à une mauvaise succion à la naissance suivie par une hyperphagie qui peut mener à une obésité précoce pendant l'enfance<sup>26</sup>. Ces patients peuvent également présenter de l'hypogonadisme, une petite taille et de petites mains et pieds suite à une déficience en hormone de croissance, des troubles neurologiques (déficience intellectuelle variable, difficultés comportementales) et une légère dysmorphie faciale<sup>10,81,121</sup>.

Les causes génétiques du syndrome de Prader-Willi sont complexes. Le PWS est un désordre chromosomique lié à l'absence physique ou fonctionnelle de la région 15q11.2-q13 sur le chromosome paternel. La plupart des gènes situés dans cette région sont soumis à l'empreinte parentale et à l'expressivité paternelle, signifiant que seuls les allèles du chromosome d'origine paternel sont actifs. Les mêmes allèles du chromosome 15 d'origine maternelle sont réduits au silence par des mécanismes épigénétiques, principalement par la méthylation. De plus, cette région comprend un grand nombre d'ARN nucléolaires (snoRNA) impliqués dans la maturation des ARN ribosomiques de protéines neurorégulatrices et hormonales. Quand cette région est altérée, cela provoque des anomalies développementales et endocrines. Au moins 4 gènes sont localisés dans cette région et ils codent au moins 5 protéines différentes dont la fonction n'a pas encore été élucidée : *makorin 3* (MRKN3), *MAGE-like2* (MAGEL2), *necdine* (NDN), *nuclear pore associated protein 1* (NPAP1) et *SNRPN upstream reading frame-small nuclear ribonucleoprotein polypeptide* (SNURF-SNRPN)<sup>121</sup>. D'autres gènes inconnus supplémentaires sont probablement mutés, ce qui peut expliquer la variabilité des phénotypes de PWS<sup>122</sup>.

L'apparition d'un PWS peut être provoquée par trois mécanismes différents. Le plus commun (environ 70% des cas de PWS) est lié à une délétion, généralement *de novo*, dans la région 15q11.2-q13 provenant du chromosome d'origine paternel. Une forme moins commune de PWS (25% des cas) est causée par une disomie uniparentale maternelle (UPD), ce qui signifie que les deux copies du chromosome 15 sont héritées de la mère en même temps sans contribution du chromosome paternel. La forme rare du PWS (5% des cas) est due à des défauts d'empreinte parentale (microdélétion ou épimutation) entraînant une inactivation anormale de certains gènes sur le chromosome 15 ou à des anomalies chromosomiques (translocation ou inversion) de la région 15q11.2-q13<sup>26,121</sup>.

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

Le BBS, décrit pour la première fois en 1866, concerne 1 naissance sur 125000 à 1/175000 dans la population européenne. Le BBS est caractérisé par une obésité sévère et précoce, une rétinite pigmentaire, une polydactylie post-axiale congénitale, des difficultés d'apprentissage, un dysfonctionnement des reins et un hypogonadisme masculin. Des signes cliniques mineurs peuvent également être observés tels qu'une fibrose hépatique, une petite taille, des défauts d'élocution, des retards du développement et des troubles endocrinologiques comme du DT2, une hypercholestérolémie et des anomalies de reproduction. Compte tenu de la diversité des symptômes, un diagnostic précis nécessite une combinaison de 3 caractéristiques majeures et 2 caractéristiques mineures ou uniquement 4 caractéristiques majeures<sup>123</sup>. La sévérité de l'obésité chez les patients avec un BBS peut varier de moyenne à sévère et peut être réversible avec des restrictions alimentaires et de l'exercice<sup>10,26</sup>.

Le BBS est une maladie hautement hétérogène qui est héritée d'une manière autosomale récessive avec une pénétrance variable. Actuellement, on dénombre plus de 20 gènes différents dont les mutations génétiques sont impliquées dans le développement du BBS : *BBS1*, *BBS2*, *BBS3/ARL6*, *BBS4*, *BBS5*, *BBS6/MKKS*, *BBS7*, *BBS8/TCC8*, *BBS9/PTHB1*, *BBS10*, *BBS11/TRIM32*, *BBS12*, *BBS13/MKSI*, *BBS14/CEP290*, *BBS15/WDPCC*, *BBS16/SDCCAG8*, *BBS17/LZTFL1*, *BBS18/BBIP1*, *BBS19/IFT27* et *BBS20/AZII*. La plupart des cas de BBS dans la population caucasienne (40-50%) présentent des mutations soit dans le gène *BBS1* ou soit dans le gène *BBS10*<sup>123</sup>. Les mécanismes exacts liant ces gènes à l'obésité n'ont pas été clairement établis mais ils sont tous impliqués dans la fonction du cil primaire. C'est pourquoi le BBS est défini souvent comme une ciliopathie. Les ciliopathies, ou dysfonctions du cil primaire, se manifestent la plupart du temps sous une forme syndromique avec une grande variabilité phénotypique liée à l'atteinte de plusieurs organes. Les cils primaires sont des organites présents à la surface des cellules qui ont comme rôle principal de détecter, d'intégrer et transmettre les signaux intra- et extracellulaires. Cependant, comment un défaut dans la fonction ciliaire peut causer l'obésité est inconnu<sup>10</sup>. Une première hypothèse est liée à un dysfonctionnement hypothalamique. En effet, il a été démontré chez la souris que la diminution de nombre de cils et leur mauvais fonctionnement dans les neurones POMC peuvent altérer la signalisation neuroendocrine, menant à une dérégulation de l'appétit qui peut conduire à de l'obésité et de l'hyperphagie<sup>26</sup>. De plus, les protéines BBS sont particulièrement requises pour la localisation de LEPR dans l'hypothalamus. Une autre hypothèse implique le tissu adipeux et d'autres tissus endocriniens. Il a été mis en évidence que les protéines BBS ont un rôle

régulateur dans la différenciation adipocytaire. En effet, les souris porteuses de mutations des gènes *BBS* avaient un développement accru du tissu adipeux malgré une alimentation contrôlée. De plus, il est possible que les récepteurs des différentes hormones intestinales soient ciliaires et par conséquent, l'inactivation des protéines BBS altèrent leur voie de signalisation<sup>81</sup>. Toutes ces hypothèses mettent en évidence que, malgré le contexte monogénique du BBS, la cause de l'obésité est une combinaison de plusieurs déséquilibres simultanés de cascades de signalisation endocriniennes et de processus de régulation de différenciation cellulaire de plusieurs tissus<sup>124</sup>.

#### 2.3.2.3. Le syndrome d'Alström

Le syndrome d'Alström affecte environ 1 naissance sur 1000000 dans la population. Il est caractérisé par une série de symptômes, parmi lesquels une obésité précoce, une dégénérescence rétinienne précoce, un DT2, une résistance à l'insuline et une surdité progressive sont les plus fréquents. De plus, des manifestations telles qu'une cardiomyopathie, une dyslipidémie et des dysfonctionnements au niveau du foie, des poumons et des reins peuvent également être présents. Les symptômes varient considérablement en termes de présentation et de gravité, rendant difficile le diagnostic<sup>125</sup>. Cependant, ces symptômes mènent généralement à des dommages aux organes, menant à une espérance de vie excédant rarement l'âge de 50 ans pour les individus atteints de ce syndrome<sup>126</sup>.

Le syndrome d'Alström, tout comme le BBS, est classé comme une ciliopathie. Il est causé par des mutations dans le gène *Alström syndrome 1 (ALMS1)* qui se transmettent de manière autosomale récessive. La protéine ALMS1 est localisée à l'intérieur des centrosomes et des corps basaux des cellules ciliaires. Bien que la fonction normale de ALMS1 et les raisons pour lesquelles son dérèglement entraîne la maladie ne soient pas claires, les données suggèrent qu'elle est impliquée dans la fonction ciliaire, particulièrement dans le trafic intracellulaire et le transport protéique<sup>127</sup>.

#### 2.3.2.4. Le syndrome lié à la délétion du chromosome 16p.11.2

Le syndrome de délétion 16p.11.2 est causé suite à la perte d'une portion spécifique du bras court du chromosome 16, précisément en position p.11.2. Ces délétions englobent plusieurs gènes impliqués dans les voies de signalisation de la leptine et de l'insuline. La forme de délétion 16p.11.2 la plus commune, avec une prévalence dans la population d'environ 1/2000, fait environ 600 kilobases (kb) dans la région proximale aux points de cassures 4 et 5 (BP4-5) comprenant 33 gènes<sup>109,128</sup>. Cette délétion est associée à une obésité précoce pouvant être accompagnée d'une déficience intellectuelle, d'un retard de développement, de troubles du

spectre autistique, de macrocéphalie et d'épilepsie. Les symptômes les plus souvent observés sont des troubles de coordination et du langage. Une délétion distale d'environ 220 kb dans la région BP2-3 située dans ce même locus a également été associée avec une obésité précoce, une macrocéphalie et un trouble du spectre autistique. Cette microdélétion comprend 9 gènes dont le gène *SH2B1*<sup>129</sup>. Dans la plupart des cas, ces délétions apparaissent *de novo*, mais elles peuvent également être transmises de manière autosomale dominante dans approximativement 7% des cas<sup>130</sup>. Les individus atteints de ce syndrome développent souvent une prise de poids vers l'âge de 9 ans qui s'accélère ensuite avec de l'hyperphagie. La prévalence de ce syndrome est difficile à évaluer en raison de la variabilité extrême des phénotypes des porteurs. En effet, certains individus peuvent présenter des symptômes très légers ou modérés, tandis que d'autres peuvent présenter des troubles beaucoup plus graves<sup>130</sup>.

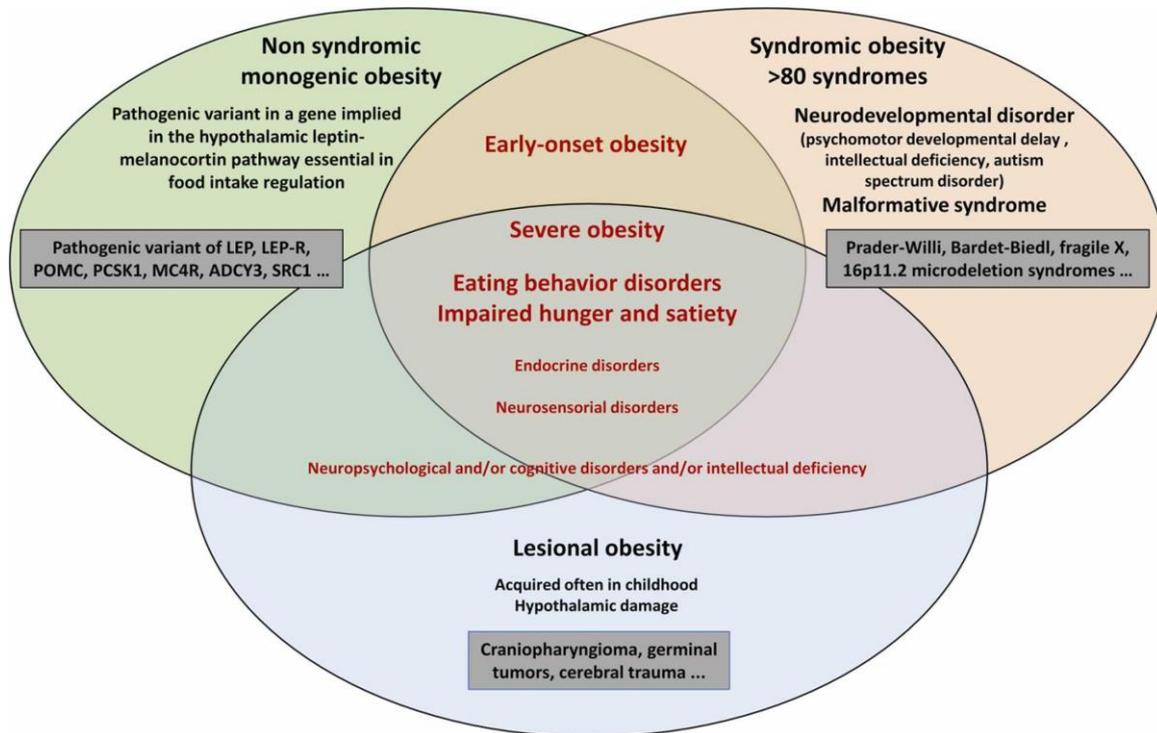
#### 2.3.2.5. Les obésités abdominales associées à un syndrome métabolique (AOMS)

En plus des obésités monogéniques associées à des problèmes neurodéveloppementaux, il existe également des obésités monogéniques liées à des syndromes métaboliques. Le syndrome regroupant une obésité abdominale, des triglycérides élevés, des niveaux bas de lipoprotéines de cholestérol à haute densité, une hypertension et un niveau de glucose élevé à jeun est appelé syndrome métabolique (ou syndrome X)<sup>131</sup> ou obésité abdominale associée à un syndrome métabolique (AOMS)<sup>132</sup>. Ce syndrome est un facteur de risque majeur pour le DT2 et les maladies cardiovasculaires. Le phénotype des patients atteints de ce syndrome est caractérisé par une maladie précoce des artères coronaires, une obésité centrale, une hypertension et un DT2. L'AOMS est transmis de manière autosomale dominante. Il existe plusieurs sortes d'AOMS selon le gène ou le locus identifié comme responsable de la maladie :

- L'**AOMS1** est lié à la région génétique située sur le chromosome 3q27<sup>133</sup>.
- La région génétique liée à l'**AOMS2** a été cartographiée sur le chromosome 17p12<sup>133</sup>.
- Les mutations causales de l'**AOMS3** se situent dans le gène *dual specificity tyrosine phosphorylation regulated kinase 1B (DYRK1B)* sur le chromosome 19q13. En 2014, il a été mis en évidence l'association autosomale dominante entre deux mutations rares dans *DYRK1B* avec ce syndrome métabolique au sein de trois familles<sup>134</sup>. Cependant, depuis 2014, cette étude n'a jamais été répliquée.
- L'**AOMS4** est causée par une mutation hétérozygote dans le gène *Chymotrypsin Like Elastase 2A (CELA2A)* situé sur le chromosome 1p36<sup>135</sup>.

### 2.3.3. Limites entre les formes d'obésités monogéniques dites « syndromiques » et « non syndromiques »

Les obésités monogéniques causées par des mutations de gènes impliqués dans la voie leptine-mélanocortine sont souvent appelées « obésités non syndromiques » afin de les distinguer des obésités syndromiques. Cependant, les limites entre ces deux catégories ne sont pas clairement définies car il existe des phénotypes similaires entre ces deux types d'obésité monogénique (Figure 9)<sup>84</sup>.



**Figure 9 : Chevauchement entre les différentes formes d'obésité monogéniques.** Source: Dubern, B., Faccioli, N., Poitou, C. & Clément, K., « Novel therapeutics in rare genetic obesities: A narrative review », *Pharmacological Research*, 2023.

Par exemple, des patients présentant des « obésités non syndromiques » peuvent présenter des troubles dans le neurodéveloppement et/ou psychiatriques ainsi que des désordres endocriniens similaires à ceux observés dans les cas du syndrome de Prader-Willi. Cela a été observé chez des patients présentant des mutations dans les gènes *SIMI*, *BDNF* ou *NTRK2*<sup>81</sup>. De plus, les patients présentant des mutations dans le gène *MRAP2* présentent une obésité monogénique qualifiée de « non syndromique », bien qu'elle soit associée à un syndrome métabolique avec une hypertriglycémie, une hypocholestérolémie, ainsi qu'une hyperglycémie et une hypertension<sup>73</sup>.

A l'inverse, des anomalies dans la voie leptine-mélanocortine ont été décrites chez des patients atteints d'obésité syndromique. Par exemple, une déficience en PC1/3 et une altération des neurones AGRP ont été décrites chez des patients atteints de PWS<sup>136</sup>. Par ailleurs, une altération du transport de LEPR et de sa localisation dans la membrane ciliaire des neurones POMC sont impliqués dans le développement de l'hyperphagie et de l'obésité chez les patients atteints de BBS<sup>137</sup>.

#### 2.3.4. Continuum entre les formes polygéniques et monogéniques de l'obésité

La découverte de nouveaux gènes liés à l'obésité est généralement classée en fonction de la fréquence allélique et de la prévalence de la maladie : les mutations rares entraînent des formes monogéniques et les variants fréquents sont associés aux formes communes polygéniques. Cependant, cette classification est en train d'évoluer vers un continuum entre ces deux formes<sup>138</sup>. En effet, il a été clairement montré que des variants pathogènes rares pouvaient être identifiés chez des patients étiquetés 'polygénique' (c'est-à-dire avec une obésité non sévère et tardive). En outre, les formes monogéniques et polygéniques élucidées jusqu'à présent partagent en partie les mêmes gènes. Par exemple, un nombre important de SNP associés à l'obésité commune et à l'IMC, identifiés par les GWAS, se retrouvent également dans les gènes impliqués dans les cas d'obésités extrêmes et précoces chez les humains ou les modèles animaux : *MC4R*, *BDNF*, *SH2B1*, *POMC*, *LEP*, *EPR*, *NPY*, *SIMI*, *NTRK2*, *PCSK1* et *KSR2*. Alors qu'une altération génétique des composants de la voie leptine-mélanocortine et de la voie de signalisation BDNF-NTRK2 entraîne une obésité sévère, les SNP présents dans ou près de ces mêmes gènes qui ont un effet moins important sur leur expression va influencer la position de l'individu dans la distribution normale d'IMC<sup>39</sup>.

Par ailleurs, l'expression de mutations responsables d'obésité monogénique est, du moins en partie, influencée par la prédisposition polygénique des individus à l'obésité<sup>39</sup>. Une récente étude a examiné les différences entre les individus porteurs de mutations pathogènes de *MC4R* ayant un poids normal (15%) et ceux présentant une obésité (85%). Les mutations pathogènes de *MC4R* avaient une pénétrance  $\geq 30\%$  et un OR  $\geq 2$ . Les chercheurs ont remarqué que les 28 porteurs avec un poids normal (sur 182 individus) avaient en fait d'autres variants génétiques qui les prédisposaient à un poids corporel plus bas, compensant ainsi le risque causé par la déficience en *MC4R*. Pour ce faire, ils ont évalué la contribution polygénique de ces porteurs en calculant un score de risque polygénique (PRS) pour l'IMC à l'échelle du génome. Le PRS de l'IMC chez les porteurs *MC4R* avec un poids normal était significativement plus bas que celui des porteurs *MC4R* avec une obésité et tendait à être plus bas que celui des non-porteurs

avec un poids normal. De plus, par rapport aux non porteurs ayant un PRS bas, les porteurs *MC4R* ayant un PRS bas avaient un risque de développer une obésité qui était multiplié par 2 et un poids environ 6% plus élevé, alors que les porteurs *MC4R* avec un PRS élevé avaient un risque qui était multiplié par 9 et un poids environ 25% plus élevé que les non porteurs ayant un PRS bas. Combinées ensemble, ces données indiquent qu'une faible prédisposition polygénique à l'obésité semble atténuer l'impact des mutations pathogènes de *MC4R*<sup>100</sup>. Ces résultats montrent l'importance des données de population à large échelle pour évaluer plus précisément l'impact des mutations de *MC4R* sur les obésités précoces extrêmes. Cette étude a démontré que malgré le rôle critique de *MC4R* dans l'obésité, les effets des mutations pathogènes de *MC4R* sur l'obésité peuvent être compensés, au moins en partie, par un score de risque polygénique bas représentant d'autres mécanismes innés impliqués dans la régulation du poids du corps<sup>100</sup>.

Cette découverte met en évidence que la régulation du poids du corps est le résultat d'une interaction complexe entre les mutations rares avec un large impact sur l'obésité et la prédisposition génétique globale déterminée par les variants génétiques communs avec un effet faible.

### 3. Traitements et médecine de précision

Une prise en charge efficace de l'obésité est importante afin d'améliorer les comorbidités, les facteurs de risque, la qualité de vie et la mobilité du patient. Le traitement proposé varie selon la sévérité de l'obésité, les comorbidités du patient et également selon sa génétique.

#### 3.1. Changement du mode de vie

Le principal traitement existant pour les patients atteints d'obésité repose sur la modification de leur environnement obésogène. Les deux grands axes de prise en charge se basent sur la modification du comportement alimentaire en réduisant et en modifiant les apports alimentaires journaliers et la lutte contre la sédentarité en favorisant l'activité physique<sup>139</sup>. Il est proposé aux patient souffrant d'obésité de perdre au minimum 10% de leur poids corporel à l'aide d'un régime alimentaire adapté, d'activité physique et de thérapie comportementale<sup>15</sup>. Une approche psychologique peut être également nécessaire à la prise en charge, particulièrement en cas de trouble du comportement alimentaire et de trouble dépressif<sup>139</sup>.

#### 3.2. Chirurgie bariatrique

La chirurgie bariatrique est le traitement de référence, elle est aujourd'hui un traitement efficace connu sur le long terme contre l'obésité commune et sévère<sup>140</sup>. Dans la majorité des cas, elle donne des résultats indiquant une diminution significative de l'IMC et des comorbidités. Cependant, elle n'est efficace qu'à condition que le patient modifie ses habitudes alimentaires et augmente son activité physique avec une prise de compléments nutritionnels tels que des protéines, des minéraux et des vitamines tout au long de sa vie. Elle n'est à envisager qu'en dernier recours car c'est une chirurgie lourde qui peut entraîner des complications au quotidien comme des problèmes liés à la chirurgie, des carences nutritionnelles et des difficultés suite à la modification de l'image du corps et des relations avec les autres<sup>141</sup>.

En France, elle est proposée uniquement aux adultes dont l'IMC est supérieur à 35 kg/m<sup>2</sup> et qui présentent au moins une comorbidité susceptible d'être améliorée par la chirurgie (le DT2, l'hypertension, le syndrome de l'apnée du sommeil ou des troubles articulaires) lorsque la modification du mode de vie et les traitements médicamenteux ont échoué<sup>15</sup>. Concernant les enfants, le recours à la chirurgie bariatrique est déconseillé. La Haute Autorité de Santé (HAS) a déclaré que la chirurgie bariatrique ne devait pas être envisagée pour les mineurs de moins de 13 ans. Pour les adolescents de plus de 15 ans, elle doit rester exceptionnelle et elle n'est

réservée qu'à des cas spécifiques d'obésité sévère lorsque des approches pluridisciplinaires de prise en charge n'ont pas abouti<sup>142</sup>.

La chirurgie bariatrique se base sur une modification du système digestif afin de diminuer la quantité d'aliments ingérée ainsi que leur assimilation par l'organisme. Les deux techniques les plus courantes qui suivent le principe de restriction et qui réduisent la taille de l'estomac sont l'anneau gastrique ajustable et la gastrectomie longitudinale (ou *sleeve gastrectomy*). Parmi les différentes techniques de chirurgie bariatrique, seule la pose de l'anneau gastrique est réversible. Elle implique la mise en place d'un anneau autour de la partie supérieure de l'estomac afin de diminuer sa taille, ce qui permet d'atteindre la sensation de satiété plus rapidement. Quant à la gastrectomie longitudinale, elle consiste à retirer les 2/3 de l'estomac afin que les aliments passent plus vite dans l'intestin. Les autres techniques existantes de chirurgie bariatrique sont mixtes. Elles associent une réduction de l'estomac avec une limitation de l'absorption intestinale : le *bypass* gastrique et la dérivation biliopancréatique. Le *bypass* de Roux-en-Y implique de court-circuiter une partie de l'estomac et de l'intestin. Les aliments vont alors directement dans la partie moyenne de l'intestin grêle où ils sont assimilés en moindre quantité. Enfin, la dérivation biliopancréatique est réservée uniquement aux patients dont l'IMC dépasse 50 kg/m<sup>2</sup> ou en cas d'échec d'une autre technique. Cette opération implique une réduction de la taille de l'estomac par gastrectomie, suivie d'un court-circuitage d'une partie de l'intestin grêle, ce qui entraîne une diminution de l'absorption intestinale<sup>141</sup>.

En outre, la chirurgie bariatrique est plutôt déconseillée pour les individus atteints d'obésité monogénique car elle se révèle moins efficace et augmente le risque de récurrence chez ces patients. En effet, des études ont démontré, à la fois chez la souris et l'homme, que les individus déficients en *MC4R* ne maintenaient pas leur réduction de poids après une chirurgie bariatrique et reprenaient leur poids après une courte période<sup>143</sup>. De plus, l'efficacité de cette pratique chez les porteurs de variants hétérozygotes de *MC4R* n'a pas été clairement démontrée<sup>33</sup>.

### 3.3. Médecine de précision

#### 3.3.1. Importance du diagnostic génétique de l'obésité monogénique

La découverte et les investigations sur les causes des formes monogéniques d'obésité permettent aujourd'hui de proposer des traitements focalisés sur le contrôle de la prise alimentaire. En fonction du gène ciblé par une mutation rare, on peut proposer un traitement adapté au patient afin de maximiser son effet et minimiser les effets secondaires (cf infra). Cela permet également de limiter les éventuelles chirurgies bariatriques<sup>84</sup>.

Dans le cas d'obésité monogénique associée à un syndrome, des tests génétiques doivent être réalisés par des spécialistes de centres de référence en fonction du syndrome suspecté. Pour le PWS, le diagnostic génétique commence par la recherche d'une anomalie du profil de méthylation de la région du chromosome 15. Le mécanisme génétique responsable de l'anomalie peut ensuite être déterminé par un caryotype de haute-résolution. Dans le cas de suspicion de BBS, la manifestation congénitale de polydactylie est souvent sous-estimée. Le diagnostic se base la plupart du temps sur le développement d'une obésité précoce et un examen oculaire vers 8-9 ans. Une échographie rénale et une pyélographie intraveineuse peuvent aider également pour diagnostic des anomalies rénales<sup>123</sup>.

En cas d'obésité sévère précoce associée à des anomalies endocriniennes suggérant une obésité monogénique, il est essentiel de confirmer le diagnostic par un séquençage direct des gènes candidats<sup>81</sup>. Le WES permet de tester rapidement un large panel de gènes identifiés comme étant impliqués dans les formes monogéniques d'obésité à moindre coût. Lorsqu'un variant identifié par séquençage de l'ADN est situé dans un des gènes connus pour être impliqués dans les formes rares d'obésité, sa pathogénicité est évaluée en se basant sur des critères établis par l'*American College of Medical Genetics and Genomics* (ACMG)<sup>144</sup>. Cette liste de critères comprend différentes preuves de pathogénicité considérées comme très fortes (PVS), fortes (PS), modérées (PM) ou secondaires (PP) basées sur des données épidémiologiques, informatiques et fonctionnelles (**Tableau 4**). Parmi ces critères, la preuve de pathogénicité la plus forte, le critère PVS1, concerne les variants dits « nuls » d'après une définition purement génétique. Ces variants regroupent les variants non-sens (insertion d'un codon stop), les variants décalant le cadre de lecture (*frameshift*), les variants d'épissage canonique situés à plus ou moins 1 ou 2 nucléotides du site d'épissage (*canonical splice donor* ou *acceptor*), les variants affectant le codon initiateur (*initiation codon*), et les variants induisant la délétion d'un ou de plusieurs exons (*single* ou *multiexon deletion*).

La détermination de la pathogénicité d'un variant (pathogène, probablement pathogène, de significativité incertaine (VUS), probablement bénin ou bénin) est basée sur la combinaison des critères ACMG, suivant des règles spécifiques (**Tableau 5**). Grâce à ces critères, il est possible de déterminer si un variant nouvellement identifié dans un gène connu a un impact sur la fonction du gène<sup>144</sup>. Cette caractérisation de la pathogénicité des variants identifiés chez les patients permet de préciser le diagnostic génétique.

**Tableau 4 : Preuve de pathogénicité suivant les critères ACMG.** Source: Richards et al., « Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology », *Genetics in Medicine*, 2014.

Preuve de pathogénicité	Critère	Définition
<b>Preuve très forte de pathogénicité (very strong)</b>	PVS1	Variant nul (variant non-sens, variant de décalage de lecture ( <i>frameshift</i> ), variant d'épissage canonique situé à plus ou moins 1 ou 2 nucléotides du site d'épissage, variant affectant le codon initiateur, variant induisant la délétion d'un ou de plusieurs exons) dans un gène pour lequel la perte de fonction est un mécanisme connu de la maladie
<b>Preuve forte de pathogénicité (strong)</b>	PS1	Variant induisant le même changement d'acide aminé qu'un autre variant déjà établi comme pathogénique selon les bases de données ClinVar ou HGMD
	PS2	Apparition du variant <i>de novo</i> (confirmation paternelle et maternelle) chez un patient malade sans histoire familiale
	PS3	Effet délétère sur le gène ou sur son produit démontré par des études fonctionnelles réalisées <i>in vitro</i> ou <i>in vivo</i>
	PS4	Prévalence plus importante du variant chez les cas que chez les contrôles
<b>Preuve modérée de pathogénicité (moderate)</b>	PM1	Variant situé dans un point chaud ( <i>hot spot</i> ) de mutations, ou dans un domaine fonctionnel critique ne contenant pas de variation bénigne (selon les bases de données ClinVar ou HGMD)
	PM2	Variant absent des individus contrôles (ou à très basse fréquence si récessif) dans des bases de données telles que Exome Sequencing Project, 1000 Genomes, gnomAD, ou ExAC
	PM3	Pour les maladies récessives, variant détecté <i>en trans</i> d'un variant pathogène
	PM4	Changement de la longueur de la protéine dû à des délétions/insertions dans le cadre de lecture dans une région non répétée ou des variants stop-loss (perte du codon STOP)
	PM5	Variant induisant une mutation non synonyme d'un acide aminé pour lequel des variants faux-sens pathogènes ont déjà été trouvés (selon les bases de données ClinVar ou HGMD)
	PM6	Variant apparu <i>de novo</i> , sans confirmation paternelle et maternelle
<b>Preuve secondaire de pathogénicité (supporting)</b>	PP1	Coségrégation avec la maladie chez plusieurs membres de la famille atteints (minimum deux membres), dans un gène connu pour provoquer la maladie
	PP2	Variant faux-sens dans un gène qui a un faible taux de variants faux-sens bénins, et dans lequel les variants faux-sens sont un mécanisme commun de la maladie
	PP3	Multiples preuves informatiques soutenant un effet délétère sur le gène ou son produit (utilisant PolyPhen-2, SIFT et Mutation Taster)
	PP4	Phénotype ou histoire familiale hautement spécifique d'une maladie ayant une étiologie génétique unique
	PP5	Variant décrit comme pathogène selon une source digne de confiance, mais pour lequel il n'est pas possible d'évaluer soi-même son effet délétère

**Tableau 5 : Règles de combinaison des critères pour la classification de variants pathogènes et probablement pathogènes, suivant les critères ACMG.** Source: Richards et al., « Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology », *Genetics in Medicine*, 2014.

Type de variant	Combinaison de critères possibles
<b>Pathogène (P)</b>	<p>(a) 1 critère très fort (PVS1) et :</p> <ul style="list-style-type: none"> <li>- ≥ 1 critère fort (PS1-4) OU</li> <li>- ≥ 2 critères modérés (PM1-6) OU</li> <li>- 1 critère modéré (PM1-6) et 1 critère secondaire (PP1-5) OU</li> <li>- ≥ 2 critères secondaires (PP1-5)</li> </ul> <p>(b) ≥ 2 critères forts (PS1-4)</p> <p>(c) 1 critère fort (PS1-4) et :</p> <ul style="list-style-type: none"> <li>- ≥ 3 critères modérés (PM1-6) OU</li> <li>- 2 critères modérés (PM1-6) et ≥ 2 critères secondaires (PP1-5) OU</li> <li>- 1 critère modéré (PM1-6) et ≥ 4 critères secondaires (PP1-5)</li> </ul>
<b>Probablement pathogène (LP)</b>	<p>(a) 1 critère très fort (PVS1) et 1 critère modéré (PM1-6)</p> <p>(b) 1 critère fort (PS1-4) et 1-2 critères modérés (PM1-6)</p> <p>(c) 1 critère fort (PS1-4) et ≥ 2 critères secondaires (PP1-5)</p> <p>(d) ≥ 3 critères modérés (PM1-6)</p> <p>(e) 2 critères modérés (PM1-6) et ≥ 2 critères secondaires (PP1-5)</p> <p>(f) 1 critère modéré (PM1-6) et ≥ 4 critères secondaires (PP1-5)</p>

Le diagnostic génétique présente de nombreux avantages : il permet d'adapter le traitement en fonction du patient pour restaurer la fonction manquante et éviter l'administration de traitements peu efficaces, d'expliquer certaines caractéristiques cliniques observées, d'anticiper les complications cliniques potentielles et les comorbidités possibles associées à l'obésité.

### 3.3.2. Leptine recombinante

Le tout premier exemple de médecine de précision appliqué à l'obésité concerne des injections sous-cutanées quotidiennes de leptine recombinante (métréleptine [Myalept], *Amylin Pharmaceuticals*) pour le traitement des individus atteints d'une déficience congénitale dans le gène *LEP* (**Figure 10**)<sup>145</sup>. Lors du premier essai clinique, une patiente âgée de 9 ans et pesant 94,4 kg a reçu des injections de leptine recombinante. Cette patiente a commencé à perdre du poids dans les deux semaines après le début du traitement à la leptine. Cette perte de poids s'est ensuite poursuivie pendant 12 mois, au cours desquels la patiente a perdu un total de 16,4 kg, perdant 1 à 2 kg par mois<sup>146</sup>. Par la suite, cette thérapie a montré des résultats remarquables chez d'autres patients déficients en *LEP*, entraînant une perte de poids significative, principalement de masse grasse, avec un effet majeur sur la réduction de la prise alimentaire et une normalisation de la fonction endocrine<sup>86,145</sup>.



**Figure 10 : Effet de l'administration de leptine recombinante chez un patient déficient en leptine.** Image reprenant un garçon de 3 ans atteint d'un déficit congénital en leptine, pesant 42 kg avant (à gauche) et 32 kg après (à droite) après 4 ans de traitement à la leptine recombinante. *Source: Farooqi, I. S., & O'Rahilly, S., « 20 YEARS OF LEPTIN: Human disorders of leptin action », Journal of Endocrinology, 2014.*

### 3.3.3. Setmélanotide

Le setmélanotide (RM-493) est un médicament développé par la société pharmaceutique *Rhythm Pharmaceuticals* dans le but de traiter les patients souffrant d'obésité monogénique. Il est commercialisé sous le nom d'*Imcivree*. Depuis novembre 2020, il a été autorisé par la *US Food and Drug Administration* (FDA) et depuis 2021 par l'Agence européenne du médicament (EMA) pour le traitement des patients déficients en *POMC*, *PCSK1* ou *LEPR* confirmé par un diagnostic génétique. Son principal avantage est qu'il peut être administré aux patients adultes ainsi qu'aux enfants dès l'âge de 6 ans. Depuis 2022, il a également été approuvé par la FDA et l'EMA pour soigner les patients atteints du BBS à partir de 6 ans<sup>84</sup>.

Le setmélanotide, un peptide cyclique composé de 8 acides aminés, est un agoniste de  $\alpha$ -MSH. Il se fixe sur le récepteur MC4R avec une constante d'inhibition de 2,1 nmol/l et son  $EC_{50}$  pour la stimulation d'AMPC après l'activation de MC4R est de 0,31 nmol/l. En se fixant sur MC4R, le setmélanotide rétablit la voie leptine-mélanocortine chez les patients déficients en *POMC*, *PCSK1* et *LEPR*, ce qui permet de restaurer la sensation de satiété après les repas. En diminuant l'appétit, le setmélanotide entraîne une réduction de la prise alimentaire ainsi que de l'obésité<sup>147</sup>. Il est délivré uniquement sur ordonnance et il doit être prescrit et suivi par un médecin expérimenté. Son administration se fait une fois par jour par voie sous-cutanée. Les principaux effets secondaires du setmélanotide sont une réaction au site d'injection et une hyperpigmentation de la peau. Cette hyperpigmentation est expliquée par le fait que le setmélanotide active également le récepteur MC1R exprimé dans les mélanocytes et le récepteur MC3R, mais avec une affinité 20 fois inférieure à celle pour MC4R. L'activation de MC1R produit une accumulation de mélanine, augmentant ainsi la pigmentation de la peau<sup>147</sup>. Cependant, il a comme avantage de ne provoquer aucune augmentation du rythme cardiaque ou de la pression sanguine, ce qui a favorisé son développement<sup>148</sup>.

Dans un premier temps, des études cliniques de phase 2 évaluant l'effet du setmélanotide sur des participants déficients dans les gènes *POMC* et *LEPR* ont montré des résultats convaincants<sup>149,150</sup>. En effet, dans la première étude clinique, deux patients avec une obésité déficients en *POMC* (un hétérozygote composite et un homozygote) ont reçu un traitement au setmélanotide : le patient 1 a perdu 25,8 kg après 13 semaines et le patient 2 a perdu 20,5 kg après 12 semaines<sup>149</sup>. De manière similaire, lors d'un traitement de 61 semaines, trois patients atteints d'obésité, homozygotes pour une mutation de *LEPR*, ont montré une réduction de poids significative: le patient 1 a perdu 21,6% de son poids (28,2 kg) après 26 semaines, le patient 2

a perdu 11,4% de son poids (13,9 kg) après 36 semaines, et le patient 3 a perdu 10 kg après 13 semaines<sup>150</sup>.

Suite aux résultats positifs obtenus avec le setmélanotide, une étude clinique de phase 3 a été réalisée sur 10 patients souffrant d'obésité liée à une déficience en *POMC* (homozygotes ou hétérozygotes composites) incluant 9 patients avec une déficience en *POMC* et 1 patient avec une déficience en *PCSK1*, ainsi que 11 patients avec une déficience en *LEPR* (homozygotes ou hétérozygotes composites). Après un an de traitement au setmélanotide, 8 individus sur 10 (80%) de l'essai sur *POMC* et 5 participants sur 11 (45%) de l'essai sur *LEPR* ont atteint une perte de poids corporel d'au minimum 10%. L'indice de faim a été évalué à l'aide d'un questionnaire basé sur une échelle de 0 à 10, rempli quotidiennement par les participants âgés d'au moins 12 ans, et dont la moyenne hebdomadaire a été calculée pour les analyses. Le pourcentage moyen de l'indice de faim était de -27,1 % pour l'essai *POMC* et de -43,7 % pour l'essai *LEPR*<sup>151</sup>. Ces résultats ont montré que le traitement au setmélanotide réduit significativement le poids corporel, ce qui peut entraîner des avantages en termes de morbidité, d'espérance de vie et de meilleure santé sur le long terme<sup>15</sup>.

Cependant, bien que le setmélanotide entraîne une perte de poids significative chez les souris hétérozygotes et homozygotes pour *MC4R*, un essai clinique de phase 1b réalisé en double aveugle sur 28 jours avec 6 patients hétérozygotes pour *MC4R* (porteurs de mutations entraînant une perte de fonction complète ou partielle) traités au setmélanotide n'a pas montré de résultats significatifs sur la perte de poids. Cette étude suggère que le setmélanotide semble être plus efficace chez les patients porteurs de mutations dans les gènes situés en amont de *MC4R*<sup>17</sup>. De plus, des études *in vitro* ont démontré que le setmélanotide pouvait restaurer la signalisation de *MC4R*, mais uniquement pour certaines mutations de *MC4R*, suggérant que l'évaluation de l'impact fonctionnel des variants génétiques sera une condition nécessaire avant de proposer le traitement chez l'homme<sup>36</sup>. L'une des conclusions de cette étude était que l'utilisation du setmélanotide présenterait probablement peu d'avantages chez les patients hétérozygotes<sup>152</sup>.

L'efficacité du setmélanotide pour le traitement des patients atteints du BBS a également été évaluée. Les résultats d'une étude de phase 2, publiée en 2020, ont montré des effets significatifs sur la perte de poids. En effet, après 3 mois de traitement, 8 patients sur 10 présentaient une diminution de poids d'au moins 5%, et après un traitement de 52 semaines, 7 patients sur 8 montraient une réduction de leur poids de 16,3% accompagnée d'une réduction de l'appétit<sup>153</sup>. Ces résultats ont conduit à la mise en place en 2019 d'une étude de phase 3 en double aveugle menée sur 38 patients (32 atteints du BBS et 6 atteints du syndrome d'Alström).

Après 52 semaines de traitement, 32,3% des patients (11 sur 31) atteints du BBS âgés de plus de 12 ans ont perdu au moins 10% de leur poids corporel, et 30,2% des participants (n = 16) avaient une diminution de leur indice de faim supérieur à 25%<sup>154</sup>. Cependant, cette étude n'a pas montré de résultats pour les patients atteints du syndrome d'Alström, car 2 des 3 patients de l'essai ont pris du poids avec le traitement, et le troisième a arrêté le traitement. Cela suggère la nécessité de mener des essais cliniques supplémentaires<sup>154</sup>.

Les résultats positifs obtenus avec l'administration du setmélanotide soulèvent la question de son utilisation potentielle pour soigner les patients hétérozygotes et/ou homozygotes pour les autres gènes de la voie leptine-mélanocortine liés à l'obésité. Dans ce contexte, une étude clinique de phase 3 (NCT05093634) est actuellement en cours, où le setmélanotide est testé sur des patients hétérozygotes pour *POMC*, *PCSK1* ou *LEPR*, ainsi que sur des patients homozygotes, hétérozygotes ou hétérozygotes composites pour *NCOAI* ou *SH2BI*.

#### 3.3.4. Autres traitements disponibles

Actuellement il existe peu d'autres traitements contre l'obésité. Parmi eux, on trouve l'orlistat (Xenical), le liraglutide (Saxenda), le semaglutide (Wegovy), le phentermine-topiramate (Qsymia), et le naltrexone-bupropion (Contrave). La plupart de ces médicaments agissent en diminuant l'appétit et en augmentant la sensation de satiété et ils sont destinés uniquement aux adultes<sup>15</sup>. En France, seuls l'orlistat, le liraglutide et le semaglutide sont acceptés comme traitement contre l'obésité. Le phentermine-topiramate et le naltrexone-bupropion sont autorisés aux USA, mais ils sont associés à des risques cardio-vasculaires et psychiatriques. L'orlistat, un inhibiteur des lipases gastro-intestinales qui permet de réduire l'absorption des graisses par l'estomac, présente une efficacité limitée (perte de 2 à 3 kg par an) et peut provoquer des effets secondaires parfois invalidants notamment au niveau digestif.

Le GLP-1 est une hormone intestinale libérée depuis l'intestin grêle et le colon en fonction des calories ingérées. Il régule la sécrétion d'insuline et de glucagon en fonction du glucose, mais il a également un effet positif sur la perte de poids et améliore les fonctions cardiovasculaires, neurologiques et rénales. Les études sur les rongeurs et les humains ont montré que l'administration de GLP-1 induit une augmentation de la satiété, favorisant ainsi la perte de poids en réduisant le bol gastrique et en supprimant la sécrétion d'acides gastriques. Les avantages thérapeutiques du GLP-1 ont conduit à la création d'analogues utilisés initialement pour traiter le DT2, mais leur potentiel pour traiter l'obésité commence à être reconnu<sup>27</sup>.

Parmi ces analogues du GLP-1, le semaglutide s'est particulièrement démarqué, induisant une perte de poids d'au minimum 13,8% sur une période de 52 semaines de traitement chez les adultes et adolescents atteints d'obésité commune. Cette perte de poids est supérieure à celle observée avec le liraglutide qui n'était que de 7,8%<sup>155</sup>. De plus, il semble que ces médicaments soient plus efficaces lorsqu'ils sont associés à des changements de mode de vie, comme des modifications alimentaires et une augmentation de l'activité physique. Dans une étude clinique de phase 3, le liraglutide, en combinaison avec un environnement moins obésogène, a entraîné une perte de poids moyenne de 11,5 % du poids initial sur un an, alors que les seules modifications du mode de vie ont mené à une diminution de 6,1%<sup>14</sup>. En comparaison, le semaglutide, associé à ces changements de mode de vie a montré une diminution de poids de 16% par rapport au poids initial, alors que sans semaglutide, la perte de poids n'était que de 5,7%.<sup>156</sup>.

Enfin, il est envisageable que ces analogues du GLP-1 puissent également être utilisés dans le cadre de l'obésité monogénique. L'effet du liraglutide sur des patients porteurs de variants *MC4R* ou atteints du PWS a été évalué : sur 16 semaines les porteurs de variants dans *MC4R* ont présenté une perte de poids similaire à celle des non-porteurs ( $\pm$  6% de perte de poids), ce qui suggère que le liraglutide peut induire la satiété malgré le dysfonctionnement de la voie leptine-mélanocortine. Les effets du liraglutide sur la perte de poids semblent résulter de son action à plusieurs niveaux *via* les GLP1R. Il agit à la fois au niveau du système nerveux périphérique, en influençant les nerfs innervant le système digestif, et au niveau du SNC, en stimulant directement les neurones POMC et en inhibant de manière indirecte les neurones NPY/AgRP<sup>157,158</sup>. Ces résultats sont encourageants mais demandent davantage d'investigations avant de considérer cette approche comme un traitement pour l'obésité monogénique<sup>84</sup>.



# **MON TRAVAIL DE THESE**

L'obésité, dont la prévalence ne cesse d'augmenter, est une maladie multifactorielle complexe dépendante d'une forte composante génétique. Bien que l'obésité commune soit définie comme une maladie polygénique, l'obésité monogénique est causée par la présence d'un seul variant rare ayant un effet suffisamment fort pour causer l'obésité. Ces formes rares, touchant environ 5% de la population obèse, surviennent très tôt dans la vie du patient et sont généralement plus sévères que les formes communes d'obésité.

Mon travail de thèse s'est porté sur l'investigation de ces variants génétiques rares en utilisant une approche de séquençage de gènes candidats suivi d'études fonctionnelles permettant d'évaluer l'impact des variants identifiés sur la fonction de la protéine associée (autrement appelée « génétique fonctionnelle »), combinée à des analyses statistiques d'association afin de déterminer le rôle de certains gènes candidats dans l'obésité. En outre, l'identification de ces variants rares chez des patients atteints d'obésité peut ouvrir la voie à des progrès dans le domaine de la médecine de précision, en proposant un traitement adapté au patient en fonction de son anomalie génétique causale.

Mon travail s'est divisé en deux parties distinctes, chacune se concentrant sur un gène candidat de l'obésité. La première partie a porté sur l'impact sur l'obésité des variants rares hétérozygotes de *PCSK1*, un gène bien connu pour son implication dans la voie leptine-mélanocortine. Alors que l'effet des mutations bialléliques sur l'obésité monogénique est bien documenté, le véritable impact des mutations hétérozygotes de *PCSK1* n'a pas été clairement établi. La seconde partie s'est concentrée sur les variants rares situés dans le gène *DYRK1B*, qui a été peu étudié dans le cadre des maladies métaboliques et qui est davantage référencé comme étant lié à une obésité « syndromique ». Jusqu'à présent, l'impact des variants rares de *DYRK1B* sur l'obésité a principalement été mis en évidence par une seule étude qui n'a jamais été répliquée à grande échelle. C'est pourquoi nous avons choisi d'examiner l'effet des variants rares situés dans ces deux gènes en particulier sur l'obésité et les traits métaboliques associés dans le cadre de cette étude.

Les variants rares étudiés au cours de ce travail proviennent de l'étude *Rare variants involved in Diabetes and Obesity* (RaDiO), un séquençage NGS des exons de plus de 1200 gènes d'intérêt chez environ 10 000 individus comprenant à la fois des individus atteints d'obésité et/ou de DT2 et des individus contrôles, adultes et enfants. Après l'identification des variants rares dans les deux gènes d'intérêt, l'effet de ces variants a été analysé à l'aide de tests fonctionnels pour déterminer les mutations entraînant une perte ou gain de fonction et dans quelle mesure. Ces variants ont ensuite été classés en différents clusters en fonction de leur

pathogénicité. L'effet de ces clusters a ensuite été évalué par des analyses statistiques d'association afin de déterminer s'ils étaient associés à un risque plus élevé de développer une obésité ou d'autres maladies métaboliques associées.



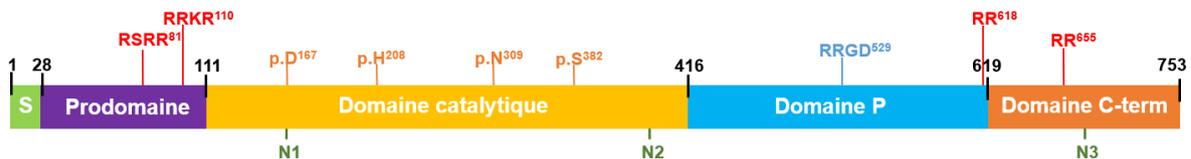
# **PRESENTATION DES TRAVAUX**

# Partie 1: Contribution des variants hétérozygotes de PCSK1 à l'obésité et implications pour la médecine de précision

## 1.1. Contexte de l'étude

### 1.1.1. La biologie de PCSK1

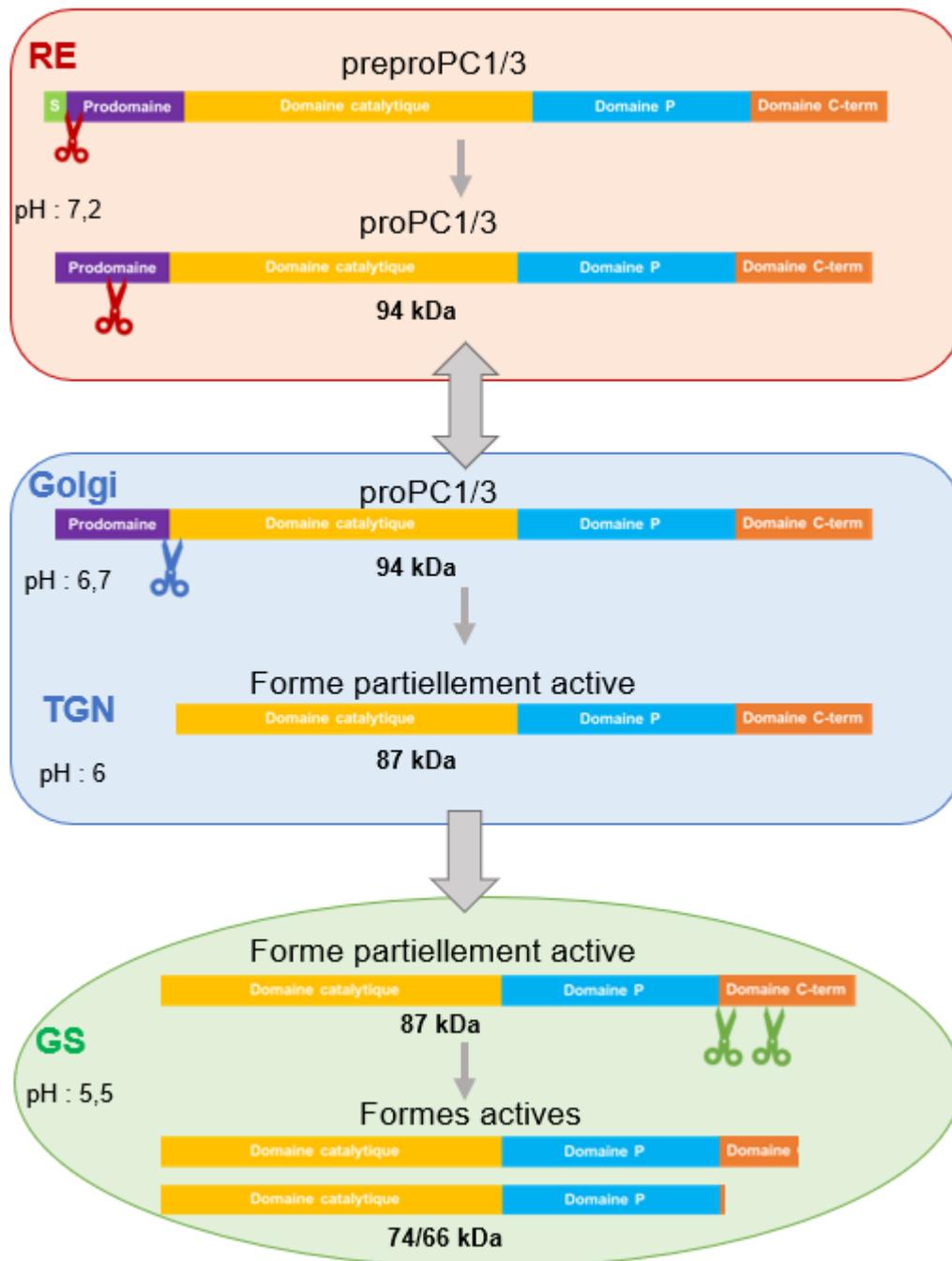
Le gène *PCSK1* appartient à la famille des *proprotein convertase subtilisin/kexin (PCSK)* contenant neuf membres différents qui partagent comme caractéristique commune un domaine catalytique hautement conservé. Cependant, ils diffèrent dans leur distribution tissulaire, leur localisation subcellulaire et leur pH d'activité optimal. Le 1<sup>er</sup> membre de cette famille, *PCSK1*, est situé sur le chromosome 5 et composé de 14 exons codants. Il est exprimé uniquement dans les neurones et les tissus endocrines<sup>93,159</sup>.



**Figure 11 : Les différents domaines de PC1/3.** Les sites de clivage de la maturation protéique sont représentés en rouge. Les acides aminés essentiels à l'activité catalytique sont représentés en orange. Les sites de glycosylation sont représentés en vert. Le motif RRGD qui est essentiel à la maturation et à la sécrétion de PC1/3 est représenté en bleu. S: peptide signal. *Source adaptée de Stijnen, P., Ramos-Molina, B., O'Rahilly, S. & Creemers, J. W. M, « PCSK1 Mutations and Human Endocrinopathies: From Obesity to Gastrointestinal Disorders », Endocrine Reviews, 2016.*

*PCSK1* code la protéine PC1/3 qui est une enzyme majeure, une endoprotéase à sérine dépendante du calcium, impliquée dans la maturation de diverses prohormones et neuropeptides qui régulent l'appétit et l'homéostasie énergétique (cf supra). Notamment, comme expliqué dans l'introduction, PC1/3 clive POMC en ACTH, le précurseur de l'hormone  $\alpha$ -MSH, c'est-à-dire l'agoniste de MC4R dans la voie leptine-mélanocortine. PC1/3 est fortement exprimée dans les tissus neuronaux et endocriniens et comprend un prodomaine, un domaine catalytique, un domaine P et un domaine carboxy-terminal (Figure 11)<sup>93</sup>. Les trois premières régions sont essentielles pour la fonction catalytique de l'enzyme alors que la région C-terminale est importante pour son transport intracellulaire et sa localisation subcellulaire. Le prodomaine, comprenant notamment le motif RRGD, est important pour la bonne conformation de la protéine ainsi que pour sa sortie du réticulum endoplasmique (RE). Le domaine catalytique, responsable de l'activité enzymatique de la protéine, est caractérisé par une triade catalytique composée des acides aminés p.D167, p.H208 et p.S382 ainsi que par un trou oxyanion p.N309. Ce dernier sert de poche catalytique, permettant la stabilisation de l'état de transition lors de la

réaction enzymatique. Enfin, le domaine P est indispensable pour la fixation du calcium ainsi que pour la dépendance de l'enzyme au pH<sup>93,159</sup>.



**Figure 12 : Le processus de maturation de PC1/3.** RE : réticulum endoplasmique, TGN : réseau Trans-Golgi ; GS : granules de sécrétion. Source adaptée de Stijnen, P., Ramos-Molina, B., O’Rahilly, S. & Creemers, J. W. M, « PCSK1 Mutations and Human Endocrinopathies: From Obesity to Gastrointestinal Disorders », *Endocrine Reviews*, 2016.

La maturation de PC1/3, qui est essentielle à son activité enzymatique, nécessite une série de clivages autocatalytiques (Figure 12). Initialement, PC1/3 est synthétisée sous forme d’un

précurseur inactif appelé preproPC1/3 comptant 753 acides aminés. Il est converti en une forme zymogène proPC1/3 (94 kDa) par autoclivage dans le RE qui élimine le peptide signal. Il est ensuite N-glycosylé (au niveau du site de glycosylation N1 chez l'humain) avant que son prodomaine soit clivé dans le RE et le Golgi où PC1/3 est alors convertie en une forme partiellement active (87 kDa). Finalement, PC1/3 est acheminée dans des vésicules de sécrétion où son domaine C-terminal est clivé en réponse à un environnement acide, libérant sa forme mature active de 74/66 kDa. Il est important de noter que cette forme active requiert des concentrations élevées en calcium ainsi qu'un bas pH (5,5) pour exercer son activité maximale. De plus, ces vésicules de sécrétion sont le lieu où PC1/3 va cliver les différents proneuropeptides et prohormones dans leur forme mature<sup>93,159</sup>.

### 1.1.2. La génétique de *PCSK1*

#### 1.1.2.1. Chez la souris

Le premier modèle de souris « *knockout* » *Pcsk1* a été créé en induisant une délétion de 900 pb au niveau du promoteur et de l'exon 1 de *Pcsk1*. Ces souris ont montré une mortalité élevée durant la période prénatale et périnatale, avec seulement 33% de survie au-delà de la première semaine de vie. Sur le plan phénotypique, les souris homozygotes pour cette modification génétique présentaient principalement un retard de croissance (atteignant environ 60% de leur taille normale) ainsi que des altérations dans la maturation des précurseurs d'hormones (incluant la proGHRH, POMC, la proinsuline et le proglucagon) accompagnée d'une hyperproinsulinémie<sup>160</sup>. En ce qui concerne les souris hétérozygotes, elles présentaient une obésité modeste avec un légère intolérance au glucose. Cependant, étant donné que le modèle homozygote ne présentait pas d'obésité, il n'est pas idéal pour l'étude de l'obésité monogénique.

Le principal modèle murin utilisé pour étudier l'impact de la déficience de PC1/3 sur l'obésité est le modèle PC1/3-p.N222D, obtenu par mutagenèse chimique. Cette mutation p.N222D, située dans le domaine catalytique de PC1/3, diminue son activité catalytique de 34 à 50% et provoque une malformation de la protéine, entraînant sa rétention dans le RE et finalement sa dégradation par le protéasome<sup>161,162</sup>. Les souris homozygotes<sup>N222D/N222D</sup>, soumises à un régime riche en graisses, développent une obésité, de l'hyperphagie et des problèmes endocriniens, notamment une réduction des niveaux d' $\alpha$ -MSH dans l'hypothalamus et des anomalies importantes dans la maturation de la proinsuline, menant à une intolérance au glucose sans résistance à l'insuline ou diabète. Les souris hétérozygotes présentent un phénotype intermédiaire et, sous un régime riche en graisses, montrent une augmentation de leur masse

grasse et une hyperproinsulinémie<sup>163</sup>. Un autre modèle murin, le modèle portant la mutation p.V96L/exon3 del, présente également une obésité. Cette mutation altère un acide aminé hautement conservé et affecte un site accepteur d'épissage, entraînant la délétion de l'exon 3. Les souris<sup>V96L/V96L</sup> développent une obésité, de l'hyperphagie, une intolérance au glucose, une résistance à l'insuline, une diarrhée transitoire et une hyperproinsulinémie<sup>164</sup>.

Plus récemment, un nouveau modèle murin pour *Pcskl* a été développé en insérant la mutation p.G209R, identifiée chez un enfant déficient en *PCSK1*, à l'aide de la technologie CRISPR-Cas9. La mutation p.G209R est également située dans le domaine catalytique, résultant en une activité catalytique complètement abolie et une rétention dans le RE. Les souris homozygotes présentent une forte mortalité postnatale et les survivantes avaient un retard de croissance. Bien que l'obésité ne soit pas présente chez ces souris, elles montrent une hypoglycémie à la naissance, des anomalies dans la maturation de l'insuline, une déficience en GH et une réduction de la production d'ACTH. Ces symptômes répliquent en partie ceux observés chez les enfants déficients en *PCSK1*, ce qui fait de ce modèle un outil pertinent pour étudier spécifiquement les anomalies endocriniennes liées à une déficience en *PCSK1*<sup>165</sup>.

#### 1.1.2.2. Chez l'humain

##### a) Les SNP identifiés dans *PCSK1* dans la population générale

Chez l'humain, trois SNP fréquents dans *PCSK1* ont été montrés comme associés à une augmentation du risque d'obésité dans la population générale<sup>58,166</sup> : rs69232 qui résulte en la substitution p.N221D, et rs6234-rs6235, deux SNP qui apparaissent généralement ensemble car ils sont en déséquilibre de liaison, et qui codent les variants p.Q665E/p.S690T. Leur MAF était de 3-5% pour rs69232 et de 24 % pour rs6234-rs6235<sup>58</sup>. Les analyses fonctionnelles ont montré que le SNP rs69232 (p.N221D) diminuait légèrement l'activité enzymatique de PC1/3 (30%) alors que le variant codé par rs6234-6235 (p.Q665E-p.S690T), situé dans le domaine C-terminal, n'a pas montré de diminution de l'activité enzymatique, mais il pourrait avoir un effet sur la sortie et la stabilité de la protéine<sup>167,168</sup>. L'association de ces trois SNP avec l'obésité commune et l'IMC a été confirmée par deux larges méta-analyses indépendantes, avec des modulations importantes liées à l'âge et l'origine ethnique<sup>169,170</sup>. Aucune association n'a été trouvée dans les cohortes d'origine asiatique<sup>169</sup>. En outre, un autre variant rare, le SNP rs1799904, présent dans *PCSK1* et codant la mutation p.R80Q, a été identifié comme ayant un effet modeste sur la fonction de PC1/3. De plus, un SNP intronique rs155971 a été associé à l'obésité dans la population chinoise<sup>171</sup> et le SNP rs261967 a été associé à l'IMC dans une

population asiatique de l'Est<sup>166</sup>, tandis que rs2570467 a montré une légère association avec le tour de taille dans une population d'origine africaine<sup>93,172</sup>.

*b) Les mutations homozygotes rares dans PCSK1 chez des patients avec une obésité extrême*

Le gène *PCSK1* est reconnu comme un des premiers gènes responsables des formes rares d'obésité monogénique chez l'humain. Le premier cas de déficience en *PCSK1* reporté a été décrit en 1995 : il s'agissait d'une patiente de 43 ans présentant une hyperphagie, une obésité précoce (36 kg à 3 ans), des hypoglycémies post-prandiales avec une intolérance au glucose, une hypocortisolémie modérée et un hypogonadisme hypogonadotrope. De plus, elle avait des concentrations plasmatiques élevées de proinsuline et de POMC avec un niveau faible d'insuline, suggérant un défaut dans le processus de maturation des prohormones<sup>94</sup>. En 1997, le séquençage a révélé que cette patiente était en réalité hétérozygote composite pour le gène *PCSK1* (une mutation non synonyme p.G593R et une mutation *splice\_donor* c.620+4A>C)<sup>173</sup>. Un deuxième cas a été reporté en 2003, impliquant un nourrisson de 18 mois hétérozygote composite pour *PCSK1*. Cet enfant souffrait de diarrhées sévères dues à une malabsorption intestinale, une obésité, une altération de la maturation de POMC et de la proinsuline, une hypoglycémie, et d'une hypocortisolémie<sup>174</sup>. Depuis ces premiers cas, d'autres patients ont été identifiés, confirmant que la déficience en *PCSK1* résultait en un phénotype clinique complexe qui inclut une obésité précoce, une malabsorption intestinale sévère, et de nombreux dysfonctionnements endocriniens<sup>175-177</sup>. La plupart des mutations ont été identifiées chez des enfants, souvent issus de familles consanguines, bien que le premier cas ait été identifié chez l'adulte<sup>178</sup>. Aujourd'hui, plus de 34 cas de déficience en *PCSK1* ont été décrits (**Tableau 6**)<sup>176,179-181</sup>, et la plupart des mutations identifiées étaient situées dans le domaine catalytique ou le domaine P de PC1/3<sup>176</sup>.

Le premier symptôme clinique observé chez les patients déficients en *PCSK1* est une malabsorption intestinale sévère, entraînant une diarrhée chronique, une déshydratation, une perte de poids et une acidose métabolique, qui peuvent conduire à une mortalité postnatale précoce. Ces diarrhées néonatales sévères nécessitent une alimentation parentérale pendant les premières années de vie, mais cette malabsorption s'améliore progressivement avec l'âge. Le traitement par GH est également souvent administré aux patients pendant les premières années de vie. Après l'âge de 2 ans, une obésité sévère se développe avec de l'hyperphagie<sup>93</sup>. En outre, les patients peuvent également développer divers troubles endocriniens, notamment une hypoglycémie postprandiale, une insuffisance corticotrope, gonadotrope, thyroïdienne et

somatotrope, un diabète insipide et une polyurie-polydipsie (**Tableau 6**)<sup>176,181</sup>. Un cas présentait également un DT2<sup>182</sup>. Ces endocrinopathies sont probablement liées à des altérations dans le processus de maturation des peptides et hormones qui ont des rôles majeurs dans l'homéostasie métabolique incluant l'insuline, l'hormone de libération de l'hormone de croissance (GHRH), l'hormone de libération des gonadotrophines hypophysaires (GNRH1), la thyrotropine (TRH), l'ACTH et l'hormone antidiurétique (ADH). De même, le phénotype de mauvaise absorption intestinale est probablement lié à l'altération du processus de maturation des prohormones qui favorisent l'assimilation des nutriments (GLP1, CCK et GIP) dans les cellules entéro-endocrines de l'intestin<sup>93</sup>.

**Tableau 6 : Traits phénotypiques observés chez 26 patients déficitaires en PC1/3 qui ont été décrits dans la littérature.** Source adaptée de Pépin, L. et al., « A New Case of PCSK1 Pathogenic Variant With Congenital Proprotein Convertase 1/3 Deficiency and Literature Review », *The Journal of Clinical Endocrinology & Metabolism*, 2019. et de Aerts, L. et al., « Novel Homozygous Inactivating Mutation in the PCSK1 Gene in an Infant with Congenital Malabsorptive Diarrhea », *Genes (Basel)*, 2021.

Trait phénotypique	Pourcentage du trait observé chez les cas déficitaires en PCSK1
<b>Diarrhée liée à une malabsorption</b>	100% (26/26)
<b>Obésité</b>	79,2% (19/24)
<b>Hypoglycémie postprandiale</b>	52,2% (12/23)
<b>Diabète insipide</b>	60,9% (14/23)
<b>Hypocortisolisme central</b>	57,1% (12/21)
<b>Déficiences en GH</b>	44,4% (8/18)
<b>Hypothyroïdisme</b>	56,5% (13/23)
<b>Hypogonadisme</b>	44,4% (8/18)
<b>Polyurie-polydipsie</b>	85% (17/20)

### c) Les mutations hétérozygotes rares dans PCSK1

L'impact potentiel des mutations hétérozygotes dans PCSK1 sur l'obésité n'a pas été clairement défini dans des études à grande échelle. Deux études précédentes ont identifié des variants rares hétérozygotes et nuls (p.R80\* et c.1095 + 1G>A) dans PCSK1 avec un effet fonctionnel délétère sur PC1/3, ce qui a été suggéré comme étant une cause d'obésité monogénique<sup>95,183</sup>. La première étude a identifié une mutation rare hétérozygote non-sens (p.R80\*) dans PCSK1 causant une obésité dans une famille sur plus de trois générations, suivant une transmission autosomique dominante. Des études *in vitro* ont montré que cette mutation inhibe partiellement

l'activité enzymatique de l'allèle *wild-type* (WT) de PC1/3 *in vitro*. Cela suggère que des mutations dans *PCSK1* pourraient avoir un effet dominant négatif sur les fonctions de la protéine mais cette étude était limitée à une seule famille<sup>95</sup>. La seconde a identifié une mutation hétérozygote *splice\_donor* (c.1095 + 1G>A) chez une adolescente de 15 ans souffrant d'obésité sévère, de taux élevés de proinsuline et d'autres anomalies cliniques (syndrome des ovaires polykystiques, *acanthosis nigricans* et hypertension) à partir d'une cohorte de 52 enfants par séquençage Sanger de *PCSK1*. Cette mutation codait une protéine tronquée au niveau de l'exon 8 qui montrait une perte totale de l'activité enzymatique et une rétention au niveau du RE par analyses *in vitro*<sup>183</sup>.

De plus, un article basé sur le séquençage de 845 individus atteints d'obésité a montré que 7 variants rares hétérozygotes de *PCSK1* avaient un effet délétère sur l'activité de PC1/3 selon les analyses *in vitro* et qu'ils contribuaient au développement de l'obésité dans la population générale<sup>96</sup>. Cependant, cette étude, ainsi que plusieurs autres qui ont examiné l'effet fonctionnel des variants de *PCSK1*, a utilisé une construction plasmidique du gène *PCSK1* qui comprenait une mutation non synonyme (codant la mutation p.S357G) qui n'avait pas une régulation sous-cellulaire normale et produisait un traitement anormal des substrats par rapport à la protéine *PCSK1* sauvage, ce qui aurait pu entraîner des biais dans l'interprétation de l'effet des variants sur l'activité de PC1/3<sup>167</sup>.

Comprendre l'impact d'une mutation hétérozygote rare de *PCSK1* sur l'obésité est pourtant crucial pour le diagnostic génétique et les traitements. En effet, la compagnie qui a développé le setmélanotide, *Rhythm Pharmaceuticals*, a lancé la phase 3 d'un essai clinique (NCT05093634) pour évaluer l'effet du setmélanotide sur les patients obèses portant des variants rares hétérozygotes de *PCSK1* sans avoir réalisé au préalable des analyses fonctionnelles *in vitro* des variants identifiés. Dans ce travail, nous avons réalisé une étude génétique fonctionnelle à large échelle en nous concentrant sur les mutations hétérozygotes rares de *PCSK1* afin de déterminer l'impact réel de ces variants sur le risque d'obésité chez l'humain. Ce travail a fait l'objet d'un article publié en mars 2023 dans *Lancet Diabetes & Endocrinology* (**Annexe 3**).

## 1.2. Matériel et Méthode

### 1.2.1. Participants de l'étude

Cette étude cas-témoins a inclus 9320 participants de l'étude RaDiO, qui comprenait les cohortes *Data from the Epidemiological Study on the Insulin Resistance Syndrome* (D.E.S.I.R.) qui comptent des hommes et des femmes d'âge moyen de l'Est de la France, de l'hôpital Corbeil-Essonnes (Corbeil-Essonnes, France), d'une étude CNRS UMR8199 qui inclut des patients recrutés et suivis par l'unité UMR8199 (Lille, France) par le département de Nutrition de l'hôpital Hôtel-Dieu (Paris, France) ou par le Centre d'étude du Polymorphisme Humain (CEPH, hôpital Saint-Louis, Paris, France) et la cohorte d'étude française Fleurbaix-Laventie Ville Santé (**Annexe 1, Tableau S1**)<sup>138</sup>. Les cas comprennent des adultes ou des enfants en surpoids ou atteints d'obésité, tandis que les témoins comprennent des adultes ou des enfants ayant un poids corporel sain. Chez les adultes de plus de 18 ans, un poids corporel sain était défini comme un IMC inférieur à 25 kg/m<sup>2</sup>, le surpoids comme un IMC de 25 à 29,9 kg/m<sup>2</sup> ou plus, et l'obésité comme un IMC de 30 kg/m<sup>2</sup> ou plus. Chez les enfants et les adolescents de moins de 18 ans, un poids corporel sain était défini comme un IMC en fonction de l'âge inférieur au 85ème percentile, et l'obésité comme un IMC en fonction de l'âge supérieur au 95ème percentile ou plus. Le sexe a été déterminé par auto-déclaration. Les participants atteints de DT2 avaient une glycémie à jeun de 7,0 mmol/l ou plus, ou étaient sous traitement pour l'hyperglycémie (ou les deux), et étaient négatifs pour les autoanticorps anti-ilots ou anti-insuline. Les participants du groupe témoin avaient une glycémie à jeun inférieure à 5,6 mmol/l à l'âge de plus de 30 ans et n'utilisaient pas de traitement pour l'hyperglycémie. Tous les participants et les parents (ou tuteurs légaux) ont fourni un consentement écrit éclairé. Les enfants et les adolescents ont donné leur assentiment oral. Les protocoles de l'étude ont été approuvés par les comités d'éthique locaux.

### 1.2.2. Séquençage de *PCSK1*

#### 1.2.2.1. Séquençage NGS

Tous les exons de *PCSK1* (NM\_000439.5) ont été séquencés par NGS. Nous avons utilisé le système de capture *NimbleGen SeqCap EZ Choice XL* (Roche Molecular Diagnosis), qui a été combiné au séquençage Illumina via le système HiSeq 4000. Brièvement, 1 µg d'ADN a été fragmenté par sonication (Covaris E220 Focused-ultrasonicator). Les extrémités des fragments d'échantillons d'ADN ont été liées à des séquences spécifiques adaptatrices (KAPA HTP Library Preparation Kits ; Roche) sur l'automate Hamilton Microlab STARlet. Ces échantillons ont été amplifiés par PCR en utilisant des amorces complémentaires aux adaptateurs. Après la sélection de la taille et la quantification d'échantillon (Perkin Elmer LabChip GX), 24 échantillons ont été combinés dans un seul *pool* d'ADN et ont été hybridés par des sondes

marquées à la biotine SeqCap EZ (Roche). Après 72h à 47°C, les captures ont été purifiées et lavées (kit SeqCap Hybridization ; Roche) à l'aide de l'appareil *Bravo Automated Liquid Handling Platform* (Agilent). Elles ont ensuite été amplifiées (KAPA HiFi HotStart ReadyMix; Roche) et quantifiées (Perkin Elmer LabChip GX et Thermo Fisher Scientific Qubit fluorometric quantitation assays). Enfin, les fragments d'ADN obtenus ont été séquencés par le système Illumina HiSeq 4000, suivant le protocole de séquençage *paired-end* (2x250 bp).

Les données obtenues par le séquençage ont été ensuite démutliplexées en utilisant le programme *bcl2fastq Conversion Software* (Illumina, version 2.17). Les séquences des fragments d'ADN, appelées également *reads*, ont été alignées sur le génome humain (hg19/GRCh37) en utilisant l'aligneur Burrows-Wheeler. Les variants ont été repérés à l'aide de l'outil *Genome Analysis ToolKit* (GATK, version 3.3). Seulement les variants avec une couverture supérieure à 8 *reads* ont été conservés. Les annotations des variants ont été réalisées en utilisant l'interface du programme *the Ensembl Perl* (version 75) et *custom Perls scripts*. Nous avons sélectionné uniquement les variants rares et codants de *PCSK1* avec une MAF inférieure à 1%. Ces variants avaient une couverture de plus de 40 *reads* et un score de qualité supérieure à 150. En outre, aucun variant n'avait plus de 5% de génotype manquant (correspondant à une couverture inférieure à 8 *reads* ou un QUAL score inférieur à 50) chez les participants, et aucun participant n'avait plus de 5 % de génotypes manquants pour *PCSK1*.

#### 1.2.2.2. Séquençage Sanger

Les exons 2,3,8 et 9 de *PCSK1* ont été séquencés pour les analyses de coségrégation par séquençage Sanger dans des familles qui étaient disponibles. Les séquences des amorces ainsi que les conditions PCR utilisées sont détaillées respectivement en **Annexe 1, Tableau S2** et **Tableau S3**. Les exons ont été séquencés dans les deux sens et ont été analysés en utilisant le 3730xl DNA Analyzer (Applied Biosystems, Thermo Fischer Scientific). Les *reads* ont été assemblés et examinés en utilisant les logiciels Variant Reporter et SeqScape (Applied Biosystems, Thermo Fischer Scientific).

#### 1.2.3. Analyses fonctionnelles

##### 1.2.3.1. Mutagenèse dirigée

Des plasmides incluant 61 différentes mutations de *PCSK1* ont été obtenus par mutagenèse (Quik Change II XL Site Directed Mutagenesis Kit ; Agilent) en utilisant des amorces portant la mutation souhaitée (**Annexe 1, Tableau S4**) ainsi qu'un plasmide *PCSK1* WT (Origene). Les séquences des plasmides ont été vérifiées par séquençage Sanger (**Annexe 1, Tableau S5**).

Deux plasmides incluant les variants p.K293R ou p.T366M dans *PCSK1* ont été achetés chez Twist Bioscience.

#### 1.2.3.2. Culture cellulaire

Les expériences *in vitro* de ce projet ont été réalisées sur le modèle cellulaire *human embryonic kidney 293* (HEK293). Les cellules HEK293 ont été cultivées dans du milieu Dulbecco's Modified Eagle's (Thermo Fisher Scientific) contenant 10% de sérum fœtal de veau (SVF) et 1% de pénicilline-streptomycine (Thermo Fisher Scientific). Les cellules ont été incubées avec 5% de CO<sub>2</sub> à 37°C.

#### 1.2.3.3. Tests enzymatiques

Les cellules HEK293 ont été transfectées à l'aide de FuGENE HD (Promega) avec 150 ng/mL de plasmide codant pour le gène de la  $\beta$ -galactosidase et 1000 ng/mL de plasmide contenant le gène *PCSK1* WT ou un des variants identifiés. Trois conditions contrôles ont été utilisées pour chaque manipulation : la condition non transfectée (NT) où seulement le plasmide portant le gène de la  $\beta$ -galactosidase est exprimé ; la condition *wild-type* (WT) exprimant le gène *PCSK1* WT ; et la condition « STOP » où la mutation p.R80\* portant un codon stop est exprimée (comme contrôle négatif). Les cellules ont été cultivées en plaque 24 puits précoatée à la poly-D-lysine (Corning) à une densité de 200 000 cellules/puits. Après 24 heures, le milieu a été remplacé par 300  $\mu$ L d'Opti-MEM (Thermo Fisher Scientific) contenant 100  $\mu$ g/mL d'aprotinine (Sigma-Aldrich). Après 48h de transfection, le surnageant a été récupéré, centrifugé puis congelé à -20°C alors que les cellules ont été lysées dans 250  $\mu$ L de *Luciferase Cell Culture Lysis Buffer* (Promega). Après 30 min sous agitation, 40  $\mu$ L du lysat cellulaire a été récupéré dans une plaque de dosage où ont été ajoutés 100  $\mu$ L de tampon  $\beta$ -galactosidase (0,479 g de sodium phosphate monobasique, 1,068 g de sodium phosphate dibasique dihydrate, 0,0746 g de potassium chloride 2M, 0,00952 g de magnésium chloride 1M complété jusqu'à 100 mL par de l'H<sub>2</sub>O ; du  $\beta$ -mercaptoéthanol (3,75  $\mu$ L/mL) et de l'ortho-nitrophenyl- $\beta$ -galactoside (ONPG) à 4 mM (250  $\mu$ L/mL) sont ajoutés au tampon extemporanément). L'activité de la  $\beta$ -galactosidase est ensuite mesurée à 450 nm par un spectrophotomètre Glomax (Promega). L'activité enzymatique de la protéine a été mesurée par le clivage d'un substrat fluorescent spécifique de PC1/3 en fonction du temps. Pour cela, 25  $\mu$ L de surnageant ont été incubés dans un volume total de 50  $\mu$ L contenant 200  $\mu$ M de substrat peptidique fluorogène pERTKR-AMC (R&D Systems), 100 mM d'acétate de sodium (pH 5,5) (Thermo Fischer Scientific), 2 mM de CaCl<sub>2</sub> (Sigma-Aldrich), 0,1% de Brij 35 (Sigma-Aldrich) et un cocktail d'inhibiteurs de protéases (1  $\mu$ M de pepstatine A, 0,28 mM de N-p-Tosyl-L-lysine chlorométhyl cétone, 10  $\mu$ M

de trans-Epoxy succinyl-L-leucylamido (4-guanidino)butane et 0,14 mM de N $\alpha$ -Tosyl-L-lysine chlorométhyl cétone hydrochloride ; tous provenant de Sigma-Aldrich). Le mélange réactionnel a ensuite été incubé à 37 °C et la fluorescence (lumière d'émission : 380 nm/ lumière d'excitation : 460 nm) est mesurée toutes les 30 minutes pendant 300 minutes à l'aide d'un spectrophotomètre Infinite F200 PRO multimode plate reader (Tecan). Les expériences ont été réalisées en triplicat technique et chaque expérience a été répétée quatre fois.

Les analyses statistiques pour les tests enzymatiques ont été réalisées en utilisant la version 4.1.3 de R. Les mesures ont été normalisées en divisant les mesures d'activité de PC1/3 par les mesures de  $\beta$ -galactosidase. L'ensemble des expériences a ensuite été normalisé en utilisant les mesures de l'activité du NT et de PC1/3<sup>WT</sup>. Le facteur « NT » a été calculé comme la moyenne géométrique des mesures de l'activité du NT. Le facteur « WT » a été calculé comme la moyenne géométrique de l'activité de PC1/3<sup>WT</sup> à 300 minutes moins les mesures du NT. Les mesures normalisées de l'activité des différents variants de PC1/3 (PC1/3<sup>mut</sup>) ont été calculées en soustrayant le facteur NT aux mesures d'activité des variants de PC1/3, puis en divisant le résultat par le facteur WT. Les trajectoires d'activité de PC1/3<sup>mut</sup> normalisées ont été clusterisées par *machine learning* en utilisant le package de R *k-means* pour les données longitudinales (*kml*) avec un nombre de clusters de deux à huit. Le nombre optimal de clusters a été déterminé comme 5 (Groupe\_A, \_B, \_C, \_D et \_E) basé sur les cinq critères de qualité : AIC, BIC, Calinski-Harabatz, Ray-Turi, et Davies-Bouldin.

#### 1.2.3.4. Western blotting

Les cellules HEK293 ont été transfectées à l'aide de FuGENE HD avec 1000 ng/mL de plasmide contenant le gène *PCSK1* WT ou un des variants identifiés du Groupe\_A ou du Groupe\_B (cf infra). Trois conditions contrôles ont été utilisées pour chaque manipulation : la condition NT ; la condition WT exprimant le gène *PCSK1* WT ; et la condition « STOP » où la mutation p.R80\* est exprimée comme contrôle négatif. Les cellules ont été cultivées en plaque 6 puits précoatée à la poly-D-lysine (Corning) à une densité de 800 000 cellules/puits. 48h après la transfection, les cellules ont été récoltées et les protéines ont été extraites en utilisant le tampon Pierce RIPA (Thermo Fisher Scientific) où du dithiothreitol 1 mM (Thermo Fisher Scientific) et des inhibiteurs de protéase (Roche) ont été ajoutés. Les protéines ont été quantifiées en utilisant le kit *Pierce BCA protein assay* (Thermo Fisher Scientific) selon les instructions du fournisseur. 20  $\mu$ g de protéines avec du buffer Laemmli 4 $\times$  (Thermo Fisher Scientific) ont été dénaturées à 95°C pendant 5 minutes. Elles ont ensuite été déposées pour la migration par électrophorèse sur un gel *precast* 10% (Bio-Rad). Après la migration, les

protéines ont été transférées sur une membrane de nitrocellulose 0,2  $\mu\text{m}$  (GE Healthcare). Les sites non-spécifiques sur la membrane ont été bloqués sous agitation pendant 1 h dans du tampon de blocage composé de *Tris buffered saline* (TBS), 0,1% Tween 20, et 5% de lait écrémé. Ensuite, les membranes ont été incubées toute la nuit à 4°C sous agitation avec l'anticorps primaire anti-PC1/3 (ab220363, Abcam ; dilué à 1/1000 dans du tampon de blocage), suivi par un anticorps secondaire anti-lapin fluorescent (SA5-35571, Thermo Fisher Scientific ; dilué à 1/5000 dans du tampon de blocage). Les membranes ont été révélées à l'aide de l'*Odyssey CLx imaging system* (LI-COR Bioscience) à une longueur d'onde de 800 nm. Afin de normaliser les données, la  $\beta$ -actine a été utilisée comme contrôle avec l'anticorps primaire anti- $\beta$ -actine (#3700 Cell Signaling Technology ; dilué à 1/5000 dans le tampon de blocage), suivi par un anticorps secondaire anti-souris fluorescent (#35518 Thermo Fisher Scientific ; dilué à 1/5000 dans du tampon de blocage). La  $\beta$ -actine a également été révélée avec l'*Odyssey CLx imaging system* à une longueur d'onde de 700 nm. Trois expériences indépendantes ont été réalisées pour chaque mutation.

#### 1.2.3.5. Chromatographie liquide-spectrométrie de masse (LC-MS)

Les surnageants provenant de cellules HEK293 transfectées avec un plasmide contenant le gène *PCSK1* WT ou un des variants du Groupe\_A (cf infra) ont été récoltées comme décrit précédemment dans la section « tests enzymatiques ». 50  $\mu\text{L}$  de surnageant ont été préparés suivant une approche protéomique *shotgun bottom-up*. 50  $\mu\text{L}$  de tampon de réduction (dithiothreitol 0,1 M) ont été ajoutés à chaque échantillon et ils ont été incubés à 56°C pendant 40 minutes. Ensuite, les échantillons ont été traités par le protocole de *filter-aided sample preparation* (FASP) en utilisant un filtre Amicon Ultra-0.5 30K (Millipore)<sup>184,185</sup>. Le dessablage des échantillons a été réalisé en utilisant des Evotips-C18 (Evosep) en suivant le manuel d'Evosep. Les Evotips ont ensuite été chargés dans l'appareil de chromatographie liquide Evosep-One (Evosep)<sup>186</sup>. Le système de chromatographie était couplé à un spectromètre de masse Q-Exactive Orbitrap (Thermo Scientific), en utilisant une source nanospray. L'identification protéique a été réalisée en comparant toutes les données MS et MS/MS avec la base de données du protéome Homo Sapiens (Uniprot, version Juin 2021) en utilisant le programme MaxQuant (version 1.6.10.43)<sup>187,188</sup>. La quantification « sans label » (LFQ) a été effectuée en conservant les paramètres par défaut de l'algorithme MaxLFQ. Nous avons utilisé un *false discovery rate* (FDR) de 1% et un minimum de deux peptides par protéine incluant un peptide unique comme paramètres d'identification des protéines et peptides. Les analyses statistiques ont été faites en utilisant le programme MaxQuant Perseus (version 1.6.10.43), avec

trois réplicats par groupe. Pour celle-ci, l'intensité LFQ de chaque échantillon a été téléchargée dans le programme Perseus et la matrice de données a été filtrée en supprimant les contaminants potentiels et les données inversées. Les données ont ensuite été transformées en logarithme et les résultats obtenus ont été normalisés par le Z-score. Les différents groupes ont été comparés à l'aide d'un test ANOVA avec une p-valeur de 0,01.

#### 1.2.3.6. Expériences d'immunofluorescence

Les cellules HEK293 ont été transfectées à l'aide de FuGENE HD avec 1000 ng/mL de plasmide contenant le gène *PCSK1* WT ou un des variants p.R80\*, p.V304I et p.P400S. Les cellules ont été cultivées dans une plaque 6 puits contenant des lamelles recouvertes de poly-D-lysine (Gibco) à une densité de 1 000 000 cellules/puits. 48 heures après la transfection, les cellules sur les lamelles ont été fixées dans une solution de PBS (1X) contenant 4% de paraformaldéhyde (Thermo Fisher Scientific) pendant 20 minutes. Les surfaces de liaison non occupées des lamelles ont été bloquées pendant 1 heure dans un tampon de blocage composé de PBS (1X), 0,1% de triton X-100 et 5% de sérum de chèvre (Sigma-Aldrich). Ensuite, les lamelles ont été incubées toute la nuit à 4°C avec l'anticorps primaire anti-PC1/3 (ab220363, Abcam ; dilué à 1/100 dans du tampon de blocage) suivi d'un anticorps secondaire anti-lapin fluorescent 2 heures à 4°C (A11034, Thermo Fisher ; dilué à 1/500 dans du tampon de blocage). L'actine-F a été marquée avec de la phalloïdine fluorescente (#21836, Thermo Fischer ; dilué à 1/300 dans du tampon de blocage), et les noyaux ont été marqués avec du DAPI (R-37606, Thermo Fisher) pendant 30 minutes à température ambiante. Les lames ont finalement été montées avec le produit de montage *Prolong Gold Antifade Mountant* (Thermo Fischer Scientific). Les cellules ont été observées à l'aide d'un microscope *Cell Observer Spinning Disk* (Zeiss). Pour chaque mutation, trois expériences indépendantes ont été réalisées.

#### 1.2.3.7. Extraction d'ARN, rétrotranscription et qPCR

Les cellules HEK293 ont été transfectées à l'aide de FuGENE HD avec 1000 ng/mL de plasmide contenant le gène *PCSK1* WT ou un des variants p.R80\*, p.V304I et p.P400S. Les cellules ont été cultivées en plaque 6 puits précoatée à la poly-D-lysine à une densité de 800 000 cellules/puits. 48h après la transfection, les cellules ont été récoltées et l'ARN a été extrait en utilisant le kit *Nucleospin RNA* (Macherey-Nagel) selon les instructions du fournisseur. 400 ng de l'ARN extrait ont ensuite été rétrotranscrits en employant le kit *High Capacity cDNA Reverse Transcription* (Thermo Fischer Scientific). Une PCR quantitative (qPCR) a ensuite été réalisée : 2 µL d'ADN complémentaire dilué au 1/8ème ont été déposés en triplicat dans une plaque 384 puits où ont été ajoutés 2,4 µL d'eau, 0,6 µL d'amorces spécifiques pour *PCSK1*

(NM\_000439.5) (Integrated DNA technologies ; **Annexe 1, Tableau S6**) et 5 µL du mix *Brilliant III Ultra-Fast SybrGreen QPCR Master with low ROX* (Agilent Technologies). La qPCR a ensuite été exécutée sur le *Viaa 7 Real-Time PCR System* (Thermo Fisher Scientific) selon les conditions détaillées en **Annexe 1, Tableau S7**. Le gène *RPLP0* a été utilisé comme gène de ménage. Pour chaque mutation, deux expériences indépendantes ont été réalisées.

#### 1.2.3.8. Analyse de l'effet dominant négatif

Les cellules HEK293 ont été transfectées à l'aide de FuGENE HD avec trois plasmides : 150 ng/mL de plasmide contenant le gène codant la β-galactosidase, 500 ng/mL de plasmide contenant le gène *PCSK1* WT et 500 ng/mL de plasmide contenant un des variants non synonyme du Groupe\_A (cf infra). Trois conditions contrôles ont été utilisées par manipulation : la condition NT où seulement le plasmide portant le gène de la β-galactosidase est exprimé ; la condition WT exprimant le gène *PCSK1* WT ; et la condition où est exprimée seulement le variant du Groupe\_A. La suite de la manipulation a été réalisée comme précédemment décrit dans la section « tests enzymatiques ». Les expériences ont été réalisées en triplicat technique et chaque expérience a été répétée quatre fois.

#### 1.2.4. Analyses statistiques pour les études d'association

Les variants rares de *PCSK1* ont été divisés en quatre clusters de pathogénicité (I = Groupe\_A + variants nuls ; II = Groupe\_B ; III = Groupe\_C ; et IV = Groupe\_D + Groupe\_E) basés sur les résultats des tests enzymatiques. Chaque cluster a été analysé en utilisant la méthode *mixed-effects score test* (MiST)<sup>189</sup>. L'analyse MiST donne un score statistique  $S(\pi)$  pour l'effet moyen ( $\pi$ ) du cluster, et un score statistique  $S(\tau)$  pour l'effet hétérogène ( $\tau$ ) du cluster. L'équation du modèle est la suivante :

$$Y = \alpha X + \pi GZ, \text{ où}$$

Y est le trait d'intérêt ; X est la matrice des covariables (comprenant l'âge, le sexe, l'IMC, et l'ethnicité pour évaluer le DT2 et les concentrations lipidiques ; et l'âge, le sexe et l'ethnicité pour évaluer l'adiposité) ; G est la matrice des variants *PCSK1* ; Z est le vecteur de la matrice des variants *PCSK1* répétés  $n$  fois avec  $n$  correspondant au nombre de variants rares de *PCSK1* trouvés chez les participants, menant à l'équation :  $\pi GZ = \pi \sum_{i=1}^n G_i$ .

L'origine ethnique des participants a été évaluée en utilisant les cinq composants génotypiques principaux (de PC1 à PC5)<sup>138</sup>. Les composantes PC1 à PC5 ont été calculées à partir des 15 020 variants fréquents disponibles dans RaDiO en commun avec le projet « 1000 Genomes » qui

inclut des patients de multiples origines ethniques. Comme aucune des études d'association entre les variants de *PCSK1* et l'obésité, le surpoids, l'IMC, le DT2 ou les lipides n'avait une hétérogénéité significative, nous avons seulement montré les p-valeurs associées avec le score  $S(\pi)$  représentant l'effet moyen du cluster. Ces analyses statistiques ont été réalisées en utilisant le logiciel R (version 4.0.2).

### 1.3. Résultats

#### 1.3.1. Détection de 65 variants rares codant de *PCSK1* dans RaDiO

Afin d'identifier les variants rares du gène *PCSK1*, nous avons réalisé un séquençage NGS de ce gène sur 9320 participants incluant 7260 adultes et 2060 enfants provenant de l'étude RaDiO. Parmi ces 9320 participants, 2568 étaient atteints d'obésité. Grâce à ce séquençage, nous avons identifié 65 mutations rares (MAF < 1%) du gène *PCSK1* (NM\_000439.5) qui étaient toutes hétérozygotes (**Tableau 7**). Parmi celles-ci, deux sont déjà identifiées comme pathogènes d'après une définition purement génétique : une mutation *frameshift* p.S340Pfs\*94, et une mutation *splice donor* c.1095+1G>A. Ces variants nuls sont difficilement modélisables dans les plasmides par mutagenèse et sont par ailleurs considérés perte-de-fonction d'après les règles génétiques. Nous avons donc réalisé les études *in vitro* sur les 63 mutations identifiées restantes (2 mutations non-sens et 61 mutations non synonymes).

**Tableau 7 : Variants rares de *PCSK1* détectés à partir de l'étude RaDiO.** *MAC* : minor allele count ; *gnomAD* : genome aggregation database (version 2.1.) ; *NA* : non disponible ; *REVEL* : rare exome variant ensemble learner ; Les références sont celles dans lesquelles des mutations de *PCSK1* ont été reportées comme pathogène, selon la *Human Gene Mutation Database Professional* (version 2022.3). Le seuil de pathogénicité pour REVEL était de 50%. \*, selon nos analyses *in vitro*.

Mutation dans <i>PCSK1</i> (NM_000439.5)	MAC dans gnomAD	MAC dans RaDiO	Références	Groupe fonctionnel*	REVEL
c.34G>C, p.A12P	0	1	/	D	Neutre
c.173T>C, p.L58S	0	1	/	B	Neutre
c.209T>G, p.F70C	0	1	/	A	Neutre
c.238C>T, p.R80*	2	1	Philippe et al (2015) <sup>95</sup> ; Qian et al (2021) <sup>180</sup>	A et Nul	NA
c.239G>A, p.R80Q	585	1	/	B	Neutre
c.243G>C, p.R81S	1	1	/	A	Neutre
c.254A>G, p.H85R	0	1	/	C	Neutre
c.281A>G, p.D94G	1	2	/	D	Neutre
c.284G>A, p.R95H	3	1	/	A	Neutre
c.334G>A, p.A112T	0	2	/	D	Neutre
c.375G>A, p.M125I	51	13	Creemers et al (2012) <sup>96</sup>	B	Neutre
c.395T>C, p.L132S	0	1	/	A	Délétère
c.470C>T, p.T157M	5	1	/	B	Délétère
c.524C>T, p.T175M	41	2	Creemers et al (2012) <sup>96</sup>	A	Délétère
c.541T>C, p.Y181H	29	4	Creemers et al (2012) <sup>96</sup>	B	Délétère
c.559T>C, p.Y187H	4	2	/	D	Délétère

c.577G>T, p.D193Y	0	1	/	A	Délétère
c.608C>T, p.T203I	0	1	/	C	Délétère
c.646A>G, p.I216V	1	1	/	C	Neutre
c.650C>T, p.A217V	15	3	/	C	Délétère
c.844C>T, p.R282W	7	1	/	D	Délétère
c.869A>G, p.Y290C	16	1	/	C	Délétère
c.878A>G, p.K293R	2	1	/	B	Neutre
c.910G>A, p.V304I	35	3	/	A	Délétère
c.934C>T, p.R312C	1	1	/	B	Délétère
c.970G>T, p.D324Y	0	1	/	A	Délétère
c.974G>A, p.S325N	0	1	Creemers et al (2012) <sup>96</sup>	C	Délétère
c.985A>G, p.I329V	1	1	/	C	Délétère
c.987C>G, p.I329M	0	2	/	C	Délétère
c.1018del, p.S340Pfs*94	0	1	/	Nul	NA
c.1019C>A, p.S340Y	0	1	/	B	Délétère
c.1061C>G, p.S354C	2	1	/	D	Délétère
c.1087C>T, p.Q363*	2	1	/	A et Nul	NA
c.1095+1G>A, p.?	0	1	Löffler et al (2016) <sup>183</sup> ; Martin et al (2013) <sup>178</sup>	Nul	NA
c.1097C>T, p.T366M	9	1	/	B	Neutre
c.1123A>G, p.T375A	0	1	/	A	Délétère
c.1130C>T, p.T377M	4	1	/	E	Délétère
c.1168G>A, p.G390S	2	2	/	A	Délétère
c.1193C>G, p.A398G	0	1	/	C	Délétère
c.1198C>T, p.P400S	0	1	/	A	Délétère
c.1283A>C, p.K428T	0	1	/	D	Neutre
c.1346T>C, p.L449P	0	1	/	A	Délétère
c.1381G>A, p.V461M	7	1	/	D	Neutre
c.1405G>A, p.V469I	6	2	/	E	Neutre
c.1406T>C, p.V469A	1	1	/	C	Neutre
c.1441G>A, p.A481T	0	1	/	C	Neutre
c.1460T>C, p.I487T	6	1	/	A	Neutre
c.1472C>T, p.T491I	0	1	/	D	Neutre
c.1582G>A, p.A528T	0	2	/	B	Neutre
c.1585G>A, p.A529T	0	1	/	C	Neutre
c.1621C>T, p.R541W	11	1	/	C	Neutre
c.1622G>A, p.R541Q	1	1	/	C	Neutre
c.1669C>T, p.H557Y	0	1	/	A	Délétère
c.1745G>C, p.G582A	0	1	/	A	Délétère
c.1813C>T, p.R605C	0	1	/	B	Neutre
c.1823C>T, p.T608M	7	1	/	D	Neutre
c.1855G>A, p.G619R	0	3	/	D	Neutre
c.1918A>G, p.T640A	368	32	/	D	Neutre
c.1961G>A, p.R654Q	8	1	/	D	Neutre
c.1978G>A, p.E660K	2	1	/	D	Neutre
c.2076G>C, p.K692N	126	1	/	D	Neutre
c.2129C>G, p.P710R	1	1	/	D	Neutre
c.2180T>A, p.V727D	0	1	/	D	Neutre
c.2219G>A, p.R740Q	7	1	/	C	Neutre
c.2236G>A, p.V746M	10	3	/	C	Neutre

### 1.3.2. Etudes fonctionnelles *in vitro* des variants rares hétérozygotes du gène *PCSK1*

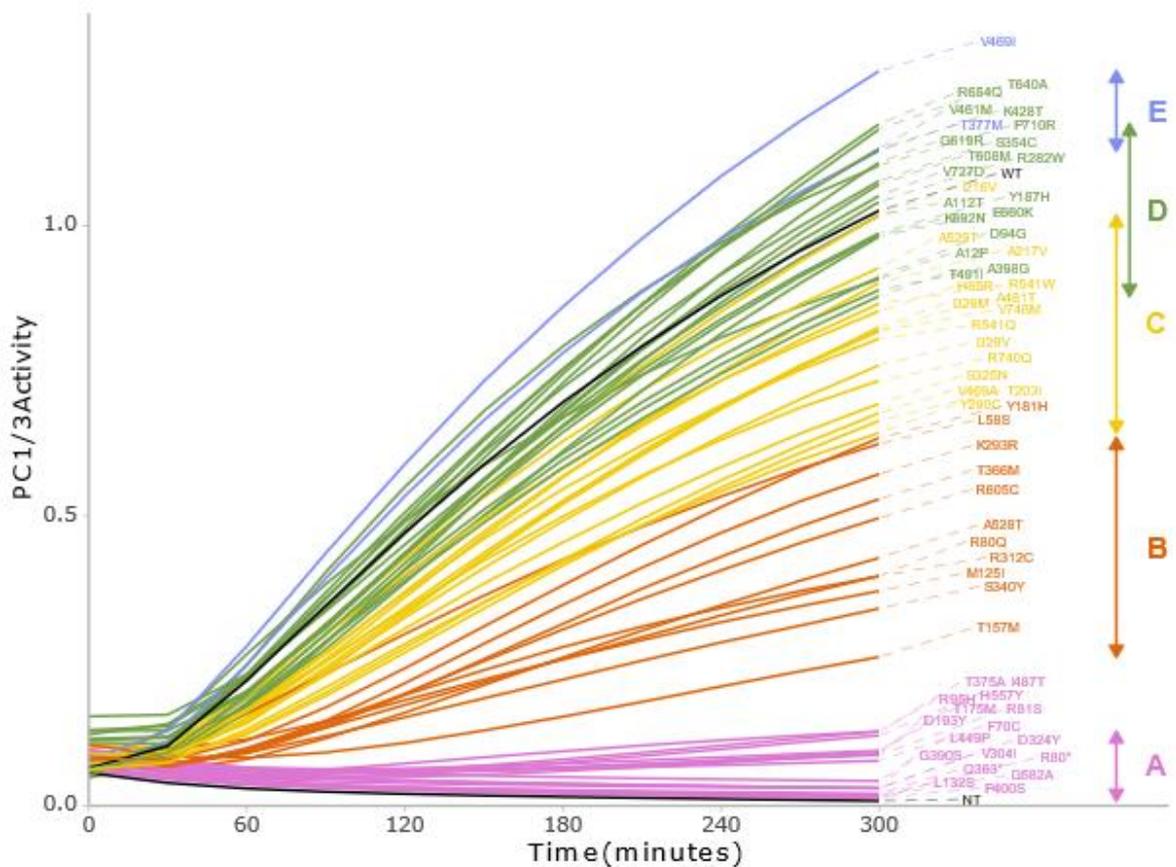
#### 1.3.2.1. Analyse de l'activité enzymatique et classification des variants par clustering

Chaque variant rare a été exprimé dans des cellules HEK293 et l'activité enzymatique a été mesurée à l'aide d'un substrat fluorescent de PC1/3 pendant 300 minutes. En comparant l'activité du WT avec celle des différents mutants, nous avons trouvé que la plupart des variants diminuait significativement l'activité enzymatique de PC1/3 mais à différents degrés (**Annexe 2, Figure S1**). Dans tous les résultats obtenus, nous n'avons observé aucune mutation gain de fonction, c'est-à-dire améliorant l'activité enzymatique par rapport à la condition WT.

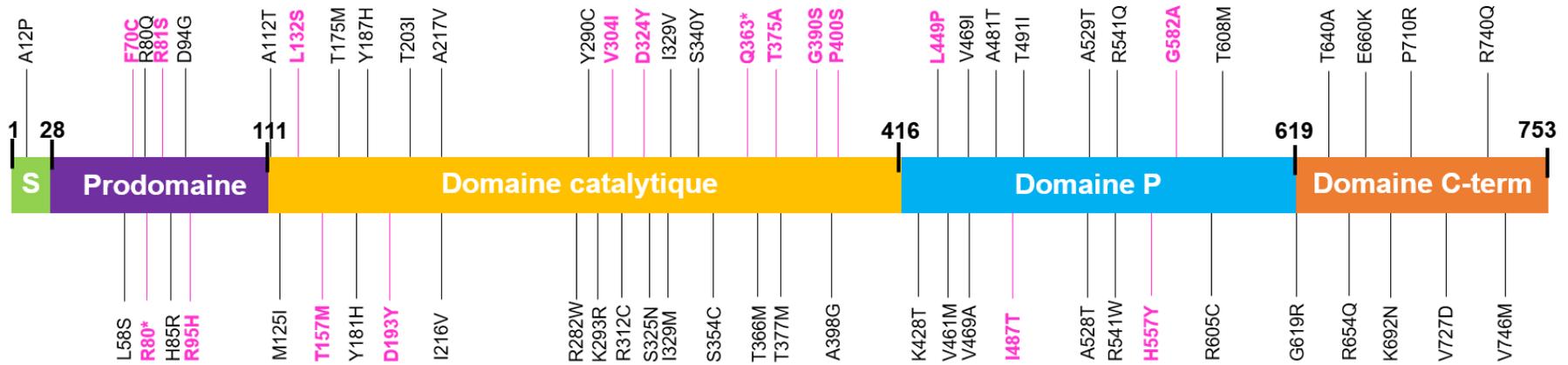
Nous avons ensuite réalisé une analyse de clusterisation par *machine learning*, nous permettant de classer les variants en cinq clusters selon leur impact sur l'activité de PC1/3 (*i.e.* A à E) (**Figure 13**). Cette technique a permis de classer les différents variants de façon objective en fonction de leur profil d'activité enzymatique. En comparaison avec la condition WT, nous avons trouvé que 17 variants menaient à une perte d'activité totale de PC1/3 (Groupe\_A), 11 variants montraient une activité enzymatique partielle (Groupe\_B), 16 variants montraient une activité très légèrement diminuée par rapport à la condition WT (Groupe\_C) et 19 variants n'affectaient pas l'activité enzymatique de PC1/3 (Groupe\_D et Groupe\_E) (**Figure 13**).

#### 1.3.2.2. Localisation des variants sur la protéine PC1/3

En regardant la localisation des variants rares identifiés dans cette étude au niveau des domaines de PC1/3, nous distinguons que la majorité des mutations du Groupe\_A se situe au niveau du domaine catalytique (neuf variants). Nous comptons également quatre mutations au niveau du prodomaine, quatre dans le domaine P mais aucune dans le domaine C-terminal de PC1/3 (**Figure 14**).



**Figure 13 : Analyse en clusters selon l'activité enzymatique des 63 variants non synonymes et non-sens de *PCSK1*.** La figure montre le clustering des trajectoires de l'activité de PC1/3 des variants de *PCSK1* par *kml* ( $k = 5$ ). L'axe des Y représente l'activité catalytique de PC1/3 normalisée par l'activité de la  $\beta$ -galactosidase. L'activité enzymatique a été mesurée pendant 300 minutes. Cinq groupes ont été déterminés selon leur activité enzymatique par *machine learning*. Les variants roses sont inclus dans le Groupe\_A ; les variants oranges sont inclus dans le Groupe\_B ; les variants jaunes sont inclus dans le Groupe\_C ; les variants verts sont inclus dans le Groupe\_D et les variants bleus dans le Groupe\_E. Les conditions NT et WT sont représentées en noir. Les données sont la moyenne de quatre expériences indépendantes réalisées en triplicat technique. *NT*, non-transfectée ; *WT*, wild-type



**Figure 14 : Localisation des variants rares non synonymes et non-sens de *PCSKI*.** Les différents domaines de la protéine PC1/3 sont représentés. Les variants en rose sont liés à une perte complète de l'activité enzymatique (Groupe\_A) selon les résultats des tests *in vitro*. Les variants en noir sont liés à un effet partiel ou neutre sur l'activité de PC1/3 (Groupe\_B, Groupe\_C, Groupe\_D et Groupe\_E) selon les résultats des tests *in vitro*. **S**, peptide signal ; **C-term**, domaine C-terminal.

### 1.3.2.3. Analyse de l'expression protéique des variants du Groupe\_A

Les western blots réalisés sur les lysats cellulaires de cellules HEK293 exprimant les différents variants du Groupe\_A ont révélé que l'expression protéique était altérée chez la plupart de ces variants (**Tableau 8** et **Annexe 2, Figure S2**). En effet, parmi les 17 mutations du Groupe\_A, seulement trois mutations (p.R81S, p.R95H et p.T375A) montraient à la fois les formes précurseurs (94 kDa et 87 kDa) et la forme active de PC1/3 (74/66 kDa) de façon similaire à la condition WT. Ces résultats suggèrent que ces trois variants n'affectent pas la conformation de la protéine mais plutôt l'activité enzymatique (**Tableau 8**).

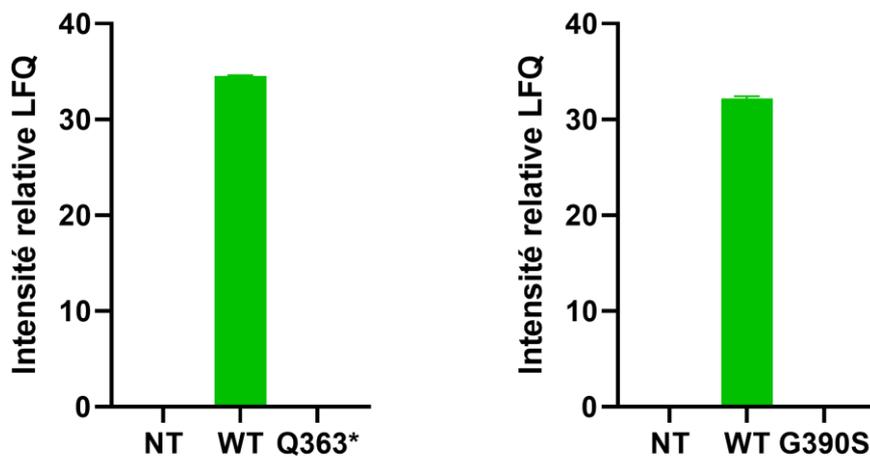
Concernant les autres variants, six mutations (p.L132S, p.D193Y, p.D324Y, p.I487T, p.H557Y et p.G582A) exprimaient les deux formes précurseurs mais n'avaient pas ou peu d'expression de la forme active de PC1/3, indiquant une altération de la sécrétion de la forme active chez ces variants (**Tableau 8**). De plus, trois mutations (p.F70C, p.T175M et p.L449P) présentaient uniquement la bande correspondant à la forme zymogène de 94 kDa, témoignant un problème dans la maturation de PC1/3 au niveau du RE (**Tableau 8**). Les analyses LC-MS réalisées sur le surnageant de cellules HEK293 exprimant les variants du Groupe\_A ont montré une sous-expression de la protéine PC1/3 dans la plupart des surnageants cellulaires par rapport à la condition WT (**Annexe 2, Figure S3**). Cette sous-expression par rapport au WT peut expliquer l'absence de détection de la forme active dans les lysats cellulaires sur les membranes des western blots, qui sont moins sensibles que l'analyse LC-MS. Cependant, il n'a pas été possible de déterminer si les protéines PC1/3 présentes dans les surnageants étaient des formes actives sécrétées ou des formes précurseurs présentes en raison de la mort cellulaire.

Trois mutations (p.R80\*, p.V304I et p.P400S) ne présentaient aucune forme de PC1/3 dans les lysats cellulaires (**Tableau 8**). Cette observation est cohérente pour le variant p.R80\* qui induit un codon stop au début de la protéine, expliquant l'absence de détection de PC1/3 dans les lysats. La surexpression génique des variants p.V304I et p.P400S dans les cellules HEK293 a été confirmée par la méthode de RT-qPCR (**Annexe 2, Figure S4**). De plus, les analyses LC-MS ont détecté la protéine PC1/3 dans le surnageant des cellules surexprimant ces deux variants, bien qu'elle ait été sous-exprimée par rapport à la condition WT (**Annexe 2, Figure S3**). L'expression intracellulaire a également été analysée par immunofluorescence (**Annexe 2, Figure S5**). La fluorescence correspondant à la protéine PC1/3 était observée dans les cellules surexprimant le variant p.P400S, mais elle était diminuée, tout comme chez le variant p.R80\* par rapport à la condition WT. De plus, alors que la fluorescence correspondant à la protéine était observée dans tout le cytoplasme cellulaire de la condition WT, seuls quelques

points lumineux étaient visibles dans les cellules surexprimant le variant p.V304I. Ces résultats suggèrent que la protéine mutée codée par p.V304I est ségrégée à certains endroits dans la cellule, probablement au niveau du RE afin d'être dégradée par le protéasome (**Annexe 2, Figure S5**).

De plus, deux variants du Groupe\_A, codant respectivement p.Q363\* et p.G390S, ne montraient ni de formes précurseurs ni aucune forme active de PC1/3, mais une seule bande d'environ 50 kDa (**Tableau 8**). Les analyses en LC-MS ont confirmé l'absence de la protéine PC1/3 dans le surnageant des cellules HEK293 exprimant ces mutations, ce qui suggère un clivage incorrect des formes précurseurs de PC1/3 (**Figure 15**).

En revanche, les western blots réalisés sur les lysats cellulaires exprimant les variants du Groupe\_B (ayant une activité enzymatique partielle) ont montré qu'un grand nombre de ces variants présentait toutes les formes de PC1/3, de façon similaire au WT (**Annexe 1, Tableau S8 et Annexe 2, Figure S2**). Il convient de noter que la bande correspondant à la forme active de la protéine était absente pour les variants codant p.T157M, p.Y181H et p.R312C, mais toutes les formes zymogènes étaient présentes.



**Figure 15 : Expression protéique des variants p.Q363\* et p.G390S de PC1/3 dans le surnageant par analyse LC-MS.** Expression protéique de PC1/3 dans les surnageants de cellules HEK293 transfectées ou non (condition non-transfectée [NT]) avec des plasmides contenant *PCSK1* WT ou le variant p.Q363\* ou le variant p.G390S. La protéine PC1/3 était considérée comme exprimée si elle était présente dans au moins deux expériences sur trois. Les données ont été transformées en log avant l'analyse. *LFQ*, label-free quantification ; *NT*, non-transfecté ; *WT*, wild-type

**Tableau 8 : Expression protéique des variants perte de fonction complète de *PCSK1* (Groupe\_A) dans les lysats cellulaires par western blotting** Ce tableau récapitule l'expression protéique des variants de *PCSK1* du Groupe\_A par western blotting. Taille des formes précurseurs de PC1/3 : 94kDa et 87kDa ; taille de la forme active de PC1/3 : 74/66 kDa. Trois expériences indépendantes ont été réalisées pour chaque mutation (les images des western blotting sont disponibles en **Annexe 2, Figure S2**). *Nnwt*, nombre de participants porteurs de la mutation avec un poids normal dans l'étude RaDiO ; *Nob*, nombre de participants porteurs de la mutation atteints d'obésité dans l'étude RaDiO ; *Nowt*, nombre de participants porteurs de la mutation atteints de surpoids dans l'étude RaDiO.

		Variants de <i>PCSK1</i> perte de fonction complète (Groupe_A)																	
Variants	WT	F70C	R80*	R81S	R95H	L132S	T175M	D193Y	V304I	D324Y	Q363*	T375A	G390S	P400S	L449P	I487T	H557Y	G582A	
<i>N<sub>ob</sub> / N<sub>owt</sub> / N<sub>nwt</sub></i>	-	0/1/0	1/0/0	1/0/0	1/0/0	1/0/0	0/2/0	0/0/1	1/1/1	1/0/0	0/1/0	1/0/0	1/1/0	1/0/0	1/0/0	1/0/0	0/1/0	0/1/0	
Formes précurseurs	Oui	Oui	Non	Oui	Oui	Oui	Oui	Oui	Non	Oui	Non	Oui	Non	Non	Oui	Oui	Oui	Oui	
Forme active	Oui	Non	Non	Oui	Oui	Non	Non	Non	Non	Non	Non	Oui	Non	Non	Non	Non	Non	Non	
Exemples																			
kD																			
100 —																			
70 —																			
55 —																			

#### 1.3.2.4. Analyse de l'effet dominant négatif sur l'activité enzymatique

L'effet dominant négatif se produit lorsqu'une protéine mutée avec une perte de fonction altère la protéine sauvage codée par l'allèle WT, entraînant une perte de la fonction totale de la protéine chez les porteurs hétérozygotes<sup>190</sup>. Plusieurs études<sup>95,161,167</sup> ont mis en évidence que les variants hétérozygotes de *PCSK1* pouvaient avoir un effet dominant négatif sur l'allèle WT, notamment parce les variants perte de fonction de *PCSK1* peuvent être retenus au niveau du RE en raison de leur mauvaise conformation, ce qui peut mener à une oligomérisation de la protéine mutée avec la protéine sauvage, empêchant son fonctionnement. C'est pourquoi l'effet dominant négatif des variants non synonymes du Groupe\_A sur l'activité enzymatique de la protéine WT a été évalué. Dans ce but, des cellules HEK293 ont été co-transfectées avec un plasmide WT et soit avec un plasmide exprimant une mutation non synonyme du Groupe\_A ou soit avec un plasmide exprimant un variant rare qui ne diminuait pas l'activité enzymatique (p.I216V [Groupe\_C] et p.T608M [Groupe\_D]) (**Annexe 2, Figures S7 et S8**). L'activité enzymatique était similaire quand la condition WT était seule ou coexprimée avec les variants contrôles p.I216V et p.T608M (**Annexe 2, Figure S7**). En revanche, l'activité enzymatique du WT était diminuée de manière plus ou moins importante dès qu'un variant non synonyme du Groupe\_A était exprimé avec la condition WT (**Annexe 2, Figure S8**). Cette diminution de l'activité enzymatique du WT était plus importante lorsque le WT était coexprimé avec les variants p.T175M ou p.L449P qui présentaient uniquement la forme zymogène dans les lysats cellulaires (**Tableau 8**) ou avec le variant p.V304I qui ne présentait aucune forme protéique en western blot et dont la protéine mutée semblait s'accumuler dans le RE (**Annexe 2, Figure S5**). Ces résultats suggèrent que les variants perte de fonction totale du Groupe\_A pourraient exercer un effet dominant négatif sur l'allèle WT, empêchant le bon fonctionnement de la protéine sauvage.

#### 1.3.3. Effet des variants rares de *PCSK1* sur différents traits métaboliques selon leur pathogénicité

Nous avons analysé l'effet de tous les variants rares de *PCSK1* identifiés dans notre étude sur l'homéostasie métabolique. En utilisant la méthode MiST ajustée au sexe, à l'âge et l'origine ethnique, l'agrégation de tous les variants rares était significativement associée avec un risque plus élevé de développer un surpoids ou une obésité ( $P = 0,018$  avec un odds ratio [OR] de 1,7 ; **Tableau 9**) et avec un IMC plus élevé ( $28 \pm 5,6$  kg/m<sup>2</sup> chez les porteurs *versus*  $27 \pm 6,5$  kg/m<sup>2</sup> chez les non-porteurs ;  $P = 0,039$  avec un effet [ $\pi$ ] de  $1,6 \pm 0,79$  ; **Tableau 9**). En revanche, en

utilisant la méthode MiST ajustée au sexe, à l'âge, à l'IMC et l'origine ethnique, l'agrégation de tous ces variants n'était pas significativement associée avec le risque de développer un DT2 ou d'avoir un plus haut niveau de cholestérol (**Tableau 9**).

Nous avons ensuite effectué des analyses similaires pour chaque cluster de pathogénicité déterminé par nos analyses *in vitro* (**Figure 13**). L'agrégation des variants perte de fonction totale et des variants nuls (*i.e.* le Cluster\_I regroupant les variants du Groupe\_A et les deux variants nuls c.1018del et c.1095+1G>A) montrait un effet significatif sur l'augmentation de l'IMC ( $P = 3,0 \times 10^{-4}$  avec  $\pi$  de  $6,9 \pm 1,9$  ; **Tableau 9**). De la même façon, ces variants du Cluster\_I étaient significativement retrouvés davantage chez les individus atteints d'obésité seule (86% des individus atteints d'obésité parmi les porteurs *versus* 35% des individus atteints d'obésité parmi les non-porteurs ;  $P = 0,014$  avec OR de 9,3 ; **Tableau 9**) et chez les individus atteints de surpoids ou d'obésité (94% parmi les porteurs *versus* 60% parmi les non-porteurs ;  $P = 0,013$  avec OR de 8,5 ; **Tableau 9**). Nous avons trouvé la même tendance d'association avec l'obésité chez les enfants bien que la puissance statistique était limitée (5 enfants atteints d'obésité sur 6 [83%] parmi les porteurs *versus* 50% des enfants atteints d'obésité parmi les non-porteurs ;  $P = 0,18$  avec OR de 4,7 ; **Tableau 10**). En revanche, les autres clusters composés des variants de *PCSK1* ayant un effet partiel ou neutre sur l'activité de PC1/3 n'avaient pas d'effet significatif sur le risque d'obésité ou de surpoids, sur l'IMC ou sur d'autres traits métaboliques (**Tableau 9**).

Finalement, nous avons analysé la coségrégation des variants du Cluster\_I (p.R81S, p.L132S, p.V304I, p.Q363\*, c.1095+1 G>A et p.T375A) avec l'obésité dans six familles dont les différents cas index provenaient de RaDiO (**Annexe 1, Tableau S9**). Dans la première famille, le cas index portant la mutation p.R81S, qui présentait une obésité sévère (IMC de  $46 \text{ kg/m}^2$ ), a transmis la mutation à trois de ses quatre enfants qui présentaient tous une obésité. Le dernier était également atteint d'obésité mais il ne portait pas la mutation. Dans la deuxième famille, le cas index était un garçon de 12 ans qui souffrait d'obésité et était porteur de la mutation p.L132S. Il a hérité cette mutation de sa mère qui ne souffrait pas d'obésité. De la même façon, dans la troisième famille, les deux cas index porteurs de la mutation p.V304I étaient un homme de 52 ans atteint d'obésité (IMC de  $38 \text{ kg/m}^2$ ) et son fils de 14 ans qui, bien qu'il soit porteur de la mutation, ne souffrait pas d'obésité. La quatrième famille était composée de deux frères : un homme de 51 ans, qui était le cas index, porteur de la mutation p.Q363\* et qui souffrait de surpoids (IMC de  $26 \text{ kg/m}^2$ ) et son frère qui était également porteur de la mutation p.Q363\* et qui présentait une obésité. La cinquième famille avait des limites plus importantes car le variant

p.1095+1 G>A était présent uniquement chez le cas index atteint de surpoids (IMC de 28 kg/m<sup>2</sup>). Enfin, dans la sixième famille, le cas index était une fille de 14 ans qui souffrait d'obésité et qui portait la mutation p.T375A héritée de sa mère qui avait un poids normal (**Annexe 1, Tableau S9**). Ces résultats nous ont montré que, bien qu'il y ait un enrichissement d'individus présentant une obésité au sein des familles porteuses des variants du Cluster\_1, la pénétrance des variants hétérozygotes qui ont une perte de fonction complète n'était pas totale car des individus porteurs de ces variants dans plusieurs de ces familles avaient un poids normal (les deuxième, troisième et sixième familles).

#### 1.3.4. Evaluation de l'utilité des analyses *in silico* pour discriminer les variants complètement perte de fonction des autres variants de *PCSK1*

Afin d'évaluer l'utilité des tests *in silico* pour discriminer les variants perte de fonction totale (Groupe\_A) des autres variants, nous avons réalisé une analyse de prédiction à l'aide de REVEL (**Tableau 7**). REVEL est un ensemble de méthodes (incluant 13 logiciels) utilisé pour prédire la pathogénicité des variants rares non synonymes<sup>191</sup>. Quatre mutations perte de fonction complète (codant pour p.F70C, p.R81S, p.R95H, p.I487T) sur les 61 variants n'ont pas été prédites comme pathogènes par REVEL et étaient prédites comme neutres (7% ; faux négatifs). En revanche, 15 variants (codant pour p.T157M, p.Y181H, p.Y187H, p.T203I, p.A217V, p.R282W, p.Y290C, p.R312C, p.S325N, p.I329V, p.I329M, p.S340Y, p.S354C, p.T377M, p.A398G) qui faisaient partie du Groupe\_B, Groupe\_C, Groupe\_D et même du Groupe\_E ont été détectés comme délétères par REVEL (25% ; faux positifs) (**Tableau 7**). Par conséquent, les analyses *in silico* ne sont pas assez sensibles et spécifiques pour prédire précisément la pathogénicité des variants hétérozygotes de *PCSK1*. La répartition des porteurs d'un variant non synonyme de *PCSK1* en fonction du sexe, de l'état d'obésité, de l'activité enzymatique des variants non synonymes (Groupe\_A *versus* Groupe\_B) et de la prédiction REVEL des variants non synonymes (délétère *versus* neutre) montre une augmentation du nombre de porteurs atteints d'obésité dans le Groupe\_A, indépendamment de la prédiction de REVEL (**Figure 16**). Ces résultats montrent l'importance des analyses *in vitro* pour discriminer véritablement les variants rares non synonymes qui ont une perte de fonction complète des autres variants de *PCSK1*.

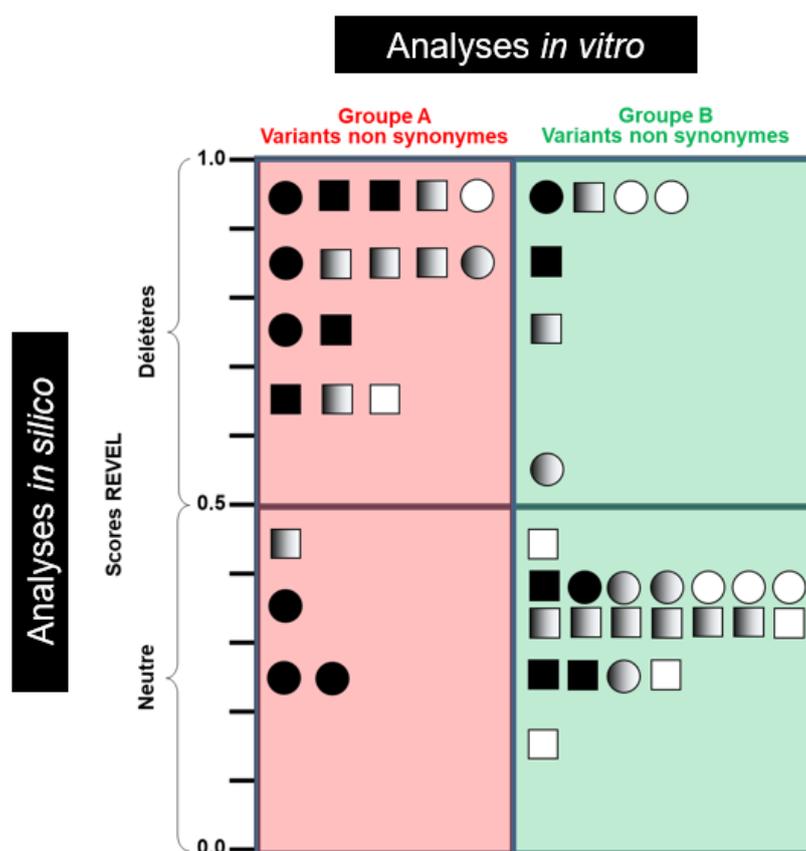
**Tableau 9 : Effet des différents clusters des variants rares de *PCSK1* sur l'obésité et d'autres traits métaboliques chez les adultes selon la pathogénicité des variants sur l'activité de PC1/3.**

Les associations entre les différentes agrégations des variants rares de *PCSK1* et l'adiposité ont été analysées en utilisant la méthode MiST ajustée à l'âge, le sexe et l'origine ethnique. Les associations entre les différentes agrégations des variants rares de *PCSK1* et le risque de DT2 ou le niveau de cholestérols ont été analysées en utilisant la méthode MiST ajustée à l'âge, le sexe, l'IMC et l'origine ethnique. Le cluster I comprend les variants du Groupe\_A et les variants nuls ; le cluster II comprend les variants du Groupe\_B ; le cluster III comprend les variants du Groupe\_C et le cluster IV comprend les variants du Groupe\_D et Groupe\_E. Les données des TG ont été transformées en log avant l'analyse. *IMC*, Indice de masse corporelle ; *n*, nombre d'individus ; *IC*, intervalle de confiance ; *HDL*, lipoprotéine de haute densité ; *IQR*, écart interquartile ;  $\pi$ , effet moyen ; *OR*, odds ratio ; *SD*, déviation standard ; *SE*, erreur standard ; *DT2*, diabète de type 2 ; *TG*, triglycérides.

Trait	<i>n</i>	Cluster	<i>n</i> (%), moyenne $\pm$ SD ou médiane [IQR] du trait chez les porteurs	<i>n</i> (%), moyenne $\pm$ SD ou médiane [IQR] du trait chez les non-porteurs	OR [IC]	$\pi \pm$ SE	<i>P</i>
IMC (kg/m <sup>2</sup> )	7260	I	32 $\pm$ 9,3	27 $\pm$ 6,5	-	6,9 $\pm$ 1,9	3,0 $\times$ 10 <sup>-4</sup>
		II	26 $\pm$ 3,6	27 $\pm$ 6,5	-	-0,99 $\pm$ 1,7	0,56
		III	27 $\pm$ 4,5	27 $\pm$ 6,5	-	0,87 $\pm$ 2,3	0,70
		IV	27 $\pm$ 4,0	27 $\pm$ 6,5	-	1,2 $\pm$ 1,1	0,30
		Tous	28 $\pm$ 5,6	27 $\pm$ 6,5	-	1,6 $\pm$ 0,79	0,039
Obésité	4402	I	6 (86%)	1,518 (35%)	9,3 [1,5–177]	-	0,014
		II	4 (36%)	1,520 (35%)	1,0 [0,27–3,4]	-	0,97
		III	3 (43%)	1,521 (35%)	1,4 [0,27–6,3]	-	0,68
		IV	10 (42%)	1,514 (35%)	1,2 [0,51–2,7]	-	0,66
		Tous	23 (47%)	1,501 (35%)	1,5 [0,85–2,7]	-	0,16
Surpoids ou obésité	7260	I	16 (94%)	4,366 (60%)	8,5 [1,7–154]	-	0,013
		II	15 (68%)	4,367 (60%)	1,4 [0,57–3,6]	-	0,50
		III	8 (67%)	4,374 (60%)	1,1 [0,34–4,3]	-	0,85
		IV	34 (71%)	4,348 (60%)	1,6 [0,85–3,1]	-	0,16
		Tous	72 (73%)	4,310 (60%)	1,7 [1,1–2,8]	-	0,018
DT2	5496	I	11 (65%)	2,167 (40%)	1,2 [0,39–3,8]	-	0,79
		II	9 (47%)	2,169 (40%)	1,7 [0,64–4,3]	-	0,29
		III	5 (56%)	2,173 (40%)	1,3 [0,29–5,9]	-	0,71
		IV	16 (39%)	2,162 (40%)	0,83 [0,39–1,7]	-	0,61
		Tous	41 (48%)	2,137 (40%)	1,11 [0,68–1,8]	-	0,67
HDL (mmol/l)	6864	I	1,4 $\pm$ 0,56	1,5 $\pm$ 0,44	-	0,074 $\pm$ 0,12	0,54
		II	1,5 $\pm$ 0,34	1,5 $\pm$ 0,44	-	-0,048 $\pm$ 0,10	0,65
		III	1,4 $\pm$ 0,41	1,5 $\pm$ 0,44	-	-0,024 $\pm$ 0,14	0,87
		IV	1,5 $\pm$ 0,45	1,5 $\pm$ 0,44	-	0,039 $\pm$ 0,070	0,58
		Tous	1,4 $\pm$ 0,44	1,5 $\pm$ 0,44	-	0,018 $\pm$ 0,049	0,71
TG (mmol/l)	6912	I	1,5 [1,1–1,7]	1,1 [0,77–1,6]	-	-0,015 $\pm$ 0,13	0,91
		II	1,1 [0,75–1,7]	1,1 [0,77–1,6]	-	0,097 $\pm$ 0,11	0,39
		III	1,0 [0,86–1,4]	1,1 [0,77–1,6]	-	-0,080 $\pm$ 0,16	0,61
		IV	1,1 [0,93–1,5]	1,1 [0,77–1,6]	-	-0,0047 $\pm$ 0,076	0,95
		Tous	1,2 [0,86–1,6]	1,1 [0,77–1,6]	-	0,0073 $\pm$ 0,053	0,89

**Tableau 10 : Effet des différents clusters des variants rares de *PCSK1* sur l'obésité chez les enfants selon la pathogénicité des variants sur l'activité de PC1/3.** Les associations entre les différentes agrégations des variants rares de *PCSK1* et le risque d'obésité ont été analysées en utilisant la méthode MiST ajustée à l'âge, le sexe et l'origine ethnique. Le cluster I comprend les variants du Groupe\_A et les variants nuls ; le cluster II comprend les variants du Groupe\_B ; le cluster III comprend les variants du Groupe\_C et le cluster IV comprend les variants du Groupe\_D et Groupe\_E. *n*, nombre d'individus ; *IC*, intervalle de confiance ; *OR*, odds ratio.

Clusters	<i>n</i>	<i>n</i> (%) enfant atteints d'obésité chez les porteurs	<i>n</i> (%) enfant atteints d'obésité chez les non-porteurs	OR [IC]	<i>P</i>
I	2060	5 (83%)	1027 (50%)	4,7 [0,57–103]	0,18
II	2060	2 (40%)	1030 (50%)	2,7 [0,34–17]	0,27
III	2060	6 (66%)	1026 (50%)	2,1 [0,48–11]	0,33
IV	2060	5 (63%)	1027 (50%)	3,6 [0,52–29]	0,21
Tous	2060	17 (63%)	1015 (50%)	2,6 [1,1–6,7]	0,028



**Figure 16 : Répartition des porteurs d'un variant *PCSK1* non synonyme du Groupe\_A et du Groupe\_B en fonction de leur sexe, de leur statut d'obésité, de l'activité enzymatique des variants non synonymes et de la prédiction REVEL de ces variants.**

## 1.4. Discussion

Nos résultats montrent qu'uniquement les variants nuls ou entraînant une perte de fonction complète du gène *PCSK1* causent l'obésité monogénique. En revanche, les variants non synonymes avec un effet neutre ou partiel sur l'activité enzymatique de PC1/3 ne contribuent pas au risque d'obésité ou à l'augmentation de l'IMC dans la population générale. Ces résultats vont dans le sens d'une étude récente de 2021 qui a montré que les variants nuls hétérozygotes de *PCSK1* étaient fortement associés avec une augmentation du risque d'obésité, à partir d'une base de données de 640 000 exomes. Dans cet article, l'effet génétique observé sur l'adiposité était plus élevé chez les porteurs de variants hétérozygotes nuls de *PCSK1* que chez les porteurs de variants hétérozygotes nuls de *MC4R*<sup>75</sup>. De plus, une publication récente a confirmé que la mutation p.Y181H (faisant partie du Groupe\_B dans notre étude) ne contribuait pas au risque d'obésité, ce qui corrobore nos propres résultats<sup>192</sup>.

De plus, nos manipulations fonctionnelles ont montré que l'expression de la protéine PC1/3 était altérée pour la plupart des variants perte de fonction complète de *PCSK1*. En effet, parmi les 17 mutations du Groupe\_A, seules trois mutations présentaient toutes les formes de PC1/3 exprimées de la même manière que la protéine sauvage. Les autres variants montraient tous une altération de l'expression protéique plus ou moins forte : trois mutations ne présentaient aucune forme de PC1/3, trois avaient uniquement la forme zymogène, six avaient les deux formes précurseurs et deux présentaient une forme de PC1/3 anormalement clivée. Ces résultats permettent de classer les mutations de *PCSK1* selon leur impact sur la protéine PC1/3 en trois catégories : les protéines retenues dans le RE suite à un mauvais repliement, ce qui les rend sujettes à la dégradation associée au RE (catégorie 1) ; celles capables de se déplacer vers les granules sécrétoires, mais qui sont altérées dans la voie sécrétoire (catégorie 2) ; et celles correctement repliées et sécrétées, mais avec une activité enzymatique altérée (catégorie 3). L'équipe de Ramos-Molina a proposé une gradation de la gravité de l'obésité chez les porteurs des mutations de *PCSK1*, de la catégorie 3 (c'est-à-dire les porteurs d'obésité moins sévère) à la catégorie 1 (c'est-à-dire les porteurs d'obésité plus sévère). Il faut noter que nous n'avons pas reproduit cette constatation, et que nous n'avons pas observé d'augmentation de la sévérité de l'obésité en fonction de la classification de la protéine mutée dans ces trois catégories. En effet, certains porteurs des mutations qui ne présentaient aucune forme de PC1/3 avaient un IMC inférieur à 30 kg/m<sup>2</sup> (**Tableau 8**). Plus généralement, nous avons observé que les porteurs de variants hétérozygotes nuls ou perte de fonction complète n'avaient pas d'obésité extrême (IMC moyen de 32 kg/m<sup>2</sup>) malgré l'effet dominant de ces mutations. Cependant, une étude

antérieure a suggéré que certains porteurs d'une mutation hétérozygote avec une perte de fonction dans *LEP*, *LEPR*, *POMC* et *PCSK1* pourraient présenter des phénotypes similaires à ceux des porteurs d'une mutation homozygote. Nos résultats concordent davantage avec d'autres études<sup>81,95,99</sup> qui ont montré que l'obésité était moins sévère chez les individus porteurs d'un variant hétérozygote de *PCSK1* et qu'elle touchait une proportion moins élevée de personnes atteintes d'obésité parmi tous les porteurs hétérozygotes que chez les porteurs d'une mutation biallélique de *PCSK1*.

Par ailleurs, les protéines codées par les mutations de *PCSK1* qui entraînent une perte de fonction totale pourraient exercer un effet dominant négatif sur la fonction normale de PC1/3 chez les individus porteurs hétérozygotes. Cet effet a été mis en évidence par des expériences *in vitro*, où des mutations de *PCSK1* connues pour être retenues dans le RE ont été co-transfectées avec la protéine WT, démontrant plusieurs résultats : une interaction entre les formes mutées et la forme sauvage de PC1/3, une réduction du transport de la protéine WT PC1/3 vers les granules sécrétoires<sup>161,167</sup> et une augmentation du stress au niveau du RE qui est reconnu comme facteur contribuant à l'obésité<sup>167,193</sup>. Nos propres manipulations suggèrent également l'existence d'un effet dominant négatif, en particulier pour les mutations générant uniquement la forme zymogène (94kDa) de la protéine, suggérant une possible rétention dans le RE, car elles altèrent plus fortement l'activité enzymatique que les autres mutations. Néanmoins, une analyse plus approfondie de cet effet dominant négatif nécessiterait des manipulations supplémentaires, notamment des analyses d'immunofluorescence afin de visualiser la rétention des protéines mutées au niveau du RE ainsi que des expériences d'immunoprécipitation afin de déterminer si les protéines WT et mutées sont ségréguées ensemble. Il est important de noter que l'effet dominant négatif n'est pas nécessairement observé pour toutes les mutations de *PCSK1*, car cela dépend de la nature spécifique de la mutation et de son impact sur la fonction de PC1/3. Cependant, cet effet pourrait être un mécanisme important dans la compréhension des mutations hétérozygotes de *PCSK1* et de leur contribution dans le développement de l'obésité.

Le mode de transmission des mutations de *PCSK1* semble suivre un schéma dominant avec une pénétrance incomplète. Cela signifie que, bien que des individus soient porteurs de mutations complètement pathogènes, tous ne développent pas systématiquement une obésité. Cette observation est cohérente avec des études antérieures qui ont décrit des porteurs de mutations hétérozygotes nuls ou conduisant à une perte de fonction totale de *PCSK1* comme ayant un poids corporel normal et étant cliniquement sains. Par exemple, la première patiente identifiée

comme porteuse hétérozygote composite pour *PCSK1* a eu quatre enfants hétérozygotes, dont trois possédaient une mutation démontrée comme fortement perte de fonction (p.G593R) et le quatrième avait la mutation *splice\_donor* (c.620+4A>C). Pourtant, tous ces enfants étaient cliniquement normaux<sup>173</sup>. De manière similaire, des parents hétérozygotes d'enfants homozygotes ou hétérozygotes composites pour *PCSK1* ont également été décrits comme ayant un poids corporel normal et cliniquement sain alors qu'ils possédaient une mutation nulle ou démontrée comme très perte de fonction<sup>174,175</sup>. De plus, notre analyse de coségrégation a montré que plusieurs membres des familles porteurs des mutations du Groupe\_A, qui sont pourtant complètement perte de fonction, ne présentaient ni surpoids ni obésité. Ce modèle de dominance avec pénétrance incomplète a également été décrit chez les individus porteurs de mutations hétérozygotes dans *MC4R*<sup>99</sup>.

En outre, comme la société *Rhythm Pharmaceuticals* a lancé des essais cliniques (NCT05093634) afin de tester le setmélanotide pour soigner les patients porteurs de mutations hétérozygotes de *PCSK1*, il est important de s'assurer que ces mutations soient réellement non fonctionnelles, d'autant plus que les porteurs de variants hétérozygotes représentent une population plus large que les porteurs des variants homozygotes. De plus, l'incapacité des tests *in silico* pour détecter précisément les variants hétérozygotes pathogènes de *PCSK1* suggère que les analyses *in vitro* sont nécessaires pour déterminer la pathogénicité des variants avant de proposer un traitement ciblant la voie leptine-mélanocortine. Notamment, cet essai en cours (NCT05093634) évalue également l'effet du setmélanotide chez les patients porteurs de variants hétérozygotes de *POMC* et *LEPR*, alors que leur véritable contribution à l'obésité monogénique est encore incertaine. Par conséquent, les orientations futures de la recherche devraient se concentrer sur la détermination de l'effet fonctionnel des variants hétérozygotes de *LEPR*, ainsi que ceux de *NCOA1* et *SH2B1*, et de leur véritable contribution à l'obésité.

Notre étude a été limitée par le fait qu'il n'était pas possible d'accéder aux concentrations d'insuline et d'ACTH des porteurs des variants perte de fonction complète de *PCSK1*. Cependant, nous n'avons pas trouvé d'association entre ces variants et le DT2 ni entre ces variants et des hauts niveaux de concentrations lipidiques, suggérant un effet minimal de *PCSK1* sur ces voies métaboliques. De plus, une perspective intéressante aurait été d'explorer si les mutations rares complètement perte de fonction de *PCSK1* étaient retrouvées chez des patients souffrant de PWS. En effet, le phénotype observé dans les cas de déficience en PC1/3 présente plusieurs similitudes avec celui des patients atteints de PWS, notamment une obésité précoce avec une hyperphagie, une déficience en GH et un hypogonadisme. Cette similarité

suggère la possibilité d'un rôle de PC1/3 dans les aspects cliniques des patients atteints de PWS<sup>194</sup>. De plus, une étude a montré que les niveaux de PC1/3 étaient réduits dans les neurones dérivés de cellules souches pluripotentes provenant de patients atteints de PWS<sup>136</sup>. Cependant, l'impact de PC1/3 sur ce phénotype n'est pas encore clairement compris.

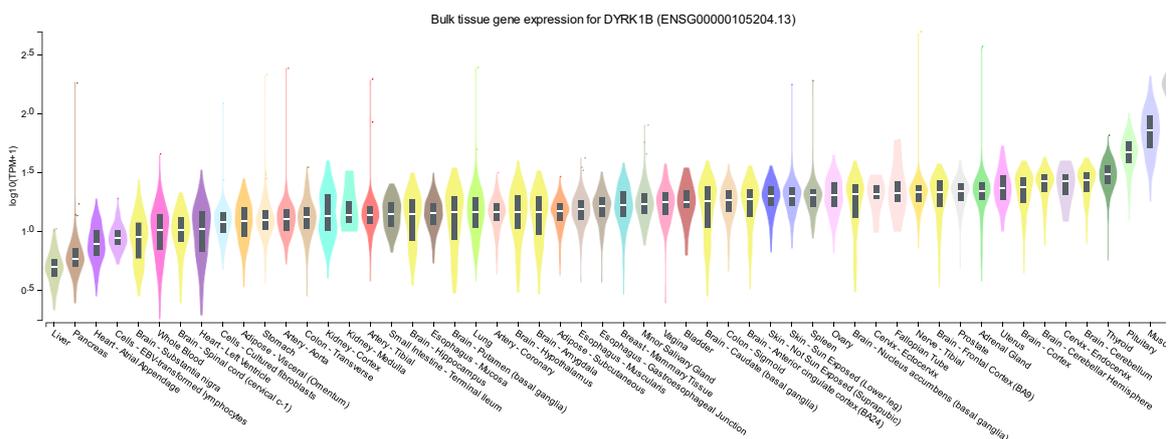
De plus, d'autres limites peuvent également être appliquées à notre étude. Notamment, les biais de cette étude sont similaires à ceux de toute biobanque basée sur un recrutement volontaire. De plus, les analyses fonctionnelles *in vitro* ont été réalisées dans des cellules HEK293, qui n'expriment pas PC1/3 dans des conditions physiologiques. Ces cellules diffèrent largement des cellules neuronales qui ont un impact important sur le phénotype de l'obésité. Cependant, cette limitation a été surmontée en concevant une grande étude de cas-témoins d'association génétique en utilisant nos essais *in vitro* standard. Enfin, il est important de noter que cette étude n'implique pas que les patients atteints d'une obésité sévère précoce portant une mutation hétérozygote neutre de *PCSK1* ne doivent pas être davantage examinés, car une autre mutation pourrait avoir un effet délétère sur la voie leptine-mélanocortine, et donc ces patients pourraient être éligibles pour un traitement au setmélanotide.

## Partie 2: Les variants pathogènes avec perte de fonction totale de *DYRK1B* causent une obésité monogénique, associée à un diabète de type 2

### 2.1. Contexte de l'étude

#### 2.1.1. La biologie de *DYRK1B*

Le gène *DYRK1B* (*dual-specificity tyrosine(Y)-phosphorylation regulated kinase 1B*), anciennement appelé *MIRK* (*minibrain-related kinase*), est compris dans la famille des protéines kinases *DYRK* comprenant 5 membres. Cette famille est caractérisée par un domaine kinase hautement conservé appelé domaine d'homologie DYRK (DH). L'autophosphorylation de leur résidu tyrosine (p.Y273), qui est hautement conservé dans une boucle d'activation du domaine catalytique, est indispensable pour l'acquisition de leur conformation catalytiquement active. Ces kinases jouent un rôle essentiel dans la régulation de divers processus cellulaires, notamment la progression du cycle cellulaire, la différenciation cellulaire, la survie cellulaire, la motilité, le développement embryonnaire et la transcription génique par phosphorylation des résidus sérine et thréonine des protéines cibles<sup>195</sup>. Le gène *DYRK1B* est situé sur le chromosome 19 et il est composé de 14 exons. Il possède trois isoformes et un gène paralogue qui est *DYRK1A*. Son expression est ubiquitaire et elle est élevée principalement dans les tissus musculaires et les testicules (**Figure 17**)<sup>196,197</sup>



**Figure 17 :** L'expression génique de *DYRK1B* au sein des différents tissus. Source des données : GTEX Analysis Release V8 (dbGaP Accession phs000424.v8.p2) (GTEX Portal, accès octobre 2023).

Le gène *DYRK1B* code la protéine kinase du même nom composée de 629 acides aminés. Elle est exprimée de façon ubiquitaire dans les noyaux chez l'humain et la souris. Elle est connue comme étant un régulateur de la différenciation cellulaire et de la progression du cycle cellulaire. Par exemple, *DYRK1B* phosphoryle directement des régulateurs cellulaires tels que la cycline

D1 (CCND1) et l'inhibiteur de kinase dépendante de la cycline 1B (CDKN1B), déclenchant leur dégradation et leur stabilité, entraînant l'arrêt de la croissance et la promotion de la quiescence<sup>198</sup>. De plus, DYRK1B est impliqué dans une interconnexion complexe avec les voies de signalisation hedgehog (Shh) et Wnt, exerçant un rôle double dans la modulation de ces voies, soit de manière positive soit négative. Cette modulation dépend du type cellulaire et de l'implication spécifique de signalisation canonique ou non canonique<sup>195,198</sup>. La fonction de DYRK1B a principalement été explorée dans la différenciation des cellules musculaires squelettiques et des adipocytes, la spermatogenèse et le cancer<sup>195,198</sup>. Notamment, l'expression de *DYRK1B* est surexprimée dans certains types de cancers incluant le cancer du sein, le cancer pancréatique, le cancer du poumon et le cancer de l'ovaire. *DYRK1B* semble jouer un rôle dans la régulation de la croissance tumorale et la progression du cancer. Par conséquent, ce gène est devenu une cible potentielle pour le développement de thérapies anticancéreuses ciblées<sup>199</sup>.

### 2.1.2. La génétique de *DYRK1B*

#### 2.1.2.1. Chez la souris

Il n'y a pas de modèle murin spécifique pour *Dyrk1b* largement reconnu dans la littérature scientifique. Le modèle murin *knockout Dyrk1b* existe mais ne présente pas de phénotype évident<sup>200</sup>. Une étude réalisée sur des souris a mis en évidence que l'augmentation de l'expression de *Dyrk1b* au niveau du foie renforçait la lipogenèse *de novo*, l'absorption des acides gras et la sécrétion des triglycérides chez la souris, ce qui pourrait provoquer la NASH et une hyperlipidémie. En revanche, une diminution de l'expression de *Dyrk1b* hépatique protégerait contre la stéatose hépatique, la fibrose induite par un régime riche en calories et l'hyperlipidémie chez la souris<sup>201</sup>.

#### 2.1.2.2. Chez l'humain

##### a) Les SNP fréquents identifiés dans *DYRK1B*

Uniquement deux SNP ont été identifiés par GWAS dans *DYRK1B* (d'après GWAS catalog). Le premier était associé avec une grande proportion de bactéries de la famille *Porphyromonadaceae* dans une étude portant sur le microbiote<sup>202</sup>. Le second était associé avec l'épaisseur de la macula, la zone centrale de la rétine<sup>203</sup>. Aucun SNP n'a été identifié comme étant lié à des caractéristiques métaboliques.

##### b) Les mutations rares identifiées dans *DYRK1B*

L'étude initiale mettant en évidence le rôle de DYRK1B dans l'AOMS décrivait deux variants rares non synonymes dans *DYRK1B* (codant pour p.R102C et p.H90P [NM\_004714.3]) dans

trois familles iraniennes et cinq individus européens non apparentés<sup>134</sup>. Ces mutations induisaient une expression accrue de l'enzyme glucose-6-phosphatase (G6PC), une enzyme hépatique importante pour la néoglucogénèse. De plus, le variant p.R102C avait un effet inhibiteur plus puissant sur la voie Wnt et un effet stimulateur plus fort sur les facteurs de transcription adipogéniques dans les préadipocytes de souris, favorisant ainsi l'adipogenèse<sup>134</sup>. Cependant, contrairement à ces effets apparemment gain de fonction, des tests *in vitro* issus d'une autre étude ont en réalité démontré une réduction de l'activité kinase du variant codant pour p.R102C, suggérant plutôt un effet perte de fonction<sup>204</sup>. Deux variants rares hétérozygotes de *DYRK1B* (codant pour p.K68Q et p.R252H [NM\_004714.3]), prédits comme étant délétères selon les programmes *in silico*, ont ensuite été découverts dans deux familles atteintes d'AOMS<sup>205</sup>. Plus récemment, un variant nul rare hétérozygote de *DYRK1B* (c.520+1G>A [NM\_004714.3]), intrinsèquement lié à la perte de la fonction protéique, a été identifié chez un père et ses deux filles qui présentaient une AOMS et une déficience intellectuelle à plusieurs degrés<sup>206</sup>.

Cependant, depuis 2014, cette étude n'a jamais été répliquée et l'impact global des variants de *DYRK1B* sur les caractéristiques métaboliques reste à explorer de manière approfondie. En particulier, les mécanismes par lesquels les variants de *DYRK1B* affectent l'homéostasie métabolique (qu'il s'agisse d'un gain ou d'une perte de fonction) restent mystérieux. Dans cette optique, nous avons réalisé une étude génomique fonctionnelle à grande échelle portant sur les variants rares de *DYRK1B* dans le but de dévoiler leurs effets sur les caractéristiques métaboliques, notamment en ce qui concerne le risque d'obésité et de diabète de type 2 chez l'humain. Un article basé sur ce travail a été publié en janvier 2024 dans *Diabetes Care* (**Annexe 4**).

## 2.2. Matériel et Méthode

### 2.2.1. Participants de l'étude

Cette étude cas-témoins a inclus 9353 participants de l'étude RaDiO, qui a été préalablement décrite dans la partie 1 (cf supra). Les cas comprennent des adultes ou des enfants en surpoids, ou atteints d'obésité et/ou diabétiques tandis que les témoins comprennent des adultes ou des enfants ayant un poids corporel sain et une glycémie normale. Chez les adultes de plus de 18 ans, un poids corporel sain était défini comme un IMC inférieur à 25 kg/m<sup>2</sup>, le surpoids comme un IMC de 25 à 29,9 kg/m<sup>2</sup> ou plus, et l'obésité comme un IMC de 30 kg/m<sup>2</sup> ou plus. Chez les enfants et les adolescents de moins de 18 ans, un poids corporel sain était défini comme un IMC-pour-âge inférieur au 85ème percentile, et l'obésité comme un IMC-pour-âge supérieur ou

égal au 95ème percentile. Les participants atteints de DT2 avaient une glycémie à jeun de 7,0 mmol/l ou plus, ou étaient sous traitement pour l'hyperglycémie (ou les deux). Les témoins avaient une glycémie à jeun inférieure à 6,1 mmol/l à l'âge de plus de 30 ans et n'utilisaient pas de traitement pour l'hyperglycémie. Tous les participants et les parents (ou tuteurs légaux) ont fourni un consentement écrit éclairé. Les enfants et les adolescents ont donné leur assentiment oral. Les protocoles de l'étude ont été approuvés par les comités d'éthique locaux.

## 2.2.2. Séquençage de *DYRK1B*

### 2.2.2.1. Séquençage NGS

Le séquençage de *DYRK1B* (NM\_004714.3) a été réalisé par NGS, comme décrit dans la section séquençage de la partie 1 (cf supra). Nous avons sélectionné uniquement les variants rares et codants de *DYRK1B* avec une MAF inférieure à 1% et avec une couverture de plus de 40 *reads*. La majorité des variants de *DYRK1B* avaient un score de qualité supérieur à 150, et ceux avec un score de qualité inférieur à 150 ont été confirmés par séquençage de Sanger. De plus, aucun variant n'avait plus de 5 % de génotypes manquants parmi les participants, et aucun participant n'avait plus de 5 % de génotypes manquants pour *DYRK1B*.

### 2.2.2.2. Séquençage Sanger

Les exons 6 et 8 de *DYRK1B* ont été séquencés pour les analyses de coségrégation par séquençage Sanger dans deux familles qui étaient disponibles. Les séquences des amorces ainsi que les conditions PCR utilisées sont détaillées respectivement en **Annexe 1, Tableau S10** et **Tableau S11**. Les exons ont été séquencés dans les deux sens, et analysés en utilisant le 3730xl DNA Analyzer. Les *reads* ont été assemblés et examinés en utilisant les logiciels Variant Reporter et SeqScape.

## 2.2.3. Analyses fonctionnelles

### 2.2.3.1. Mutagénèse dirigée

Des plasmides incluant 65 différentes mutations de *DYRK1B* qui ont été identifiées dans RaDiO et un variant contrôle négatif (portant les deux mutations c.812A>T and c.818A>T codant respectivement p.Y271F et p.Y273F [noté p.Y271/273F pour la suite] avec un domaine kinase inactivé<sup>207</sup> ont été créés par mutagénèse en utilisant des amorces portant la mutation souhaitée (**Annexe 1, Tableau S12**) ainsi qu'un plasmide *DYRK1B* WT (Origene). Les séquences des plasmides ont été vérifiées par séquençage Sanger (**Annexe 1, Tableau S13**).

### 2.2.3.2. Culture cellulaire

#### a) Lignée cellulaire HEK293

Les manipulations de luciférase et de western blotting ont été réalisées sur le modèle cellulaire HEK293. Les conditions de culture des HEK293 étaient les mêmes que celles de la partie 1 (cf supra).

#### b) Lignée cellulaire HepG2

Des expériences de RT-qPCR et de western blotting ont été réalisées sur la lignée cellulaire de carcinome hépatocellulaire humain HepG2. Des cellules HepG2 ont été maintenues dans du milieu Dulbecco's Modified Eagle's (Thermo Fisher Scientific) contenant 10% de SVF et 1% de pénicilline-streptomycine et 1% d'acides aminés non-essentiels (NEAA) (Thermo Fisher Scientific). Les cellules ont été incubées avec 5% de CO<sub>2</sub> à 37°C.

### 2.2.3.3. Expériences de luciférase

Les cellules HEK293 ont été transfectées à l'aide de FuGENE HD avec 200 ng/mL de plasmide codant pour le gène de la  $\beta$ -galactosidase, 350 ng/mL de plasmide TOPflash (c'est-à-dire le plasmide rapporteur TCF ; Sigma-Aldrich) et 450 ng/mL de plasmide contenant le gène *DYRK1B* WT ou un des variants identifiés. Trois conditions contrôles ont été utilisées pour chaque manipulation : la condition non transfectée (NT) où seulement le plasmide portant le gène de la  $\beta$ -galactosidase est exprimé ; la condition WT exprimant le gène *DYRK1B* WT ; et la condition exprimant la mutation p.Y271/273F où le domaine kinase est inactivé (comme contrôle négatif). Les cellules ont été cultivées en plaque 96 puits préalablement coâtée à la poly-D-lysine (Gibco) à une concentration de 500 000 cellules/mL. Après 48h de transfection, les cellules ont été traitées avec des concentrations croissantes (0, 10, 30 et 100 ng/mL) du ligand WNT3A (R&D Systems) pendant 6 heures. Après le traitement, les cellules sont lysées dans 250  $\mu$ L de *Luciferase Cell Culture Lysis Buffer* (Promega). Après 30 min sous agitation, les lectures de luciférase et de  $\beta$ -galactosidase sont réalisées en parallèle. La luminescence émise par l'activité luciférase a été évaluée en ajoutant 25  $\mu$ L de *Luciferase Assay System reagent* (Promega) aux 40  $\mu$ L de lysat cellulaire, puis en la mesurant à l'aide de luminomètre Glomax. L'activité de la  $\beta$ -galactosidase a été mesurée de façon similaire à ce qui a été décrit pour la partie 1 (cf supra). Les expériences ont été réalisées en triplicat technique et chaque expérience a été répétée quatre fois.

Les mesures d'activité de luciférase ont été normalisées par les mesures de  $\beta$ -galactosidase. Le *fold change* a été calculé en divisant les mesures d'activité luciférase normalisées par la

moyenne de l'activité basale du WT (sous la condition contrôle [0 ng/mL]). L'effet de chaque variants (*versus* la condition WT) a été analysé en utilisant un modèle de régression linéaire ajusté aux concentrations de WNT3A et à l'effet des plaques. En raison de la déviation de la normalité observée dans les résidus de certains modèles de régression linéaire, nous avons assuré la validité des résultats statistiques en utilisant le test robuste de Scheirer-Ray-Hare, un test ANOVA non paramétrique, ajusté à la concentration de WNT3A, à l'interaction entre le variant et la concentration de WNT3A. Ces analyses ont été réalisées en utilisant la version 4.2.1 de R. Les variants pathogènes ou probablement pathogènes (P/LP) avec un  $\beta < -0.6$  et une p-valeur significative à la fois en modèle de régression linéaire et en ANOVA, ont été catégorisés avec les variants nuls dans les variants P/LP totalement inhibiteur (noté P/LP-full).

#### 2.2.3.4. Western blotting

##### a) *Analyse de l'expression protéique de la CCND1 phosphorylée*

Les cellules HEK293 ont été transfectées à l'aide de FuGENE HD avec 500 ng/mL du plasmide pcDNA3.1-HA-CCND1 (Addgene, 172649<sup>208</sup>) et 500 ng/mL de plasmide contenant le gène *DYRK1B* (vecteur vide [VV], WT ou avec un P/LP-full variant). Trois conditions contrôles ont été utilisées pour chaque manipulation : un contrôle positif contenant le plasmide *DYRK1B* WT, un contrôle négatif contenant le plasmide exprimant la mutation Y271/273F et un autre contrôle négatif avec un vecteur vide (VV) sans le gène *DYRK1B* (Origene). Les cellules ont été cultivées en plaque 6 puits précoatée à la poly-D-lysine à une densité de 500 000 cellules/mL. 48h après la transfection, les cellules ont été récoltées et les protéines ont été extraites, quantifiées et dénaturées de façon similaire à ce qui a été décrit dans la section « Western blotting » de la partie 1 (cf supra). 20  $\mu$ g de protéines ont été déposées pour la migration par électrophorèse sur un gel precast 10%. Après la migration, les protéines ont été transférées sur une membrane de nitrocellulose et les sites non-spécifiques sur la membrane ont été bloqués sous agitation pendant 1 h dans du tampon de blocage (composé de TBS, 0,1% Tween 20, et 5% d'albumine de sérum bovin [BSA]). Ensuite, les membranes ont été incubées toute la nuit à 4°C sous agitation avec les anticorps primaires : DYRK1B (#5672 Cell signaling Technology; dilué à 1/1000 dans du tampon de blocage), phospho-CCND1 (à l'acide aminé p.T286) (#3300 Cell signaling Technology; dilué à 1/1000 dans du tampon de blocage) and CCND1 (Origene; dilué à 1/2000 dans du tampon de blocage), suivi par un anticorps secondaire anti-lapin fluorescent (SA5-35571, Thermo Fisher Scientific ; dilué à 1/5000 dans du tampon de blocage) et un anticorps secondaire anti-souris fluorescent (#35518 Thermo Fisher Scientific ; dilué à 1/5000 dans du tampon de blocage). Les membranes de nitrocellulose ont été révélées à l'aide

de l'Odyssey CLx imaging system. Afin de normaliser les données, la  $\beta$ -actine a été utilisée comme contrôle avec l'anticorps primaire anti- $\beta$ -actine (#3700 Cell Signaling Technology ; dilué à 1/4000 dans le tampon de blocage). Quatre expériences indépendantes ont été réalisées pour chaque mutation. Les niveaux d'expression protéique ont été quantifiés en utilisant le programme ImageJ. L'expression relative de la phospho-CCND1 a été calculée en divisant l'intensité de la phospho-CCND1 par l'intensité de CCND1, suivi par la normalisation avec la  $\beta$ -actine. Le *fold change* a été obtenu en divisant l'expression de la protéine normalisée par la moyenne de l'expression protéique de la condition WT. L'effet de chaque variant (*versus* la condition WT) a été analysé à l'aide du test Mann-Whitney.

#### b) Analyse de l'expression protéique de la G6PC

Les cellules HepG2 ont été cultivées à une densité de 350 000 cellules/mL dans une plaque 6 puits. 24h après avoir été platées, les HepG2 ont été transfectées à l'aide de FuGENE HD avec 1000 ng/mL de plasmide contenant le vecteur vide [VV], le gène *DYRK1B* WT ou un des variants p.Y271/273F, p.R102C, et p.R358\*. 48h après la transfection, les cellules ont été récoltées et les protéines ont été extraites, quantifiées et dénaturées de façon similaire à ce qui a été décrit dans la section « Western blotting » de la partie 1 (cf supra). 35  $\mu$ g de protéines ont été déposées pour la migration par électrophorèse sur un gel precast 10%. Après la migration, les protéines ont été transférées sur une membrane de nitrocellulose et les sites non-spécifiques sur la membrane ont été bloqués sous agitation pendant 1 h dans du tampon de blocage (composé de TBS, 0,1% Tween 20, et 5% de lait écrémé). Ensuite, les membranes ont été incubées toute la nuit à 4°C sous agitation avec les anticorps primaires : DYRK1B (#5672 Cell signaling Technology; dilué à 1/1000 dans du tampon de blocage), G6PC (ab93857, Abcam ; dilué à 1/1000 dans du tampon de blocage) et  $\beta$ -actine (#3700 Cell Signaling Technology ; dilué à 1/4000 dans le tampon de blocage) suivi par un anticorps secondaire anti-lapin fluorescent (SA5-35571, Thermo Fisher Scientific ; dilué à 1/5000 dans du tampon de blocage) et un anticorps secondaire anti-souris fluorescent (#35518 Thermo Fisher Scientific ; dilué à 1/5000 dans du tampon de blocage). Les membranes de nitrocellulose ont été révélées à l'aide de l'Odyssey CLx imaging system. Trois expériences indépendantes ont été réalisées pour chaque mutation.

#### 2.2.3.5. Extraction d'ARN, rétrotranscription et qPCR

Les cellules HepG2 ont été cultivées à une densité de 350 000 cellules/mL dans une plaque 6 puits. 24h après avoir été platées, les cellules ont été transfectées à l'aide de FuGENE HD avec 1000 ng/mL de plasmide contenant le gène *DYRK1B* WT ou un des variants p.Y271/273F,

p.R102C, et p.R358\*. 48h après la transfection, les cellules ont été récoltées et l'ARN a été extrait en utilisant le kit *Nucleospin RNA* selon les instructions du fournisseur. 400 ng de l'ARN extrait ont ensuite été rétrotranscrits en employant le mélange *Transcriptor Universal cDNA Master* (Roche). Une PCR quantitative (qPCR) a ensuite été réalisée : 2  $\mu$ L d'ADN complémentaire dilué au 1/8ème ont été déposés en triplicat dans une plaque 384 puits où ont été ajoutés 2,4  $\mu$ L d'eau, 0,6  $\mu$ L d'amorces spécifiques pour *DYRK1B* (NM\_004714.3) (Integrated DNA technologies ; **Annexe 1, Tableau S14**) et 5  $\mu$ L du mix *Power SYBR Green PCR Master* (Applied Biosystems). La qPCR a ensuite été exécutée sur le QuantStudio 7 Pro (Thermo Fisher Scientific) selon les conditions détaillées en **Annexe 1, Tableau S7**. Le gène *RPLP0* a été utilisé comme gène de ménage. Pour chaque condition, trois expériences indépendantes ont été réalisées. Les résultats ont été analysés à l'aide du logiciel Design & Analysis 2.6.0. (Thermo Fisher Scientific).

#### 2.2.4. Analyses statistiques pour les études d'association

Nous avons évalué l'effet de trois clusters différents (ALL = tous les variants rares ; P/LP = tous les variants P/LP ; P/LP-full = tous les variants P/LP totalement inhibiteurs) sur des traits métaboliques. Chaque cluster a été analysé en utilisant la méthode MiST (cf supra) ajustée à l'âge, au sexe et l'origine ethnique. L'analyse MiST donne un score statistique  $S(\pi)$  pour l'effet moyen ( $\pi$ ) du cluster, et un score statistique  $S(\tau)$  pour l'effet hétérogène ( $\tau$ ) du cluster. Ces analyses statistiques ont été réalisées en utilisant le logiciel R (version 4.0.2).

La méta-analyse entre notre étude et les cohortes 52K et TopMED (qui sont disponibles sur le *Type 2 Diabetes knowledge portal*<sup>209</sup> (accès en aout 2023) a été réalisée en utilisant la méthode générique de la variance inverse provenant du package R meta<sup>210</sup>. Uniquement les variants P/LP-full ont été inclus dans cette méta-analyse.

## 2.3. Résultats

### 2.3.1. Détection de 65 variants rares codant de *DYRK1B* dans RaDiO

De façon similaire au projet *PCSK1*, nous avons réalisé un séquençage NGS de *DYRK1B* chez 9353 participants (incluant 7268 adultes et 2085 enfants) provenant de l'étude RaDiO afin d'identifier des variants rares du gène *DYRK1B*. Grâce à ce séquençage, nous avons identifié 65 mutations rares (MAF < 1%) du gène *DYRK1B* (NM\_00714.2) qui étaient toutes hétérozygotes. Parmi ces variants, nous comptons une seule mutation non-sens (codant p.R358\*) et 64 mutations non synonymes (**Tableau 11**). Nous avons modélisé tous ces

différents variants dans des plasmides par mutagenèse afin de réaliser nos analyses fonctionnelles *in vitro*.

**Tableau 11 : Variants rares de *DYRK1B* détectés à partir de l'étude RaDiO.** *MAC* : *minor allele count* ; *gnomAD* : *genome aggregation database* (version 2.1.1) ; *NA* : non disponible ; *REVEL* : *rare exome variant ensemble learner* ; *P/LP* : variant pathogène/probablement pathogène ; *P/LP-full* : variant pathogène/probablement pathogène avec un effet totalement inhibiteur. Le seuil de pathogénicité pour REVEL était de 50 %.

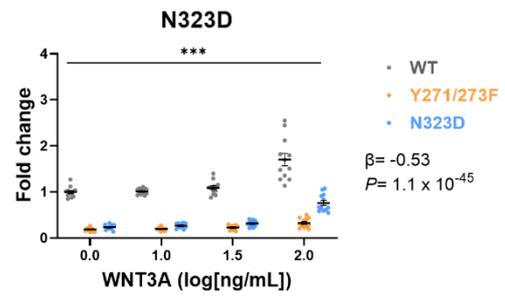
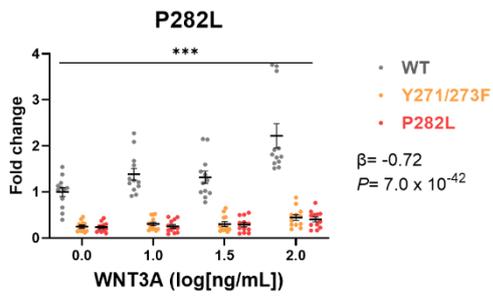
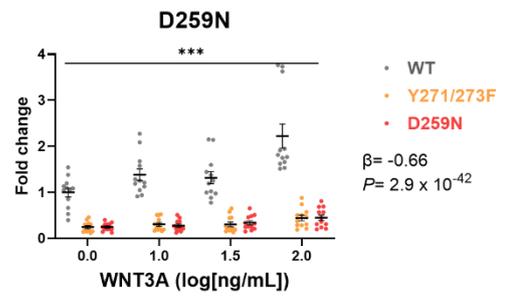
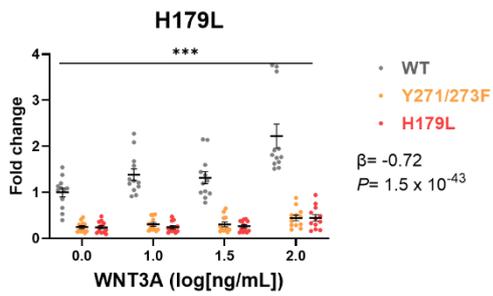
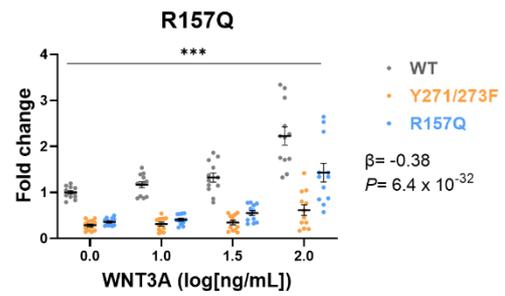
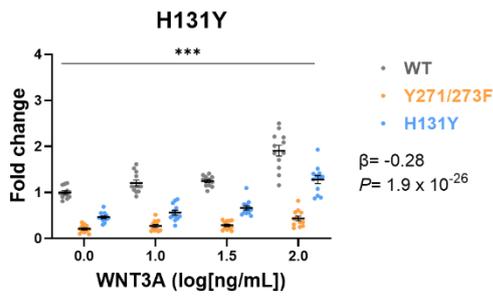
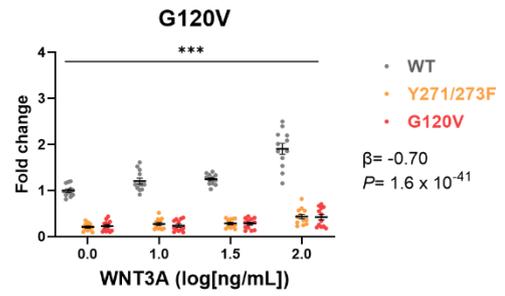
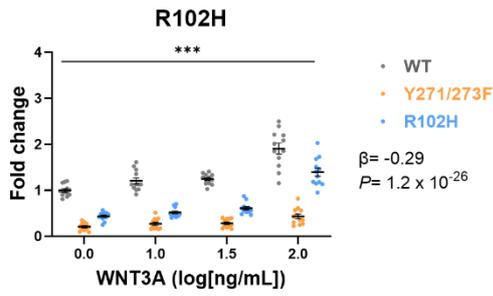
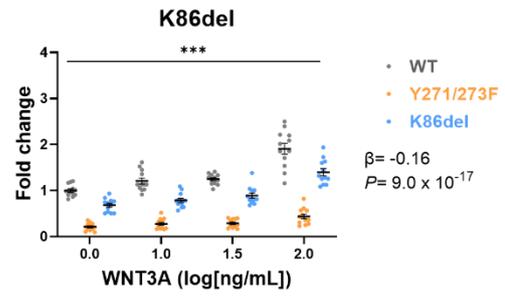
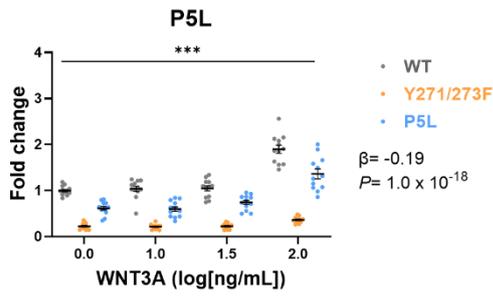
Rare variants in <i>DYRK1B</i> (NM_004714.3)	MAC dans gnomAD	MAC dans RaDiO	Catégories	REVEL
c.7G>A, p.V3I	7	1	Neutre	Neutre
c.14C>T, p.P5L	1	1	P/LP	Neutre
c.80G>A, p.R27Q	6	2	Neutre	Neutre
c.92G>C, p.R31P	0	1	Neutre	Neutre
c.118G>T, p.A40S	0	1	Neutre	Neutre
c.170A>G, p.K57R	3	1	Neutre	Neutre
c.202A>C, p.K68Q	5	1	Neutre	Neutre
c.209G>A, p.R70Q	17	1	Neutre	Neutre
c.236C>T, p.S79L	2	1	Neutre	Neutre
c.256_258del, p.K86del	0	1	P/LP	NA
c.305G>A, p.R102H	8	2	P/LP	Neutre
c.359G>T, p.G120V	0	1	P/LP-full	Délétère
c.391C>T, p.H131Y	0	1	P/LP	Neutre
c.470G>A, p.R157Q	1	1	P/LP	Neutre
c.500C>T, p.T167M	10	1	Neutre	Neutre
c.506T>G, p.M169R	0	2	Neutre	Neutre
c.515A>G, p.Y172C	3	1	Neutre	Neutre
c.526C>A, p.L176M	16	1	Neutre	Neutre
c.536A>T, p.H179L	0	1	P/LP-full	Délétère
c.668C>T, p.T223M	30	3	Neutre	Neutre
c.746A>G, p.N249S	1	2	Neutre	Neutre
c.775G>A, p.D259N	0	1	P/LP-full	Délétère
c.845C>T, p.P282L	0	1	P/LP-full	Délétère
c.967A>G, p.N323D	0	1	P/LP	Neutre
c.971G>T, p.R324L	5	1	Neutre	Neutre
c.1003G>A, p.A335T	3	4	Neutre	Neutre
c.1030C>T, p.R344C	8	2	Neutre	Neutre
c.1031G>A, p.R344H	14	2	Neutre	Neutre
c.1045C>T, p.R349W	1	1	P/LP-full	Neutre
c.1046G>A, p.R349Q	2	1	Neutre	Neutre
c.1054G>A, p.G352R	0	1	P/LP	Neutre
c.1055G>C, p.G352A	155	5	Neutre	Neutre
c.1057G>T, p.G353C	1	1	P/LP	Neutre
c.1072C>T, p.R358*	0	1	P/LP-full	NA
c.1073G>A, p.R358Q	4	1	P/LP	Neutre
c.1079C>A, p.T360K	0	1	P/LP	Neutre
c.1111G>A, p.G371R	2	3	P/LP	Neutre
c.1111G>C, p.G371R	1	1	P/LP	Neutre
c.1196C>A, p.A399D	0	1	Neutre	Neutre
c.1208G>A, p.R403H	1	1	Neutre	Neutre
c.1229G>A, p.R410H	0	1	P/LP	Neutre
c.1252G>A, p.A418T	22	9	Neutre	Neutre
c.1285G>A, p.G429S	103	3	Neutre	Neutre
c.1285G>C, p.G429R	5	3	Neutre	Neutre
c.1295G>A, p.R432H	1	1	Neutre	Neutre
c.1328C>T, p.P443L	0	1	Neutre	Neutre
c.1336A>G, p.S446G	11	3	Neutre	Neutre
c.1341T>A, p.S447R	0	1	P/LP	Neutre

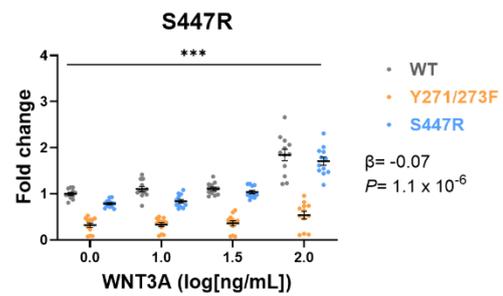
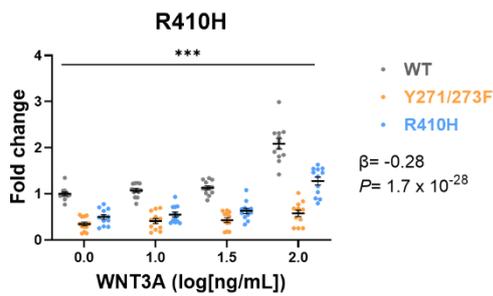
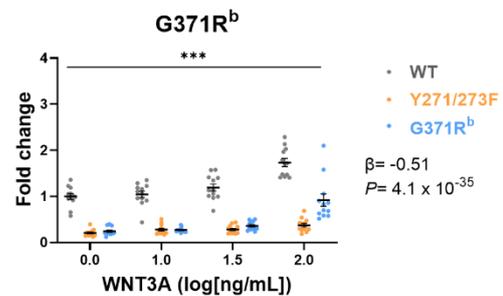
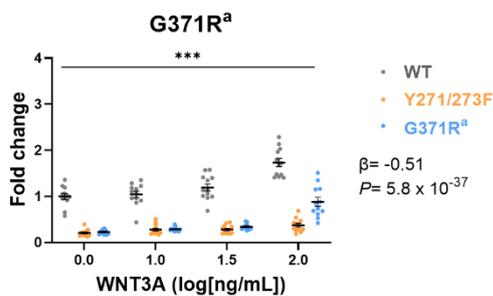
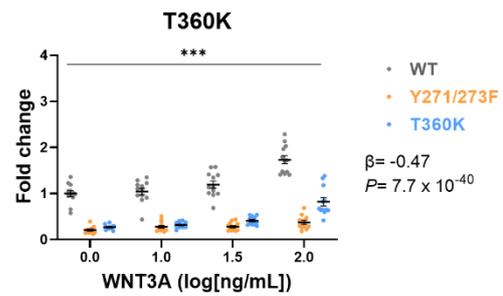
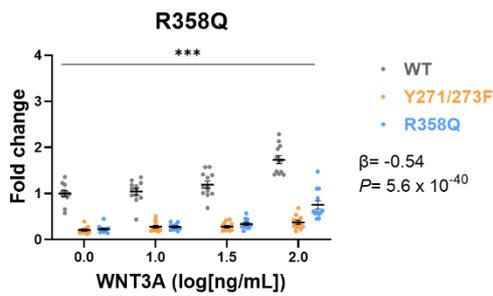
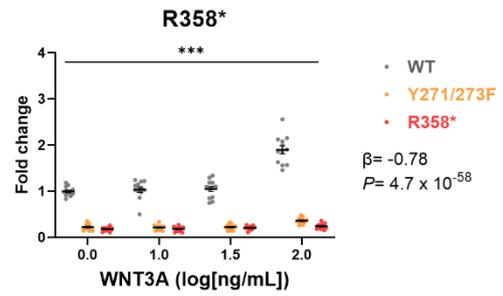
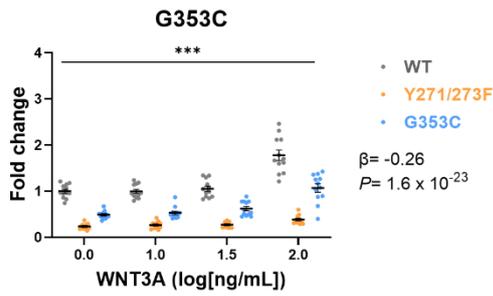
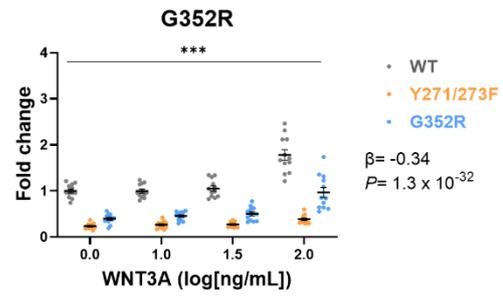
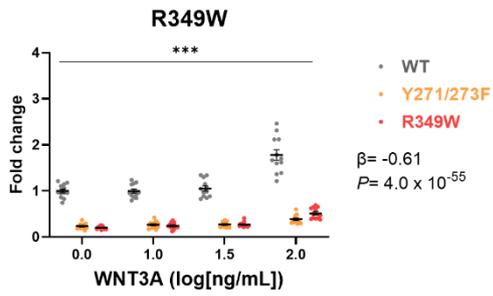
c.1349C>T, p.T450I	0	1	Neutre	Neutre
c.1358C>T, p.A453V	5	1	Neutre	Neutre
c.1386C>G, p.S462R	101	5	Neutre	Neutre
c.1414G>A, p.G472S	0	2	Neutre	Neutre
c.1441C>T, p.R481W	0	1	Neutre	Neutre
c.1450C>T, p.R484C	7	1	Neutre	Neutre
c.1463G>A, p.R488Q	14	1	Neutre	Neutre
c.1469G>T, p.C490F	0	1	Neutre	Neutre
c.1470T>G, p.C490W	8	3	Neutre	Neutre
c.1481G>A, p.G494E	11	1	Neutre	Neutre
c.1675C>T, p.P559S	2	1	Neutre	Neutre
c.1732C>T, p.P578S	294	31	Neutre	Neutre
c.1742C>T, p.A581V	2	1	Neutre	Neutre
c.1799G>A, p.R600H	5	1	Neutre	Neutre
c.1823A>T, p.D608V	0	1	Neutre	Neutre
c.1840C>A, p.P614T	38	2	Neutre	Neutre
c.1855C>T, p.R619C	6	1	Neutre	Neutre

### 2.3.2. Etudes fonctionnelles *in vitro* des variants rares du gène *DYRK1B*

#### 2.3.2.1. Analyse de l'effet sur la voie Wnt et classification des variants P/LP selon les critères ACMG

Nous avons utilisé les critères ACMG pour évaluer la pathogénicité de chaque variant détecté. Pour déterminer le critère de pathogénicité fort PS3, nous avons évalué l'effet fonctionnel de chaque variant de *DYRK1B* sur la voie de signalisation Wnt à l'aide du test de gène rapporteur luciférase TCF/LEF. Des cellules HEK293 surexprimant les différents variants de *DYRK1B*, la  $\beta$ -galactosidase ainsi que le plasmide TOPflash ont été soumises à des concentrations croissantes du ligand WNT3A qui va activer la signalisation Wnt (**Figure 18** et **Annexe 2, Figure S8**). En comparaison avec la condition WT, nous trouvons que 31 variants de *DYRK1B* inhibent significativement la voie Wnt. En utilisant les critères ACMG intégrant le critère PS3 (**Annexe 1, Tableau S15**), nous avons identifié 20 variants P/LP de *DYRK1B* (**Figure 18**). Parmi ces variants P/LP, six variants (codant p.G120V, p.H179L, p.D259N, p.P282L, p.R349W et p.R358\*) ont montré un effet entièrement inhibiteur (P/LP-full) sur la voie de signalisation Wnt. Cet effet inhibiteur était similaire à celui observé pour le contrôle négatif, c'est-à-dire le plasmide *DYRK1B* portant la double mutation artificielle p.Y271/273F où le domaine kinase est inactivé (**Figure 18**).





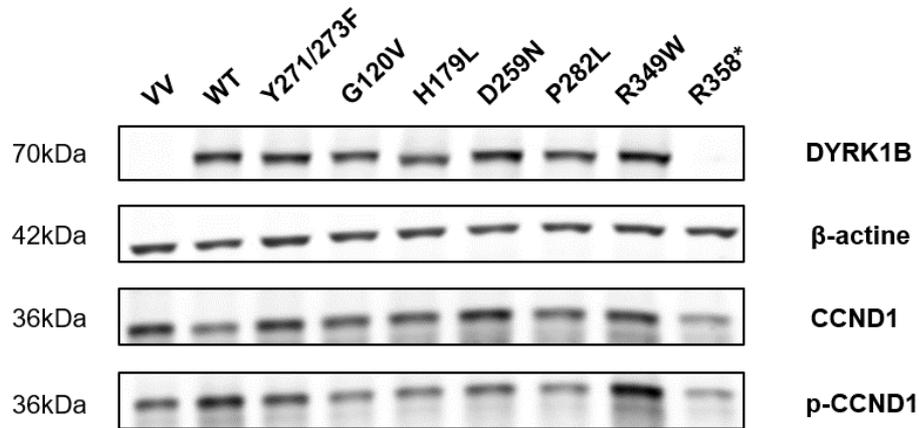
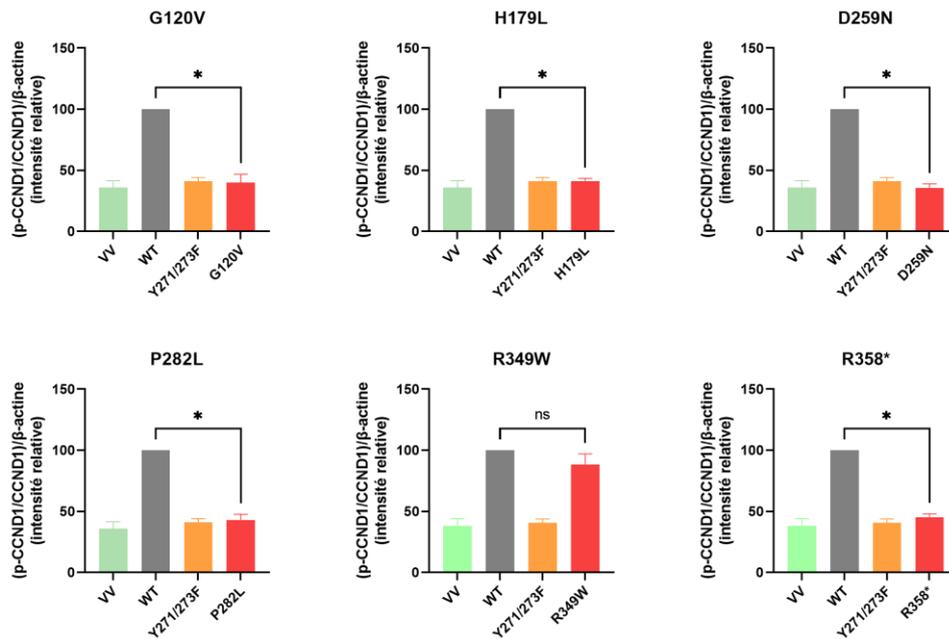
**Figure 18 : Effet des variants P/LP de *DYRK1B* sur la signalisation Wnt, selon les analyses luciférase.** La figure montre les *fold changes* de l'activité luciférase normalisés par les mesures de  $\beta$ -galactosidase, réalisés dans des cellules HEK293 transfectées. La transfection implique l'utilisation d'un plasmide exprimant *DYRK1B* WT ou un des variant P/LP de *DYRK1B*, ainsi que du plasmide TOPflash (rapporteur TCF). La réponse a été mesurée pour des concentrations croissantes de WNT3A (0, 10, 30, et 100 ng/mL), par rapport à l'activité basale observée avec *DYRK1B* WT. Les variants P/LP sont représentés en bleu et les variants P/LP-full en rouge. Les contrôles positif et négatif, c'est à dire WT et Y271/273F sont représentés respectivement en gris et orange. Les données sont la moyenne  $\pm$  SEM des *fold changes* de quatre expériences différentes réalisées en triplicat technique. L'effet de chaque variant de *DYRK1B* a été analysé en utilisant un modèle de régression linéaire (l'*estimates*  $[\beta]$  et la p-valeur sont indiquées à droite des graphiques) et confirmé avec le modèle ANOVA ( $***P < 0,001$  versus WT). Le variant G371R<sup>a</sup> correspond à la substitution nucléotidique c.1111G>A ; le variant G371R<sup>b</sup> correspond à la substitution nucléotidique c.1111G>C. *WT*, wild type.

#### 2.3.2.2. Analyse de l'activité kinase des variants P/LP-full sur la CCND1

Afin d'évaluer l'effet des variants P/LP-full de *DYRK1B* sur une autre cible protéique, nous avons évalué l'activité kinase des variants P/LP-full sur la phosphorylation de la protéine CCND1 par western blotting. En comparaison avec la condition WT, cinq des six variants P/LP-full (codant p.G120V, p.H179L, p.D259N, p.P282L et p.R358\*) ont montré une diminution significative de l'expression de la CCND1 phosphorylée, suggérant une réduction de l'activité kinase sur la CCND1 (**Figure 19** et **Annexe 2, Figure S9**).

#### 2.3.2.3. Analyse de l'effet des variants sur la G6PC

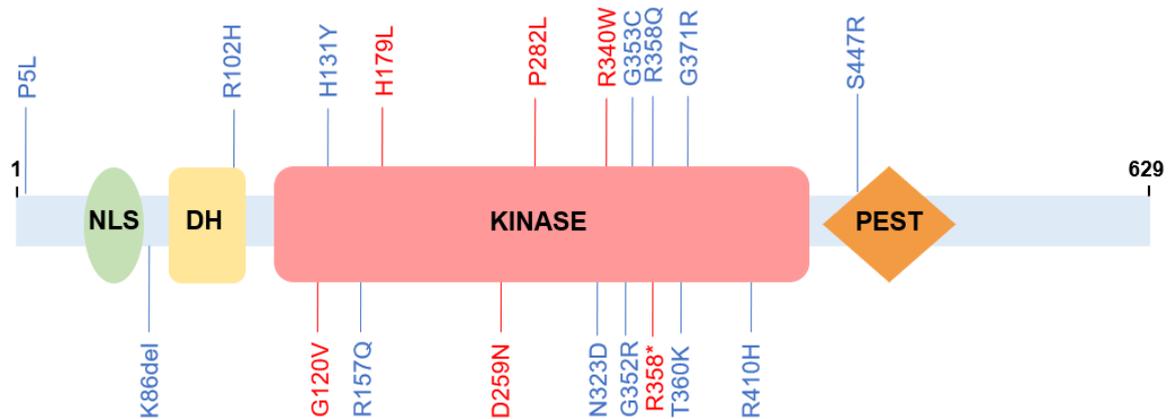
Nous avons également évalué l'effet des variants P/LP-full de *DYRK1B* sur la G6PC, de façon similaire à ce qui a été décrit dans l'étude initiale sur *DYRK1B*<sup>134</sup>. Pour cela, les cellules HepG2, qui expriment constitutivement la G6PC, ont été transfectées avec le plasmide exprimant *DYRK1B* WT, ou une des mutations de *DYRK1B* p.Y271/273F, p.R102C (qui est la mutation identifiée dans l'étude initiale<sup>134</sup>) ou p.R358\*. Nous avons ensuite mesuré l'expression du gène *G6PC* par RT-qPCR ainsi que l'expression protéique par western blotting (**Annexe 2, Figure S10**). Contrairement à ce qui avait été obtenu dans l'étude, où la surexpression de *DYRK1B* augmentait l'expression de la *G6PC*, l'expression génique et protéique de G6PC restait stable entre toutes les conditions. Nous n'avons donc pas appliqué ce test *in vitro* aux variants P/LP.

**A****B**

**Figure 19 : Effet des variants P/LP-full de *DYRK1B* sur la phosphorylation de CCND1 par western blotting.** La figure montre l'expression protéique de p-CCND1 dans des cellules HEK293 transfectées. Cette transfection implique l'utilisation d'un vecteur vide [VV] ou d'un plasmide *DYRK1B* WT ou d'un plasmide portant un variant P/LP-full de *DYRK1B* avec un plasmide codant CCND1. (A) L'expression protéique de DYRK1B, CCND1, p-CCND1 et de la  $\beta$ -actine a été analysée par western blotting. (B) La quantification de l'expression protéique a été réalisée en divisant les valeurs d'intensité de p-CCND1 par les valeurs d'intensité de CCND1, normalisée par la  $\beta$ -actine, par rapport à l'activité basale de DYRK1B WT. Les variants P/LP-full sont représentés en rouge. Les conditions contrôles VV, WT et Y271/273F sont représentées en vert, gris et orange, respectivement. Les données sont la moyenne  $\pm$  SEM des *fold changes* provenant de quatre expériences indépendantes (les images de western blotting sont disponibles dans l'Annexe, Figure S9). L'effet de chaque variant de *DYRK1B* (*versus* la condition WT) a été analysé en utilisant un test de Mann-Whitney ( $*P < 0.05$  *versus* WT). VV, vecteur vide; WT, wild-type; CCND1, cycline D1; p-CCND1, CCND1 phosphorylée.

#### 2.3.2.4. Localisation des variants P/LP sur la protéine DYRK1B

En regardant la localisation des variants rares identifiés dans cette étude au niveau des différents domaines de DYRK1B, nous avons observé que tous les variants P/LP-full étaient situés dans le domaine kinase de DYRK1B, un domaine hautement conservé (**Figure 20**).



**Figure 20 : Localisation des variants P/LP et P/LP-full de *DYRK1B*.** La figure montre les différents domaines et motifs de la protéine DYRK1B. Les variants bleus représentent les variants P/LP qui ont été définis par nos analyses *in vitro* et les critères ACMG. Les variants rouges représentent les variants P/LP-full qui sont des variants P/LP avec un effet totalement inhibiteur sur l'activité de DYRK1B, selon nos analyses *in vitro* et les critères ACMG. *NLS*, nucleus localization signals; *DH*, domaine d'homologie DYRK. *PEST*, motif riche en résidus proline, acide glutamique, sérine et thréonine.

#### 2.3.3. Effet des variants de *DYRK1B* sur différents traits métaboliques selon leur pathogénicité

Avant les analyses fonctionnelles, l'agrégation de tous les variants rares de *DYRK1B* n'était pas significativement associée à l'adiposité, au risque de DT2 ou à une hyperglycémie à jeun (**Tableau 12**). Après avoir réalisé les tests fonctionnels et la classification par les critères ACMG, nous avons évalué l'effet des variants P/LP et P/LP-full de *DYRK1B* sur l'adiposité et le DT2 en utilisant la méthode MiST ajustée au sexe, à l'âge et l'origine ethnique (**Tableau 12**).

**Tableau 12 : Effet des différents clusters des variants rares de *DYRK1B* sur l'homéostasie métabolique chez les adultes provenant de l'étude RaDiO.**

Les associations entre les agrégations des variants rares de *DYRK1B* en clusters et les traits métaboliques ont été analysées en utilisant la méthode MiST ajustée à l'âge, le sexe et l'origine ethnique. **IMC**, indice de masse corporelle ; **P/LP**, variants pathogènes/probablement pathogènes ; **P/LP-full**, variants pathogènes/probablement pathogènes avec un effet totalement inhibiteur ; **n**, nombre de participants ; **IC**, intervalle de confiance ;  **$\pi$** , effet moyen ; **OR**, odds ratio ; **SD**, déviation standard ; **SE**, erreur standard ; **DT2**, diabète de type 2 ;  **$\tau$**  : effet hétérogène

Trait	n	Cluster	%, moyenne $\pm$ SD du trait chez les porteurs	%, moyenne $\pm$ SD du trait chez les non-porteurs	OR [IC]	$\pi \pm SE$	$P_{\tau}$	$P_{\tau}$	$P_{overall}$
IMC (kg m <sup>-2</sup> )	7260	P/LP-full	36 $\pm$ 15	27 $\pm$ 6,5	-	8,0 $\pm$ 3,2	0,013	0,016	0,0020
		P/LP-all	31 $\pm$ 9,1	27 $\pm$ 6,5	-	4,2 $\pm$ 1,6	0,0084	0,20	0,013
		All	28 $\pm$ 7,1	27 $\pm$ 6,5	-	0,025 $\pm$ 0,28	0,93	0,99	0,99
Obésité ( $\geq 30$ kg m <sup>-2</sup> )	4401	P/LP-full	4/5 (80%)	1522/4396 (35%)	7,9 [1,2-155]	-	0,028	0,28	0,046
		P/LP-all	9/14 (64%)	1517/4387 (35%)	3,3 [1,2-10]	-	0,016	0,40	0,039
		All	22/50 (44%)	1504/4351 (35%)	1,1 [0,94-1,4]	-	0,17	0,41	0,26
DT2	6356	P/LP-full	4/6 (67%)	2175/6350 (34%)	4,8 [0,85-37]	-	0,066	0,38	0,12
		P/LP-all	6/16 (38%)	2173/6340 (34%)	1,1 [0,43-3,0]	-	0,78	0,38	0,65
		All	25/78 (32%)	2154/6278 (34%)	0,94 [0,79-1,1]	-	0,45	0,68	0,67
Glycémie à jeun (mmol l <sup>-1</sup> )	6906	P/LP-full	9,4 $\pm$ 5,9	6,2 $\pm$ 2,2	-	2,9 $\pm$ 1,0	0,0051	5,8 $\times 10^{-5}$	4,7 $\times 10^{-6}$
		P/LP-all	6,5 $\pm$ 3,6	6,2 $\pm$ 2,2	-	0,39 $\pm$ 0,52	0,45	0,0032	0,011
		All	5,9 $\pm$ 2,2	6,2 $\pm$ 2,2	-	-0,16 $\pm$ 0,082	0,047	0,84	0,17

### 2.3.3.1. Effets des variants rares P/LP de *DYRK1B* sur l'adiposité

Chez les adultes de l'étude RaDiO, nous avons constaté que l'agrégation des variants P/LP était significativement associée à un IMC plus élevé ( $31 \pm 9,1$  kg/m<sup>2</sup> chez les porteurs *versus*  $27 \pm 6,5$  kg/m<sup>2</sup> chez les contrôles ;  $P_{\pi} = 0,0084$  avec un effet [ $\pi$ ] de  $4,2 \pm 1,6$ ; **Tableau 12**) et était significativement associée à un risque accru de développer une obésité (9 sur 14 [64%] individus avec une obésité parmi les porteurs *versus* 35% des individus avec une obésité parmi les non-porteurs;  $P_{\pi} = 0,016$  avec un OR de 3,3; **Tableau 12**). De plus, il est important de noter que l'agrégation des variants P/LP-full avait un effet plus fort sur l'IMC que les variants P/LP ( $36 \pm 15$  kg/m<sup>2</sup> chez les porteurs *versus*  $27 \pm 6,5$  kg/m<sup>2</sup> chez les non-porteurs ;  $P_{\pi} = 0,013$  avec un effet [ $\pi$ ] de  $8,0 \pm 3,2$ ), malgré une légère hétérogénéité du cluster ( $P_{\tau} = 0,016$ ; **Tableau 12**). De même, l'effet de l'agrégation des variants P/LP-full de *DYRK1B* sur le risque d'obésité était plus important que celui des variants P/LP (4 des 5 [80%] individus avec une obésité parmi les porteurs *versus* 35% d'individus avec une obésité parmi les non-porteurs ;  $P_{\pi} = 0,028$  avec un OR de 7,9 ; **Tableau 12**). Parmi les 2085 enfants et adolescents provenant de l'étude RaDiO, seuls trois enfants étaient porteurs de variants P/LP, sans aucun porteur de variants P/LP-full, ce qui a considérablement réduit la puissance statistique de notre étude. Néanmoins, deux des trois enfants porteurs de mutations P/LP de *DYRK1B* présentaient une obésité.

### 2.3.3.2. Effets des variants rares P/LP de *DYRK1B* sur le DT2

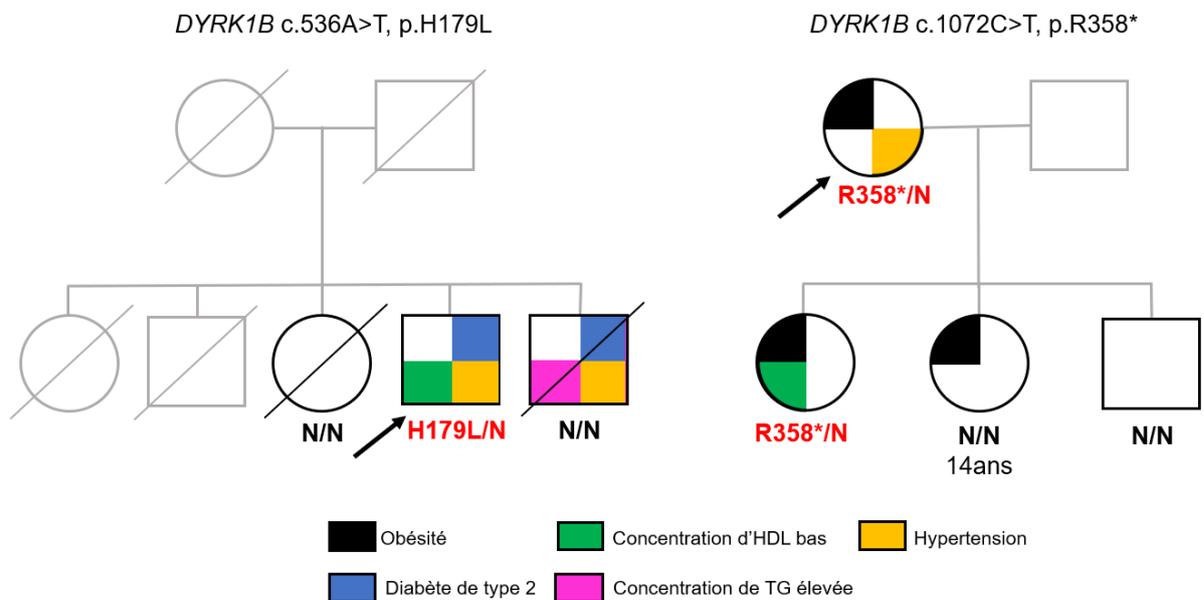
Nous avons observé une tendance d'association entre les variants P/LP-full de *DYRK1B* et le risque de DT2 (4 sur 6 [67%] individus atteints de DT2 parmi les porteurs *versus* 34% des individus atteints de DT2 parmi les non-porteurs ;  $P_{\pi} = 0,066$  avec un OR de 4,8, **Tableau 12**) ; alors que les variants P/LP n'avaient aucun effet (**Tableau 12**). Afin de répliquer les résultats obtenus, nous avons ensuite analysé l'association entre les variants P/LP-full de *DYRK1B* (transcrit NM\_004714.3 [ENST00000323039]) et le risque de DT2 dans le *Type 2 Diabetes Knowledge Portal* (à l'aide de l'outil interactif d'association génétique)<sup>209</sup> (**Tableau 13**). Parmi les 43 125 individus provenant de la 52K et les 44 083 individus provenant de la TOPMed, nous avons identifié deux et sept variants P/LP-full dans la 52K et la TOPMed, respectivement (c'est-à-dire des variants nuls avec une MAF très faible dans le serveur gnomAD) (**Annexe 1, Tableau S16**). Aucun des variants P/LP ou P/LP-full identifiés dans notre étude n'a été retrouvé dans la TOPMed ou dans la 52K. Alors que les variants P/LP-full dans la 52K n'ont pas montré d'effet significatif sur le risque de DT2 en raison d'une faible puissance statistique ( $P = 0,13$  avec un OR de 1,7 ; *collapsing burden test*), les variants P/LP-full dans TOPMed ont montré une association significative avec le DT2 ( $P = 0,030$  avec un OR de 4,2 ; *collapsing burden test*)

(Tableau 13). De plus, grâce à une méta-analyse des études RaDiO, TOPMed et 52K, nous avons constaté un enrichissement des variants P/LP-full de *DYRK1B* parmi les cas de DT2 ( $P = 0,0040$  avec un OR de 2,9, IC 95% : 1,4-6,0 ; *random effects model*) (Tableau 13). Par ailleurs, nous avons constaté que les variants P/LP-full de *DYRK1B* étaient significativement associés à des taux de glucose à jeun plus élevés ( $9,4 \pm 5,9$  mmol/l chez les porteurs *versus*  $6,2 \pm 2,2$  mmol/l chez les non-porteurs ;  $P_{\pi} = 0,0051$  avec un effet [ $\pi$ ] de  $2,9 \pm 1,0$  ; Tableau 12), bien que le cluster présente une hétérogénéité significative ( $P_{\tau} = 5,8 \times 10^{-5}$  ; Tableau 12).

**Tableau 13 : Effet des variants P/LP-full de *DYRK1B* sur le risque de DT2 par une méta-analyse des études RaDiO, TOPMed et 52K.** *n*, nombre de participants ; *IC*, intervalle de confiance ; *OR*, odds ratio

Etude	n	Variants P/LP-full	OR [IC]	P
RaDiO	6356	6	4,8 [0,85-36,5]	0,066
52K	43 125	2	1,7 [1-2,4]	0,13
TOPMed	44 083	7	4,2 [2,9-5,5]	0,030
<b>Méta-analyse</b>			2,9 [1,4-6,0]	0,0040

Enfin, nous avons eu l'occasion d'analyser la coségrégation des variants P/LP-full (codant pour p.H179L et p.R358\*) avec l'obésité, le DT2 et d'autres traits métaboliques dans deux familles contenant des cas index issus de l'étude RaDiO (Figure 21). Le cas index portant le variant codant p.R358\*, qui présentait une obésité sévère (IMC de 39 kg/m<sup>2</sup>), a transmis la mutation à sa fille qui présentait également une obésité sévère (IMC de 39 kg/m<sup>2</sup>). Ses deux fils ne portaient pas la mutation ; l'un avait un poids normal et l'autre était atteint d'obésité (Figure 21). La deuxième famille présentait des limites plus importantes, le variant codant p.H179L étant trouvé exclusivement chez le cas index (Figure 21).



**Figure 21 : Coségrégation des variants p.H179L et p.R358\* avec des traits métaboliques dans deux familles.** La figure montre la coségrégation de deux variants P/LP-full (p.H179L et p.R358\*) avec des traits métaboliques dans deux familles. La flèche indique l'individu séquencé dans l'étude RaDiO. Les membres des familles ont été séquencés par séquençage Sanger. L'obésité a été définie comme un IMC  $\geq 30$  kg/m<sup>2</sup>, DT2 comme une glycémie à jeun  $\geq 7,0$  mmol/l et/ou l'utilisation d'un traitement contre l'hyperglycémie, un niveau bas de HDL comme  $\leq 1,04$  mmol/l pour les hommes et  $\leq 1,30$  mmol/l pour les femmes, des hauts niveaux de TG comme  $\geq 1,70$  mmol/l et l'hypertension par une pression systolique  $\geq 130$  mmHg ou par une pression diastolique  $\geq 85$  mmHg. *NN*, *wild-type* ; *HDL*, lipoprotéine à haute densité ; *TG*, triglycéride.

## 2.4. Discussion

Nos résultats ont mis en évidence que les variants P/LP de *DYRK1B* qui entraînent une perte totale de fonction causent une forme monogénique d'obésité (avec une moyenne d'IMC de 36 kg/m<sup>2</sup>) associée au DT2 dans une large population. Il est important de noter que, sur base des analyses *in vitro*, nous avons démontré que les variants P/LP de *DYRK1B* qui entraînent une perte totale de la fonction de la protéine, ont un impact beaucoup plus important sur l'obésité et le DT2 par rapport aux mutations P/LP de *DYRK1B*. Ces résultats soulignent l'importance des analyses fonctionnelles *in vitro* pour distinguer les variants P/LP qui sont complètement perte de fonction des variants P/LP de *DYRK1B*. L'identification des individus porteurs de ces variants P/LP complètement perte de fonction, est cruciale car les progrès futurs dans les interventions pharmacologiques visant à moduler la fonction métabolique de *DYRK1B* pourraient permettre une médecine de précision chez ces patients.

Il est intéressant de noter que toutes les mutations P/LP perte de fonction totale de *DYRK1B* identifiées dans cette étude étaient situées dans le domaine catalytique (**Figure 18**), suggérant un impact direct sur l'activité kinase. En revanche, la majorité des mutations de *DYRK1B* précédemment identifiées dans d'autres études se trouvaient dans le domaine DH ou dans le domaine N-terminal de *DYRK1B*. Les effets de ces mutations peuvent s'expliquer par des altérations structurales qui induisent une instabilité conformationnelle et augmentent le repliement incorrect de *DYRK1B*, conduisant à l'agrégation intracellulaire d'une protéine non fonctionnelle<sup>197,205</sup>

A noter que le variant codant p.K68Q, que nous avons classé comme neutre, a été identifié comme pathogène dans une étude antérieure<sup>205</sup>. Cependant, dans notre étude, le porteur de ce variant était un homme de 45 ans qui ne présentait pas de surpoids ni de DT2.

Notre étude a posé la question sur l'impact réel des variants de *DYRK1B* causant l'AOMS, en particulier s'ils entraînent un gain ou une perte de fonction sur l'activité de *DYRK1B*. La première étude a décrit des variants *DYRK1B* (qui étaient P/LP) dans l'AOMS comme étant des variants gain de fonction<sup>134</sup>. Néanmoins, d'autres études ont démontré plutôt un effet perte de fonction des variants *DYRK1B*, se traduisant par une activité kinase réduite, une diminution de l'expression de *DYRK1B*, une augmentation du repliement incorrect, de la dégradation et de l'agrégation de *DYRK1B*<sup>197,204,206</sup>. Ces hypothèses qui suggèrent que les variants P/LP soient perte de fonction plutôt que gain de fonction, sont en accord avec nos résultats. En effet, nous avons observé que nos variants P/LP-full avaient un effet similaire sur la signalisation Wnt que

le contrôle négatif où le domaine kinase était désactivé (p.Y271/273F) ainsi que le variant non-sens (p.R358\*), qui entraîne intrinsèquement une perte de fonction. De plus, les deux variants codant pour p.Y271/273F et p.R358\* ont fortement réduit l'expression de la CCND1 phosphorylée, suggérant un effet perte de fonction. Il convient cependant de noter que *DYRK1B* peut également agir comme un facteur de transcription indépendamment de son activité kinase<sup>134</sup>.

Le lien entre les mutations pathogènes dans *DYRK1B* et l'obésité et/ou le DT2 est mal compris. En effet, bien que *DYRK1B* ait principalement été étudié dans le contexte du cancer et de la myogenèse, son rôle dans les troubles métaboliques a été peu exploré. *DYRK1B* pourrait favoriser l'adipogenèse pendant l'embryogenèse en inhibant les voies de signalisation Hh et Wnt et en stimulant les facteurs de transcription adipogéniques tels que *CCAAT/enhancer-binding protein  $\alpha$*  (C/EBP $\alpha$ ) et *peroxysome proliferator-activated receptor  $\gamma$*  (PPAR $\gamma$ ), comme démontré dans les préadipocytes de souris 3T3-L1<sup>134</sup>. De plus, *DYRK1B* pourrait contribuer à la promotion de la G6PC, indépendamment de son activité kinase, régulant ainsi l'homéostasie du glucose<sup>134,211</sup>. Cependant, il convient de noter que nos expériences n'ont pas réussi à répliquer cet effet. Par ailleurs, l'effet inhibiteur de *DYRK1B* sur la voie RAS-RAF-MEK, mis en évidence dans le cancer<sup>212</sup>, pourrait également être important pour la régulation de l'absorption de glucose et de la glycolyse, notamment car des mutations perte de fonction identifiées dans *KSR2*, qui est impliqué également dans cette même voie, ont été liées à une obésité monogénique avec une insulino-résistance<sup>114</sup>.

De plus, *DYRK1B* partage 85% d'identité de séquence avec son paralogue *DYRK1A*, connu pour son rôle dans le développement neuronal. Les mutations de perte de fonction de *DYRK1A* ont été associées à un syndrome de retard mental autosomique dominant 7<sup>213</sup>. L'expression de *DYRK1B* dans des régions cérébrales, notamment l'hypophyse et le cervelet, suggère un rôle potentiel dans le développement neuronal<sup>198</sup>. De plus, l'implication de *DYRK1A* dans la régulation centrale de l'apport alimentaire par le biais de la voie Sir-2-FOXO-NPY chez *Drosophila melanogaster* et la souris implique la possibilité d'une fonction comparable pour *DYRK1B*<sup>214,215</sup>. Enfin, *DYRK1B* est également exprimé dans les cellules bêta pancréatiques, où il a été démontré que l'inhibition à la fois de *DYRK1A* et de *DYRK1B* est nécessaire pour atteindre les taux de prolifération des cellules bêta maximaux. Cependant, la manière dont *DYRK1B* inhibe la prolifération des cellules bêta ou d'autres mécanismes est inconnue<sup>216</sup>.

La principale limitation de notre étude était l'absence de données sur le statut cardiaque des porteurs de l'étude RaDiO. Par conséquent, nous ne disposons pas d'informations sur le fait

que les variants P/LP-full étaient associés à un risque élevé de développer une maladie coronarienne précoce. De plus, nous nous sommes concentrés uniquement sur des voies de signalisation spécifiques impliquant DYRK1B, malgré son expression ubiquitaire. Par conséquent, l'effet potentiel sur d'autres protéines cibles ne peut être exclu et nécessite des investigations supplémentaires.

# **DISCUSSION GENERALE ET PERSPECTIVES**

En combinant le séquençage NGS avec des manipulations fonctionnelles *in vitro* des variants génétiques détectés (sorte de génétique fonctionnelle) nous avons mis en évidence l'impact des variants rares situés dans les gènes *PCSK1* et *DYRK1B* sur l'obésité et les traits phénotypiques associés. Cette approche combinée nous a permis d'identifier de nouvelles mutations rares et d'établir les spécificités phénotypiques des individus porteurs de ces mutations, contribuant ainsi à une meilleure compréhension des liens mécanistiques de *PCSK1* et *DYRK1B* dans l'obésité. En outre, nous avons démontré l'importance de discriminer les variants présentant une perte de fonction totale de ceux ayant un effet plus modéré grâce à nos analyses fonctionnelles *in vitro*. Dans le cas du projet sur *PCSK1*, seules les mutations nulles et identifiées comme entraînant une perte de fonction totale causaient une obésité monogénique alors que les variants rares avec un effet partiel ou neutre sur l'activité enzymatique de PC1/3 ne contribuaient pas à l'obésité. De manière similaire, dans l'étude sur *DYRK1B*, les mutations P/LP nulles ou ayant un effet complètement inhibiteur sur la protéine avaient un effet plus important sur l'obésité monogénique et le DT2 que les mutations P/LP qui avaient un effet plus modéré sur l'adiposité et aucune incidence sur l'homéostasie glucidique. Il est important de noter que les tests de prédiction *in silico*, tels que le programme REVEL (regroupant 13 logiciels de prédiction *in silico*), ne sont pas parvenus à déterminer précisément les variants non synonymes délétères. En effet, les prédictions des variants rares hétérozygotes de *PCSK1* dans notre étude montraient à la fois une faible sensibilité (7% de faux négatifs) et une faible spécificité (25% de faux positifs). En ce qui concerne les variants rares de *DYRK1B*, REVEL a montré une bonne spécificité (0% de faux positifs) mais il manquait de sensibilité, en prédisant correctement seulement quatre des cinq variants non synonymes P/LP-full (2% de faux négatifs). Ces résultats mettent en évidence l'importance des analyses *in vitro* pour évaluer la pathogénicité des variants rares non synonymes. Ces découvertes ont des implications directes pour l'amélioration du diagnostic génétique et de la prise en charge de l'obésité monogénique, et peuvent ouvrir la voie à l'identification et au développement de nouvelles cibles thérapeutiques.

Le point essentiel de ces deux études était de déterminer les différents groupes de pathogénicité des variants rares en partant de cohortes à grande échelle, afin de discriminer objectivement les variants avec une perte de fonction totale des autres variants. Habituellement, l'évaluation de la pathogénicité des variants rares s'appuie sur les critères ACMG. Cependant, dans le cas du projet *PCSK1*, de nombreux variants rares hétérozygotes étaient considérés comme pathogènes dans la littérature (critère PS1), même s'ils n'affectaient que partiellement l'activité

enzymatique. De plus, certaines études ont utilisé un vecteur plasmidique qui portait une mutation activatrice, ce qui a biaisé les résultats des analyses enzymatiques. Ainsi, nous avons choisi de ne pas appliquer les critères ACMG aux variants hétérozygotes de *PCSK1*. A la place, nous avons opté pour une approche basée sur des tests *in vitro* de l'activité enzymatique analysés par la méthode du clustering *k-means* par *machine learning*, validée au moyen de différents critères de qualité, afin de déterminer objectivement cinq groupes de pathogénicité. Ces résultats suggèrent qu'il est nécessaire de reclasser les variants hétérozygotes de *PCSK1* étudiés précédemment dans la littérature et actuellement classés comme pathogènes. Concernant le projet *DYRK1B*, le volume de données provenant des tests fonctionnels n'était pas suffisant pour utiliser la méthode de clustering. Par conséquent, nous avons utilisé en parallèle les critères ACMG et les résultats des tests fonctionnels analysés à l'aide de deux méthodes statistiques différentes (suivant l'intercept [ $\beta$ ]) afin de déterminer les variants avec une perte de fonction totale (P/LP-full). Nous avons donc utilisé deux approches différentes pour discriminer les variants avec une perte de fonction totale des autres variants. Il serait intéressant de développer une méthode statistique objective, qui puisse être adaptée à différents tests fonctionnels afin de différencier précisément les variants fortement perte de fonction des autres variants.

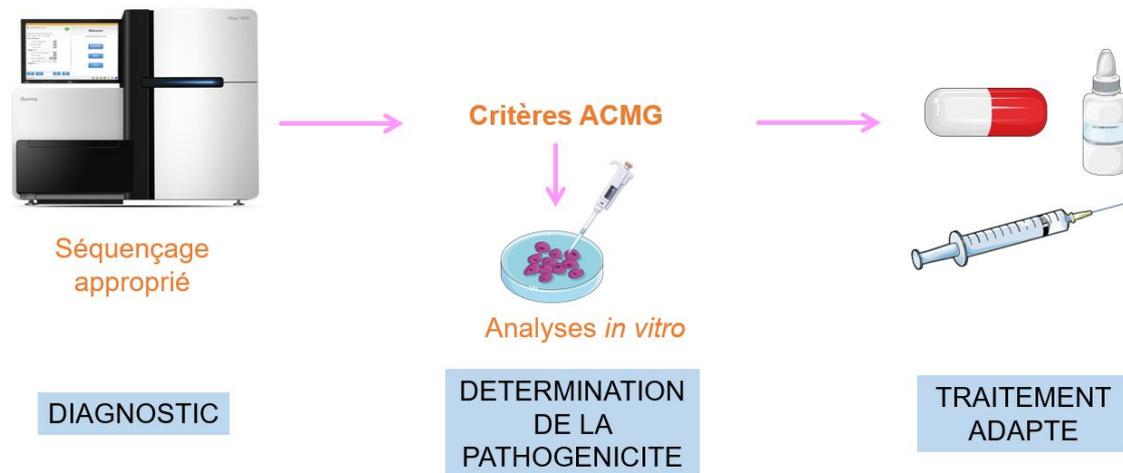
Ce travail présente également des limites liées à l'utilisation de cohortes. Tout d'abord, ces deux études se sont appuyées sur une seule cohorte d'origine caucasienne, ce qui restreint l'applicabilité des résultats aux individus de la même ethnie. Des études sur *PCSK1* ou *DYRK1B* dans d'autres ethnies peuvent apporter des nouveaux éléments sur leur impact sur l'obésité car les populations ne partagent pas forcément les mêmes variants, la même prédisposition et le même environnement. De plus, le projet RaDiO contient des cohortes assez anciennes, et une partie des patients recrutés est malheureusement décédée. Nous avons constaté que l'ADN de certains échantillons était partiellement dégradé lors du séquençage Sanger effectué pour les analyses de coségrégation. De plus, certaines données phénotypiques potentiellement pertinentes dans le cadre de ces projets manquaient : une évaluation de l'hyperphagie des patients, les concentrations d'ACTH et de proinsuline ainsi que des informations sur d'éventuels troubles intestinaux pour le projet *PCSK1*, et des données concernant le statut cardiaque pour le projet *DYRK1B*. Enfin, les études de cohortes peuvent présenter un biais de sélection lié au recrutement des participants qui peuvent ne pas être représentatifs de la population générale. En effet, les personnes qui décident de participer à une étude de cohorte peuvent être plus conscientes de leur santé ou présenter des caractéristiques particulières les

distinguant de la population générale. De plus, la cohorte d'enfants incluse dans notre étude présentait une taille d'échantillon relativement modeste pour ce type d'étude, ce qui peut s'expliquer par la difficulté de recruter des enfants et contribuer au manque de puissance statistique observée dans nos résultats. Malgré ces limitations, les études de cohorte restent essentielles pour la recherche en santé publique, fournissant des informations précieuses et contribuant à améliorer notre compréhension des maladies de manière générale.

Il est également important de souligner l'importance de sélectionner un modèle cellulaire approprié pour la réalisation des analyses *in vitro*, tout en reconnaissant que les lignées cellulaires cultivées présentent des limitations en termes de représentation de la réalité physiologique. Par exemple, notre étude sur *PCSK1* peut présenter certaines limitations en raison du fait que les analyses *in vitro* ont été effectuées sur une lignée cellulaire non neuronale. Cependant, les cellules HEK293 présentaient l'avantage d'être des cellules humaines qui n'expriment pas constitutivement *PCSK1* en plus d'avoir été utilisées par plusieurs études pour étudier la voie de signalisation MC4R<sup>96,175,217</sup>, ce qui a validé notre choix de ce modèle. Il serait intéressant de répliquer notre analyse fonctionnelle dans une lignée cellulaire neuronale, telle que les N2a, afin de se rapprocher des conditions physiologiques réelles. La tâche est plus complexe pour les gènes à expression ubiquitaire, comme *DYRK1B*. Nous avons également choisi des cellules HEK293 en raison de leur facilité de transfection et de leur réponse à la voie Wnt, et nous avons validé notre analyse *in vitro* grâce à différents contrôles. Cependant, il serait intéressant de reproduire cette analyse dans d'autres types cellulaires pour approfondir notre compréhension du rôle biologique de *DYRK1B*.

Les variants rares sont souvent plus difficiles à détecter que les SNP fréquents, mais leur identification peut avoir un impact majeur sur la prise en charge des patients atteints d'obésité monogénique. En effet, les avancées dans l'étude de l'impact de ces variants génétiques rares ont permis de mieux comprendre les liens mécanistiques entre les gènes porteurs de mutations et la régulation du poids corporel. Ces découvertes permettent d'ouvrir de nouvelles perspectives en matière de prévention et de traitement de l'obésité monogénique. Le processus de prise en charge de l'obésité monogénique implique un diagnostic génétique, une évaluation de la pathogénicité des variants identifiés et la possibilité d'un traitement personnalisé, introduisant ainsi la médecine de précision pour les patients atteints d'obésité monogénique (**Figure 22**). L'étude de l'impact des variants rares sur l'obésité monogénique est un domaine de recherche avec un fort potentiel pour améliorer la compréhension et la gestion de l'obésité. Le développement de la médecine de précision a pu progresser grâce aux avancées dans le

domaine de la génétique et au développement des outils de séquençage, jouant un rôle essentiel dans les perspectives de prévention et de traitement.



**Figure 22 : Les différentes étapes de la mise en place de la médecine de précision chez les patients atteints d'obésité monogénique.**

L'identification de la mutation ou de l'anomalie génétique susceptible de causer l'obésité est la première étape de la prise en charge de l'obésité monogénique. Le diagnostic génétique est envisagé chez les patients présentant des caractéristiques typiques de l'obésité monogénique, telles qu'une obésité précoce (avant 6 ans) accompagnée d'hyperphagie la plupart du temps. Il peut également être systématique dans les familles touchées par une obésité héréditaire et particulièrement dans les familles consanguines. Le diagnostic génétique s'est considérablement amélioré grâce aux développements des outils de séquençage. La technique la plus couramment utilisée à des fins thérapeutiques est celle du NGS effectué sur un panel de gènes connus qui ont été sélectionnés pour être la cause potentielle de l'obésité monogénique d'après la littérature ou d'autres patients phénotypiquement semblables à celui qui est diagnostiqué. Cependant, ce séquençage ciblé est limité par l'évolution constante de la compréhension scientifique, car de nouveaux gènes impliqués dans l'obésité monogénique sont régulièrement découverts, rendant les kits de panels de gènes obsolètes. Le WES permet de surmonter ces limitations en fournissant le séquençage de l'exome entier sans avoir d'a priori. Néanmoins, il génère une quantité importante de données à traiter nécessitant un équipement informatique performant. De plus, le WES est peu sensible à la détection des CNV. Quant au WGS, il est pour l'instant limité au domaine de la recherche en raison de son coût élevé et de la complexité de l'analyse des vastes quantités de données générées. Cependant, le WGS pourrait devenir utile en clinique pour détecter les CNV et les réarrangements chromosomiques

susceptibles de provoquer une obésité monogénique comme c'est le cas pour la délétion du chromosome 16p11.2 ou la duplication en tandem dans le gène *ASIP* qui ne peuvent pas être détectées par les méthodes de séquençage de routine traditionnel. Pour améliorer la détection des CNV et des mutations ponctuelles en une seule étape, une technique de séquençage nommée CoDE-SEQ a été développée par le laboratoire. Elle est basée sur un WES amélioré enrichi en sondes spécifiques couvrant l'ensemble des chromosomes, offrant une approche plus rapide et moins coûteuse que le WGS<sup>218</sup>.

Une fois la cause génétique identifiée par le diagnostic, il est nécessaire d'évaluer si le variant génétique est responsable de l'obésité. Si le variant est déjà décrit dans la littérature, l'interprétation de sa pathogénicité est relativement simple. Cependant, lorsqu'il s'agit d'un variant inconnu, il est nécessaire d'estimer sa pathogénicité en utilisant les critères ACMG (cf supra). Les résultats de ce travail ont mis en évidence l'importance des analyses *in vitro* pour évaluer la pathogénicité des variants rares non synonymes (critère PS3). Intégrer des tests fonctionnels *in vitro* dans le domaine du diagnostic génétique pourrait permettre de déterminer l'impact fonctionnel des nouveaux variants rares<sup>219</sup>. Cette information serait essentielle pour proposer un traitement efficace en fonction de la pathogénicité de la mutation. Pour atteindre cet objectif, il est nécessaire de mettre au point des tests fonctionnels *in vitro* standardisés qui soient rapides, simples et efficaces. Trois analyses *in vitro* ont été proposées par Sha et al, 2022 afin de caractériser fonctionnellement les variants rares situés dans *LEPR*, *POMC* et *PCSK1*, en se basant sur ce qui a été établi et validé dans les études précédentes : une analyse de la signalisation STAT3 pour les variants *LEPR*, de l'activité enzymatique pour *PCSK1* et de la signalisation MC4R pour *POMC*<sup>217</sup>. De plus, l'utilisation en parallèle d'une technique de clustering standardisée avec des analyses permettrait de discriminer objectivement les variants avec une perte de fonction totale de ceux avec un effet partiel. Ces méthodes, bien que nécessitant des adaptations spécifiques pour chaque gène étudié, pourraient devenir un outil précieux pour déterminer précisément la pathogénicité des variants rares en routine, améliorant le diagnostic ainsi que les stratégies de traitement personnalisés pour les patients atteints d'obésité monogénique. Afin de développer cette stratégie, il est impératif de comprendre comment un variant délétère peut mener à l'obésité. Cela souligne la nécessité de poursuivre la recherche sur les variants génétiques rares dans le contexte de l'obésité, tel que dans notre étude.

Le diagnostic génétique peut non seulement éclairer le pronostic du patient, mais aussi, dans certains cas, déterminer le traitement approprié en fonction de la génétique. Il permet d'identifier les individus présentant un risque plus élevé de développer une obésité, ce qui à son

tour permet la mise en place de mesures de prévention personnalisée afin de limiter le développement potentiel d'une obésité. L'identification des individus à risque dès leur plus jeune âge est essentielle pour des interventions précoces et des stratégies de gestion de poids personnalisées. C'est pourquoi le projet OBELISK, un large projet européen conduit par le laboratoire, vise à réduire de manière significative l'obésité infantile en Europe en développant une véritable médecine de précision. Il s'appuie sur des outils de prédiction de l'obésité dès la naissance, des programmes de prévention impliquant des familles, des professionnels de la santé, des travailleurs sociaux et des communautés médicales, ainsi que sur de nouveaux traitements médicamenteux<sup>220</sup>.

A ce jour, il existe deux traitements pour l'obésité monogénique adaptés à la génétique du patient : la leptine recombinante et le setmélanoïde. Même si ces thérapies ne concernent qu'une minorité de patients, elles sont disponibles et elles ont prouvé leur efficacité (cf supra). Elles représentent un premier pas vers une prise en charge médicamenteuse ciblée de l'obésité monogénique. Concernant l'obésité polygénique, le seul traitement efficace est la chirurgie bariatrique (cf supra). Le diagnostic génétique d'une obésité monogénique permet de préciser les indications pour la chirurgie bariatrique, car l'obésité monogénique est considérée comme une contre-indication. Ainsi, le génotypage systématique pourrait améliorer l'évaluation des indications de la chirurgie. Cela démontre comment la compréhension des mécanismes biologiques sous-jacents peut orienter le développement des molécules et des médicaments afin de rétablir les voies altérées dans les formes monogéniques d'obésité et comment la médecine de précision peut être un moyen efficace de prévenir et traiter l'obésité. Néanmoins, il reste d'importants obstacles dans la transition de la médecine conventionnelle à la médecine de précision dans le contexte de l'obésité monogénique. En effet, cela nécessiterait la mise en œuvre systématique du WES pour tous les individus soupçonnés de porter une mutation délétère, voire un dépistage standardisé à la naissance, une pratique qui n'est pas encore généralisée, en grande partie en raison des exigences de gestion des données et des préoccupations éthiques qui y sont liées<sup>39</sup>.

En outre, il est important de poursuivre le travail de la recherche afin d'identifier de nouveaux gènes associés à l'obésité. Actuellement, la plupart des gènes impliqués dans l'obésité monogénique sont situés dans la voie leptine-mélanocortine, ce qui a conduit au développement d'un traitement ciblant cette voie, le setmélanoïde. Cependant, en étudiant l'impact des variants rares situés dans *DYRK1B*, nous avons montré que d'autres gènes extérieurs à la voie leptine-mélanocortine pourraient également contribuer à l'obésité. Il est donc essentiel d'explorer

d'autres voies moléculaires susceptibles d'être impliquées dans l'obésité afin de découvrir de nouveaux gènes. Par exemple, l'effet de la voie de signalisation Raf-MEK-ERK qui comprend les gènes *KSR2* et *DYRK1B* demanderait davantage d'investigation dans le contexte de l'obésité. De nouvelles connaissances des mécanismes biologiques des voies impliquées dans l'obésité monogénique pourraient mener au développement des thérapies personnalisées qui ciblent spécifiquement ces voies perturbées par les mutations. Plus récemment, des mutations dans le gène *GLP1R* ayant un impact sur son expression à la surface cellulaire ont été mises en évidence comme associées à une augmentation de l'IMC et à un mauvais contrôle de la glycémie. Les patients porteurs de ces mutations spécifiques pourraient bénéficier d'un traitement personnalisé, à l'aide d'agonistes du GLP1R ou de modulateurs allostériques positifs (PAM) du GLP1R<sup>221</sup>.

Nous avons démontré que les variants rares hétérozygotes de *PCSK1* avec un effet partiellement délétère ne contribuent pas significativement à l'obésité monogénique. Cependant, cela n'exclut pas que les patients porteurs d'un variant avec un effet partiel (catégorie B ou C) ne pourraient pas également posséder une autre mutation favorisant l'obésité, et que l'obésité apparente soit due à l'effet cumulatif de plusieurs variants. A ce sujet, une obésité plus sévère est observée chez les individus porteurs de deux variants hétérozygotes dans la voie leptine-mélanocortine par rapport à ceux porteurs d'un seul variant hétérozygote dans cette même voie<sup>222</sup>. Par conséquent, il est important de noter qu'un seul diagnostic génétique ne peut exclure la possibilité d'un second diagnostic génétique lié à un effet cumulatif de plusieurs variants. En outre, les mutations liées à l'obésité monogénique sont influencées, au moins en partie, par la prédisposition polygénique des individus. Connaitre la susceptibilité génétique individuelle à l'obésité dès le plus jeune âge pourrait contribuer à prédire de manière plus précise quelles personnes sont les plus à risque à la prise de poids. Cette prédisposition polygénique est évaluée à l'aide du PRS qui est calculé à partir de larges études génétiques. Le PRS le plus récent repose sur une GWAS sur l'IMC impliquant des données sur plus de 2 millions de variants. Il a été constaté que l'IMC moyen des individus ayant un haut PRS était supérieur de 2,9 kg m<sup>-2</sup> et qu'ils avaient un risque d'obésité 4,2 fois plus élevé que les individus avec un PRS faible<sup>223</sup>. Cependant, malgré les fortes associations observées avec l'IMC et l'obésité dans la UK Biobank, la capacité prédictive du PRS-IMC est encore limitée : un PRS-IMC supérieur au 90ème percentile présente à la fois des faux positifs (43% des individus prédits pour développer une obésité l'ont effectivement développée) et des faux négatifs (parmi les individus développant réellement une obésité, seuls 19% sont correctement prédits<sup>39</sup>). Néanmoins, cette

capacité prédictive du PRS devrait s'améliorer à mesure que les GWAS augmentent en taille d'échantillon et que les algorithmes de calcul des scores deviennent plus précis. Le diagnostic génétique continuera probablement de s'améliorer à l'avenir grâce à des études génétiques de plus grande envergure et à des algorithmes informatiques améliorés qui prendront en compte les variants rares à effet fort, la prédisposition génétique et les interactions entre les variants. Les informations génétiques seront alors de plus en plus utilisées pour identifier les personnes à risque d'obésité et mettre en place des mesures de prévention et des traitements personnalisés<sup>39</sup>.

En résumé, l'étude des variants rares dans le contexte de l'obésité, en particulier l'obésité monogénique, offre de nombreuses opportunités pour améliorer notre compréhension et élaborer des stratégies de prévention et de traitement plus précises et personnalisées. Cependant, il est important de noter que l'obésité est multifactorielle, incluant des facteurs génétiques, environnementaux et comportementaux qui interagissent de manière complexe. Cela souligne la nécessité d'adopter une approche globale de la lutte contre l'obésité. L'intégration de diverses sources de données provenant de la génétique, l'épigénétique, du microbiome et des expositions environnementales spécifiques pourraient contribuer à identifier les jeunes patients à risque dans le contexte de l'obésité commune. La poursuite des recherches sur ces différents facteurs dans le domaine de l'obésité devrait amener à de nouvelles perspectives dans la compréhension de l'étiologie de l'obésité.



# **BIBLIOGRAPHIE**

1. Générique obésités de causes rares. *Haute Autorité de Santé* [https://www.has-sante.fr/jcms/p\\_3280217/fr/generique-obesites-de-causes-rares](https://www.has-sante.fr/jcms/p_3280217/fr/generique-obesites-de-causes-rares).
2. Obesity and overweight. *World Health Organization* <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight>.
3. BMI for Children and Teens. *Centers for Disease Control and Prevention* <https://www.cdc.gov/obesity/basics/childhood-defining.html> (2023).
4. Cole, T. J., Bellizzi, M. C., Flegal, K. M. & Dietz, W. H. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ* **320**, 1240–1243 (2000).
5. Cole, T. J. & Lobstein, T. Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. *Pediatric Obesity* **7**, 284–294 (2012).
6. Paley, C. A. & Johnson, M. I. Abdominal obesity and metabolic syndrome: exercise as medicine? *BMC Sports Sci Med Rehabil* **10**, 7 (2018).
7. Han, T. S., Sattar, N. & Lean, M. Assessment of obesity and its clinical implications. *BMJ* **333**, 695–698 (2006).
8. Duren, D. L. *et al.* Body Composition Methods: Comparisons and Interpretation. *J Diabetes Sci Technol* **2**, 1139–1146 (2008).
9. World Health Organization. World Obesity Day 2022 – Accelerating action to stop obesity. <https://www.who.int/news/item/04-03-2022-world-obesity-day-2022-accelerating-action-to-stop-obesity>.
10. Ranadive, S. A. & Vaisse, C. Lessons from Extreme Human Obesity: Monogenic Disorders. *Endocrinol Metab Clin North Am* **37**, 733–x (2008).
11. Chew, N. W. S. *et al.* The global burden of metabolic disease: Data from 2000 to 2019. *Cell Metabolism* **35**, 414–428.e3 (2023).
12. World Obesity Atlas 2023. *World Obesity Federation* <https://www.worldobesity.org/resources/resource-library/world-obesity-atlas-2023>.
13. GHO | By category | Prevalence of obesity among adults, BMI ≥ 30, age-standardized - Estimates by country. *WHO* <https://apps.who.int/gho/data/node.main.A900A?lang=en>.
14. Presentation maps. *World Obesity Federation Global Obesity Observatory* <https://data.worldobesity.org/maps/>.
15. Lin, X. & Li, H. Obesity: Epidemiology, Pathophysiology, and Therapeutics. *Front Endocrinol (Lausanne)* **12**, 706978 (2021).
16. Fontbonne, A. *et al.* Prevalence of Overweight and Obesity in France: The 2020 Obepi-Roche Study by the “Ligue Contre l’Obésité”. *J Clin Med* **12**, 925 (2023).
17. Lister, N. B. *et al.* Child and adolescent obesity. *Nat Rev Dis Primers* **9**, 1–19 (2023).
18. Djalalinia, S., Qorbani, M., Peykari, N. & Kelishadi, R. Health impacts of Obesity. *Pak J Med Sci* **31**, 239–242 (2015).
19. Sharma, V. *et al.* A systematic review and meta-analysis estimating the population prevalence of comorbidities in children and adolescents aged 5 to 18 years. *Obes Rev* **20**, 1341–1349 (2019).
20. Sherling, D. H., Perumareddi, P. & Hennekens, C. H. Metabolic Syndrome: Clinical and Policy Implications of the New Silent Killer. *J Cardiovasc Pharmacol Ther* **22**, 365–367 (2017).
21. Masson, E. Physiopathologie de l’obésité. *EM-Consulte* <https://www.em-consulte.com/article/75115/physiopathologie-de-l-obesite>.
22. Arora, S. & Anubhuti, null. Role of neuropeptides in appetite regulation and obesity--a review. *Neuropeptides* **40**, 375–401 (2006).
23. Maffei, M. *et al.* Leptin levels in human and rodent: Measurement of plasma leptin and ob RNA in obese and weight-reduced subjects. *Nat Med* **1**, 1155–1161 (1995).
24. Park, H.-K. & Ahima, R. S. Leptin signaling. *F1000Prime Rep* **6**, 73 (2014).
25. Singh, R. K., Kumar, P. & Mahalingam, K. Molecular genetics of human obesity: A comprehensive review. *Comptes Rendus Biologies* **340**, 87–108 (2017).
26. Mahmoud, R., Kimonis, V. & Butler, M. G. Genetics of Obesity in Humans: A Clinical Review. *Int J Mol Sci* **23**, 11005 (2022).
27. Caroline M. Apovian, M. D. Obesity: Definition, Comorbidities, Causes, and Burden. **22**, (2016).
28. Perry, B. & Wang, Y. Appetite regulation and weight control: the role of gut hormones. *Nutr Diabetes* **2**, e26 (2012).

29. Cowley, M. A. *et al.* Leptin activates anorexigenic POMC neurons through a neural network in the arcuate nucleus. *Nature* **411**, 480–484 (2001).
30. Nasrallah, C. M. & Horvath, T. L. Mitochondrial dynamics in the central regulation of metabolism. *Nat Rev Endocrinol* **10**, 650–658 (2014).
31. Baldini, G. & Phelan, K. D. The melanocortin pathway and control of appetite- progress and therapeutic implications. *J Endocrinol* **241**, R1–R33 (2019).
32. Harno, E., Gali Ramamoorthy, T., Coll, A. P. & White, A. POMC: The Physiological Power of Hormone Processing. *Physiol Rev* **98**, 2381–2430 (2018).
33. Kühnen, P., Krude, H. & Biebermann, H. Melanocortin-4 Receptor Signalling: Importance for Weight Regulation and Obesity Treatment. *Trends in Molecular Medicine* **25**, 136–148 (2019).
34. Silvestri, C. & Di Marzo, V. The Endocannabinoid System in Energy Homeostasis and the Etiopathology of Metabolic Disorders. *Cell Metabolism* **17**, 475–490 (2013).
35. Ramos, E. J. B., Meguid, M. M., Campos, A. C. L. & Coelho, J. C. U. Neuropeptide Y, alpha-melanocyte-stimulating hormone, and monoamines in food intake regulation. *Nutrition* **21**, 269–279 (2005).
36. Baron, M., Froguel, P. & Bonnefond, A. Du nouveau dans la génétique des formes monogéniques d'obésité et son impact pour mieux en comprendre la physiopathologie. *Med Sci (Paris)* **36**, 859–865 (2020).
37. Morris, D. L. & Rui, L. Recent advances in understanding leptin signaling and leptin resistance. *Am J Physiol Endocrinol Metab* **297**, E1247–E1259 (2009).
38. Jiang, L., Li, Z. & Rui, L. Leptin stimulates both JAK2-dependent and JAK2-independent signaling pathways. *J Biol Chem* **283**, 28066–28073 (2008).
39. Loos, R. J. F. & Yeo, G. S. H. The genetics of obesity: from discovery to biology. *Nat Rev Genet* **23**, 120–133 (2022).
40. Tao, Y.-X. The Melanocortin-4 Receptor: Physiology, Pharmacology, and Pathophysiology. *Endocr Rev* **31**, 506–543 (2010).
41. Fatima, M. T., Ahmed, I., Fakhro, K. A. & Akil, A. S. A. Melanocortin-4 receptor complexity in energy homeostasis, obesity and drug development strategies. *Diabetes Obes Metab* **24**, 583–598 (2022).
42. Xu, B. & Xie, X. Neurotrophic Factor Control of Satiety and Body Weight. *Nat Rev Neurosci* **17**, 282–292 (2016).
43. Xu, B. *et al.* Brain-derived neurotrophic factor regulates energy balance downstream of melanocortin-4 receptor. *Nat Neurosci* **6**, 736–742 (2003).
44. El-Sayed Moustafa, J. S. & Froguel, P. From obesity genetics to the future of personalized obesity therapy. *Nat Rev Endocrinol* **9**, 402–413 (2013).
45. van der Klaauw, A. A. & Farooqi, I. S. The Hunger Genes: Pathways to Obesity. *Cell* **161**, 119–132 (2015).
46. Pulst, S. M. Genetic Linkage Analysis. *Archives of Neurology* **56**, 667–672 (1999).
47. Walley, A. J., Asher, J. E. & Froguel, P. The genetic contribution to non-syndromic human obesity. *Nat Rev Genet* **10**, 431–442 (2009).
48. Kwon, J. M. & Goate, A. M. The Candidate Gene Approach. *Alcohol Research & Health* **24**, 164 (2000).
49. Herrera, B. M. & Lindgren, C. M. The Genetics of Obesity. *Curr Diab Rep* **10**, 498–505 (2010).
50. Yengo, L. *et al.* Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. *Human Molecular Genetics* **27**, 3641–3649 (2018).
51. Bonnefond, A. & Froguel, P. Rare and Common Genetic Events in Type 2 Diabetes: What Should Biologists Know? *Cell Metabolism* **21**, 357–368 (2015).
52. Rankinen, T. *et al.* The Human Obesity Gene Map: The 2005 Update. *Obesity* **14**, 529–644 (2006).
53. Saunders, C. L. *et al.* Meta-Analysis of Genome-wide Linkage Studies in BMI and Obesity. *Obesity* **15**, 2263–2275 (2007).
54. Kurokawa, N. *et al.* The ADRB3 Trp64Arg variant and BMI: a meta-analysis of 44 833 individuals. *Int J Obes* **32**, 1240–1249 (2008).

55. Shugart, Y. Y. *et al.* Two British women studies replicated the association between the Val66Met polymorphism in the brain-derived neurotrophic factor (BDNF) and BMI. *Eur J Hum Genet* **17**, 1050–1055 (2009).
56. Benzinou, M. *et al.* Endocannabinoid receptor 1 gene variations increase risk for obesity and modulate body mass index in European populations. *Hum Mol Genet* **17**, 1916–1921 (2008).
57. Wang, D. *et al.* Association of the MC4R V103I polymorphism with obesity: a Chinese case-control study and meta-analysis in 55,195 individuals. *Obesity (Silver Spring)* **18**, 573–579 (2010).
58. Benzinou, M. *et al.* Common nonsynonymous variants in PCSK1 confer risk of obesity. *Nat Genet* **40**, 943–945 (2008).
59. Tönjes, A., Scholz, M., Loeffler, M. & Stumvoll, M. Association of Pro12Ala polymorphism in peroxisome proliferator-activated receptor gamma with Pre-diabetic phenotypes: meta-analysis of 57 studies on nondiabetic individuals. *Diabetes Care* **29**, 2489–2497 (2006).
60. Frayling, T. M. *et al.* A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* **316**, 889–894 (2007).
61. Scuteri, A. *et al.* Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet* **3**, e115 (2007).
62. Loos, R. J. F. & Yeo, G. S. H. The bigger picture of FTO: the first GWAS-identified obesity gene. *Nat Rev Endocrinol* **10**, 51–61 (2014).
63. Dina, C. *et al.* Variation in FTO contributes to childhood obesity and severe adult obesity. *Nat Genet* **39**, 724–726 (2007).
64. Turcot, V. *et al.* Protein-altering variants associated with body mass index implicate pathways that control energy intake and expenditure underpinning obesity. *Nat Genet* **50**, 26–41 (2018).
65. Loos, R. J. F. *et al.* Common variants near MC4R are associated with fat mass, weight and risk of obesity. *Nat Genet* **40**, 768–775 (2008).
66. Chambers, J. C. *et al.* Common genetic variation near MC4R is associated with waist circumference and insulin resistance. *Nat Genet* **40**, 716–718 (2008).
67. Willer, C. J. *et al.* Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat Genet* **41**, 25–34 (2009).
68. Meyre, D. *et al.* Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. *Nat Genet* **41**, 157–159 (2009).
69. Lindgren, C. M. *et al.* Genome-wide association scan meta-analysis identifies three Loci influencing adiposity and fat distribution. *PLoS Genet* **5**, e1000508 (2009).
70. Ndiaye, F. K. *et al.* The expression of genes in top obesity-associated loci is enriched in insula and substantia nigra brain regions involved in addiction and reward. *Int J Obes (Lond)* **44**, 539–543 (2020).
71. Bonnefond, A. *et al.* Rare MTNR1B variants impairing melatonin receptor 1B function contribute to type 2 diabetes. *Nature Genetics* **44**, 297–301 (2012).
72. Ichimura, A. *et al.* Dysfunction of lipid sensor GPR120 leads to obesity in both mouse and human. *Nature* **483**, 350–354 (2012).
73. Baron, M. *et al.* Loss-of-function mutations in MRAP2 are pathogenic in hyperphagic obesity with hyperglycemia and hypertension. *Nat Med* **25**, 1733–1738 (2019).
74. Emdin, C. A. *et al.* Analysis of predicted loss-of-function variants in UK Biobank identifies variants protective for disease. *Nat Commun* **9**, 1613 (2018).
75. Akbari, P. *et al.* Sequencing of 640,000 exomes identifies GPR75 variants associated with protection from obesity. *Science* **373**, eabf8683 (2021).
76. Speliotes, E. K. *et al.* Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* **42**, 937–948 (2010).
77. Jarick, I. *et al.* Novel common copy number variation for early onset extreme obesity on chromosome 11q11 identified by a genome-wide analysis. *Hum Mol Genet* **20**, 840–852 (2011).
78. Falchi, M. *et al.* Low copy number of the salivary amylase gene predisposes to obesity. *Nat Genet* **46**, 492–497 (2014).
79. Xie, H., Sun, L. & Lodish, H. F. Targeting microRNAs in obesity. *Expert Opin Ther Targets* **13**, 1227–1238 (2009).
80. Stöger, R. Epigenetics and obesity. *Pharmacogenomics* **9**, 1851–1860 (2008).

81. Huvenne, H., Dubern, B., Clément, K. & Poitou, C. Rare Genetic Forms of Obesity: Clinical Approach and Current Treatments in 2016. *Obes Facts* **9**, 158–173 (2016).
82. Saeed, S. *et al.* Genetic variants in LEP, LEPR, and MC4R explain 30% of severe obesity in children from a consanguineous population. *Obesity (Silver Spring)* **23**, 1687–1695 (2015).
83. Saeed, S., Arslan, M. & Froguel, P. Genetics of Obesity in Consanguineous Populations: Toward Precision Medicine and the Discovery of Novel Obesity Genes. *Obesity (Silver Spring)* **26**, 474–484 (2018).
84. Dubern, B., Faccioli, N., Poitou, C. & Clément, K. Novel therapeutics in rare genetic obesities: A narrative review. *Pharmacol Res* **191**, 106763 (2023).
85. Zhang, Y. *et al.* Positional cloning of the mouse obese gene and its human homologue. *Nature* **372**, 425–432 (1994).
86. Dubern, B. & Clement, K. Leptin and leptin receptor-related monogenic obesity. *Biochimie* **94**, 2111–2115 (2012).
87. Chen, H. *et al.* Evidence that the diabetes gene encodes the leptin receptor: identification of a mutation in the leptin receptor gene in db/db mice. *Cell* **84**, 491–495 (1996).
88. Montague, C. T. *et al.* Congenital leptin deficiency is associated with severe early-onset obesity in humans. *Nature* **387**, 903–908 (1997).
89. Clément, K. *et al.* A mutation in the human leptin receptor gene causes obesity and pituitary dysfunction. *Nature* **392**, 398–401 (1998).
90. Krude, H. *et al.* Severe early-onset obesity, adrenal insufficiency and red hair pigmentation caused by POMC mutations in humans. *Nat Genet* **19**, 155–157 (1998).
91. Krude, H. & Grüters, A. Implications of proopiomelanocortin (POMC) mutations in humans: the POMC deficiency syndrome. *Trends Endocrinol Metab* **11**, 15–22 (2000).
92. Le Collen, L. *et al.* Heterozygous pathogenic variants in POMC are not responsible for monogenic obesity: Implication for MC4R agonist use. *Genet Med* **25**, 100857 (2023).
93. Stijnen, P., Ramos-Molina, B., O’Rahilly, S. & Creemers, J. W. M. PCSK1 Mutations and Human Endocrinopathies: From Obesity to Gastrointestinal Disorders. *Endocr Rev* **37**, 347–371 (2016).
94. O’Rahilly, S. *et al.* Impaired Processing of Prohormones Associated with Abnormalities of Glucose Homeostasis and Adrenal Function. *N Engl J Med* **333**, 1386–1391 (1995).
95. Philippe, J. *et al.* A nonsense loss-of-function mutation in PCSK1 contributes to dominantly inherited human obesity. *Int J Obes* **39**, 295–302 (2015).
96. Creemers, J. W. M. *et al.* Heterozygous Mutations Causing Partial Prohormone Convertase 1 Deficiency Contribute to Human Obesity. *Diabetes* **61**, 383–390 (2012).
97. Vaisse, C., Clement, K., Guy-Grand, B. & Froguel, P. A frameshift mutation in human MC4R is associated with a dominant form of obesity. *Nat Genet* **20**, 113–114 (1998).
98. Yeo, G. S. *et al.* A frameshift mutation in MC4R associated with dominantly inherited human obesity. *Nat Genet* **20**, 111–112 (1998).
99. Farooqi, I. S. *et al.* Clinical Spectrum of Obesity and Mutations in the Melanocortin 4 Receptor Gene. *N Engl J Med* **348**, 1085–1095 (2003).
100. Chami, N., Preuss, M., Walker, R. W., Moscati, A. & Loos, R. J. F. The role of polygenic susceptibility to obesity among carriers of pathogenic mutations in MC4R in the UK Biobank population. *PLoS Med* **17**, e1003196 (2020).
101. Ramachandrappa, S. *et al.* Rare variants in single-minded 1 (SIM1) are associated with severe obesity. *J Clin Invest* **123**, 3042–3050 (2013).
102. Bonnefond, A. *et al.* Loss-of-function mutations in SIM1 contribute to obesity and Prader-Willi-like features. *J Clin Invest* **123**, 3037–3041 (2013).
103. Gray, J. *et al.* Hyperphagia, Severe Obesity, Impaired Cognitive Function, and Hyperactivity Associated With Functional Loss of One Copy of the Brain-Derived Neurotrophic Factor (BDNF) Gene. *Diabetes* **55**, 3366–3371 (2006).
104. Harcourt, B. E. *et al.* Maternal inheritance of BDNF deletion, with phenotype of obesity and developmental delay in mother and child. *Am J Med Genet* **176**, 194–200 (2018).
105. Sandrini, L. *et al.* Association between Obesity and Circulating Brain-Derived Neurotrophic Factor (BDNF) Levels: Systematic Review of Literature and Meta-Analysis. *Int J Mol Sci* **19**, 2281 (2018).

106. Han, J. C. Rare Syndromes and Common Variants of the Brain-Derived Neurotrophic Factor Gene in Human Obesity. *Prog Mol Biol Transl Sci* **140**, 75–95 (2016).
107. Yeo, G. S. H. *et al.* A de novo mutation affecting human TrkB associated with severe obesity and developmental delay. *Nat Neurosci* **7**, 1187–1189 (2004).
108. Doche, M. E. *et al.* Human SH2B1 mutations are associated with maladaptive behaviors and obesity. *J Clin Invest* **122**, 4732–4736 (2012).
109. Bochukova, E. G. *et al.* Large, rare chromosomal deletions associated with severe early-onset obesity. *Nature* **463**, 666–670 (2010).
110. Siljee, J. E. *et al.* Subcellular localization of MC4R with ADCY3 at neuronal primary cilia underlies a common pathway for genetic predisposition to obesity. *Nat Genet* **50**, 180–185 (2018).
111. Saeed, S. *et al.* Loss-of-function mutations in ADCY3 cause monogenic severe obesity. *Nat Genet* **50**, 175–179 (2018).
112. Miltenberger, R. J., Mynatt, R. L., Wilkinson, J. E. & Woychik, R. P. The Role of the agouti Gene in the Yellow Obese Syndrome<sup>1,2</sup>. *The Journal of Nutrition* **127**, 1902S-1907S (1997).
113. Kempf, E. *et al.* Aberrant expression of agouti signaling protein (ASIP) as a cause of monogenic severe childhood obesity. *Nat Metab* **4**, 1697–1712 (2022).
114. Pearce, L. R. *et al.* KSR2 Mutations Are Associated with Obesity, Insulin Resistance, and Impaired Cellular Fuel Oxidation. *Cell* **155**, 765–777 (2013).
115. van der Klaauw, A. A. *et al.* Human Semaphorin 3 Variants Link Melanocortin Circuit Development and Energy Balance. *Cell* **176**, 729-742.e18 (2019).
116. Yang, Y. *et al.* Steroid receptor coactivator-1 modulates the function of Pomc neurons and energy homeostasis. *Nat Commun* **10**, 1718 (2019).
117. He, Y. *et al.* Human loss-of-function variants in the serotonin 2C receptor associated with obesity and maladaptive behavior. *Nat Med* **28**, 2537–2546 (2022).
118. Borman, A. D. *et al.* A homozygous mutation in the TUB gene associated with retinal dystrophy and obesity. *Hum Mutat* **35**, 289–293 (2014).
119. Alsters, S. I. M. *et al.* Truncating Homozygous Mutation of Carboxypeptidase E (CPE) in a Morbidly Obese Female with Type 2 Diabetes Mellitus, Intellectual Disability and Hypogonadotrophic Hypogonadism. *PLoS One* **10**, e0131417 (2015).
120. Kaur, Y., de Souza, R. J., Gibson, W. T. & Meyre, D. A systematic review of genetic syndromes with obesity. *Obes Rev* **18**, 603–634 (2017).
121. Butler, M. G., Miller, J. L. & Forster, J. L. Prader-Willi Syndrome - Clinical Genetics, Diagnosis and Treatment Approaches: An Update. *Curr Pediatr Rev* **15**, 207–244 (2019).
122. Schaaf, C. P. *et al.* Truncating mutations of MAGEL2 cause Prader-Willi phenotypes and autism. *Nat Genet* **45**, 1405–1408 (2013).
123. Priya, S., Nampoothiri, S., Sen, P. & Sriprya, S. Bardet–Biedl syndrome: Genetics, molecular pathophysiology, and disease management. *Indian J Ophthalmol* **64**, 620–627 (2016).
124. Chennen, K., Scerbo, M. J., Dollfus, H., Poch, O. & Marion, V. Syndrome de Bardet-Biedl : cils et obésité - De la génétique aux approches intégratives. *Med Sci (Paris)* **30**, 1034–1039 (2014).
125. Choudhury, A. R., Munonye, I., Sanu, K. P., Islam, N. & Gadaga, C. A review of Alström syndrome: a rare monogenic ciliopathy. *Intractable & Rare Diseases Research* **10**, 257 (2021).
126. Marshall, J. D., Maffei, P., Collin, G. B. & Naggert, J. K. Alström Syndrome: Genetics and Clinical Overview. *Curr Genomics* **12**, 225–235 (2011).
127. Girard, D. & Petrovsky, N. Alström syndrome: insights into the pathogenesis of metabolic disorders. *Nat Rev Endocrinol* **7**, 77–88 (2011).
128. Walters, R. G. *et al.* A novel highly-penetrant form of obesity due to microdeletions on chromosome 16p11.2. *Nature* **463**, 671–675 (2010).
129. Bachmann-Gagescu, R. *et al.* Recurrent 200-kb deletions of 16p11.2 that include the SH2B1 gene are associated with developmental delay and obesity. *Genet Med* **12**, 641–647 (2010).
130. Chung, W. K., Roberts, T. P., Sherr, E. H., Snyder, L. G. & Spiro, J. E. 16p11.2 deletion syndrome. *Curr Opin Genet Dev* **68**, 49–56 (2021).
131. Reaven, G. M. Role of Insulin Resistance in Human Disease. *Diabetes* **37**, 1595–1607 (1988).
132. Björntorp, P. Metabolic Implications of Body Fat Distribution. *Diabetes Care* **14**, 1132–1143 (1991).

133. Kissebah, A. H. *et al.* Quantitative trait loci on chromosomes 3 and 17 influence phenotypes of the metabolic syndrome. *Proc Natl Acad Sci U S A* **97**, 14478–14483 (2000).
134. Keramati, A. R. *et al.* A Form of the Metabolic Syndrome Associated with Mutations in DYRK1B. *N Engl J Med* **370**, 1909–1919 (2014).
135. Esteghamat, F. *et al.* CELA2A mutations predispose to early-onset atherosclerosis and metabolic syndrome and affect plasma insulin and platelet activation. *Nat Genet* **51**, 1233–1243 (2019).
136. Burnett, L. C. *et al.* Deficiency in prohormone convertase PC1 impairs prohormone processing in Prader-Willi syndrome. *J Clin Invest* **127**, 293–305 (2017).
137. Guo, D.-F. *et al.* The BBSome Controls Energy Homeostasis by Mediating the Transport of the Leptin Receptor to the Plasma Membrane. *PLoS Genetics* **12**, e1005890 (2016).
138. Bonnefond, A. *et al.* Pathogenic variants in actionable MODY genes are associated with type 2 diabetes. *Nat Metab* **2**, 1126–1134 (2020).
139. Obésité de l’adulte : améliorer la prise en charge médicale des cas les plus sévères. *Haute Autorité de Santé* [https://www.has-sante.fr/jcms/p\\_3346198/fr/obesite-de-l-adulte-ameliorer-la-prise-en-charge-medicaledes-cas-les-plus-severes](https://www.has-sante.fr/jcms/p_3346198/fr/obesite-de-l-adulte-ameliorer-la-prise-en-charge-medicaledes-cas-les-plus-severes).
140. Sjöström, L. Review of the key results from the Swedish Obese Subjects (SOS) trial - a prospective controlled intervention study of bariatric surgery. *J Intern Med* **273**, 219–234 (2013).
141. Obésité : prise en charge chirurgicale chez l’adulte - Brochure patients. *Haute Autorité de Santé* [https://www.has-sante.fr/jcms/c\\_849636/fr/obesite-prise-en-charge-chirurgicale-chez-l-adulte-brochure-patients](https://www.has-sante.fr/jcms/c_849636/fr/obesite-prise-en-charge-chirurgicale-chez-l-adulte-brochure-patients).
142. Chirurgie de l’obésité pour les moins de 18 ans : à n’envisager que dans des cas très particuliers. *Haute Autorité de Santé* [https://www.has-sante.fr/jcms/c\\_2621051/fr/chirurgie-de-l-obesite-pour-les-moins-de-18-ans-a-n-envisager-que-dans-des-cas-tres-particuliers](https://www.has-sante.fr/jcms/c_2621051/fr/chirurgie-de-l-obesite-pour-les-moins-de-18-ans-a-n-envisager-que-dans-des-cas-tres-particuliers).
143. Aslan, I. R. *et al.* Bariatric surgery in a patient with complete MC4R deficiency. *Int J Obes* **35**, 457–461 (2011).
144. Richards, S. *et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* **17**, 405–424 (2015).
145. Farooqi, I. S. & O’Rahilly, S. 20 YEARS OF LEPTIN: Human disorders of leptin action. *Journal of Endocrinology* **223**, T63–T70 (2014).
146. Farooqi, I. S. *et al.* Effects of Recombinant Leptin Therapy in a Child with Congenital Leptin Deficiency. *New England Journal of Medicine* **341**, 879–884 (1999).
147. Markham, A. Setmelanotide: First Approval. *Drugs* **81**, 397–403 (2021).
148. Trapp, C. M. & Censani, M. Setmelanotide: a promising advancement for pediatric patients with rare forms of genetic obesity. *Curr Opin Endocrinol Diabetes Obes* **30**, 136–140 (2023).
149. Kühnen, P. *et al.* Proopiomelanocortin Deficiency Treated with a Melanocortin-4 Receptor Agonist. *N Engl J Med* **375**, 240–246 (2016).
150. Clément, K. *et al.* MC4R agonism promotes durable weight loss in patients with leptin receptor deficiency. *Nat Med* **24**, 551–555 (2018).
151. Clément, K. *et al.* Efficacy and safety of setmelanotide, an MC4R agonist, in individuals with severe obesity due to LEPR or POMC deficiency: single-arm, open-label, multicentre, phase 3 trials. *The Lancet Diabetes & Endocrinology* **8**, 960–970 (2020).
152. Collet, T.-H. *et al.* Evaluation of a melanocortin-4 receptor (MC4R) agonist (Setmelanotide) in MC4R deficiency. *Mol Metab* **6**, 1321–1329 (2017).
153. Haws, R. *et al.* Effect of setmelanotide, a melanocortin-4 receptor agonist, on obesity in Bardet-Biedl syndrome. *Diabetes Obes Metab* **22**, 2133–2140 (2020).
154. Haqq, A. M. *et al.* Efficacy and safety of setmelanotide, a melanocortin-4 receptor agonist, in patients with Bardet-Biedl syndrome and Alström syndrome: a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial with an open-label period. *Lancet Diabetes Endocrinol* **10**, 859–868 (2022).
155. O’Neil, P. M. *et al.* Efficacy and safety of semaglutide compared with liraglutide and placebo for weight loss in patients with obesity: a randomised, double-blind, placebo and active controlled, dose-ranging, phase 2 trial. *The Lancet* **392**, 637–649 (2018).

156. Wadden, T. A. *et al.* Effect of Subcutaneous Semaglutide vs Placebo as an Adjunct to Intensive Behavioral Therapy on Body Weight in Adults With Overweight or Obesity: The STEP 3 Randomized Clinical Trial. *JAMA* **325**, 1403–1413 (2021).
157. He, Z. *et al.* Direct and indirect effects of liraglutide on hypothalamic POMC and NPY/AgRP neurons – Implications for energy balance and glucose control. *Mol Metab* **28**, 120–134 (2019).
158. Iepsen, E. W. *et al.* Patients with Obesity Caused by Melanocortin-4 Receptor Mutations Can Be Treated with a Glucagon-like Peptide-1 Receptor Agonist. *Cell Metabolism* **28**, 23–32.e3 (2018).
159. Ramos-Molina, B., Martin, M. G. & Lindberg, I. PCSK1 Variants and Human Obesity. in *Progress in Molecular Biology and Translational Science* vol. 140 47–74 (Elsevier, 2016).
160. Zhu, X. *et al.* Disruption of PC1/3 expression in mice causes dwarfism and multiple neuroendocrine peptide processing defects. *Proc Natl Acad Sci U S A* **99**, 10293–10298 (2002).
161. Prabhu, Y. *et al.* Defective Transport of the Obesity Mutant PC1/3 N222D Contributes to Loss of Function. *Endocrinology* **155**, 2391–2401 (2014).
162. Stijnen, P. *et al.* Endoplasmic reticulum-associated degradation of the mouse PC1/3-N222D hypomorph and human PCSK1 mutations contributes to obesity. *Int J Obes* **40**, 973–981 (2016).
163. Lloyd, D. J., Bohan, S. & Gekakis, N. Obesity, hyperphagia and increased metabolic efficiency in Pcl mutant mice. *Human Molecular Genetics* **15**, 1884–1893 (2006).
164. Muhsin, N. I. A., Bentley, L., Bai, Y., Goldsworthy, M. & Cox, R. D. A novel mutation in the mouse Pcsk1 gene showing obesity and diabetes. *Mamm Genome* **31**, 17–29 (2020).
165. Shakya, M. *et al.* The G209R mutant mouse as a model for human PCSK1 polyendocrinopathy. *Endocrinology* **163**, bqac024 (2022).
166. Wen, W. *et al.* Meta-analysis identifies common variants associated with body mass index in East Asians. *Nat Genet* **44**, 307–311 (2012).
167. Blanco, E. H., Ramos-Molina, B. & Lindberg, I. Revisiting PC1/3 Mutants: Dominant-Negative Effect of Endoplasmic Reticulum-Retained Mutants. *Endocrinology* **156**, 3625–3637 (2015).
168. Mbikay, M., Sirois, F., Nkongolo, K. K., Basak, A. & Chrétien, M. Effects of rs6234/rs6235 and rs6232/rs6234/rs6235 PCSK1 single-nucleotide polymorphism clusters on proprotein convertase 1/3 biosynthesis and activity. *Molecular Genetics and Metabolism* **104**, 682–687 (2011).
169. Stijnen, P. *et al.* The association of common variants in PCSK1 with obesity: a HuGE review and meta-analysis. *Am J Epidemiol* **180**, 1051–1065 (2014).
170. Nead, K. T. *et al.* Contribution of common non-synonymous variants in PCSK1 to body mass index variation and risk of obesity: a systematic review and meta-analysis with evidence from up to 331 175 individuals. *Hum Mol Genet* **24**, 3582–3594 (2015).
171. Chang, Y.-C. *et al.* Common PCSK1 haplotypes are associated with obesity in the Chinese population. *Obesity (Silver Spring)* **18**, 1404–1409 (2010).
172. Liu, C.-T. *et al.* Genome-wide association of body fat distribution in African ancestry populations suggests new loci. *PLoS Genet* **9**, e1003681 (2013).
173. Jackson, R. S. *et al.* Obesity and impaired prohormone processing associated with mutations in the human prohormone convertase 1 gene. *Nat Genet* **16**, 303–306 (1997).
174. Jackson, R. S. *et al.* Small-intestinal dysfunction accompanies the complex endocrinopathy of human proprotein convertase 1 deficiency. *J. Clin. Invest.* **112**, 1550–1560 (2003).
175. Farooqi, I. S. *et al.* Hyperphagia and Early-Onset Obesity due to a Novel Homozygous Missense Mutation in Prohormone Convertase 1/3. *The Journal of Clinical Endocrinology & Metabolism* **92**, 3369–3373 (2007).
176. Pépin, L. *et al.* A New Case of PCSK1 Pathogenic Variant With Congenital Proprotein Convertase 1/3 Deficiency and Literature Review. *The Journal of Clinical Endocrinology & Metabolism* **104**, 985–993 (2019).
177. Frank, G. R. *et al.* Severe obesity and diabetes insipidus in a patient with PCSK1 deficiency. *Mol Genet Metab* **110**, 191–194 (2013).
178. Martín, M. G. *et al.* Congenital proprotein convertase 1/3 deficiency causes malabsorptive diarrhea and other endocrinopathies in a pediatric cohort. *Gastroenterology* **145**, 138–148 (2013).
179. Ni, Y. *et al.* Modulation of PC1/3 activity by a rare double-site homozygous mutation. *Front Pediatr* **10**, 1026707 (2022).

180. Qian, Y. *et al.* Case Report: Complete Maternal Uniparental Isodisomy of Chromosome 5 (iUPD(5)mat) With PCSK1 Nonsense Variant in an Infant With Recurrent Diarrhea. *Front Genet* **12**, 668326 (2021).
181. Aerts, L. *et al.* Novel Homozygous Inactivating Mutation in the PCSK1 Gene in an Infant with Congenital Malabsorptive Diarrhea. *Genes (Basel)* **12**, 710 (2021).
182. Karacan Küçükali, G. *et al.* A case of prohormone convertase deficiency diagnosed with type 2 diabetes. *Turk Arch Pediatr* **56**, 81–84 (2021).
183. Löffler, D. Functional and clinical relevance of novel and known PCSK1 variants for childhood obesity and glucose metabolism. 11.
184. Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat Methods* **6**, 359–362 (2009).
185. Wiśniewski, J., Ostasiewicz, P. & Mann, M. High recovery FASP applied to the proteomic analysis of microdissected formalin fixed paraffin embedded cancer tissues retrieves known colon cancer markers. *Journal of proteome research* (2011) doi:10.1021/pr200019m.
186. Bache, N. *et al.* A Novel LC System Embeds Analytes in Pre-formed Gradients for Rapid, Ultra-robust Proteomics \*. *Molecular & Cellular Proteomics* **17**, 2284–2296 (2018).
187. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367–1372 (2008).
188. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**, 731–740 (2016).
189. Sun, J., Zheng, Y. & Hsu, L. A unified mixed-effects model for rare-variant association in sequencing studies. *Genet Epidemiol* **37**, 334–344 (2013).
190. Wright, J. *et al.* Dominant protein interactions that influence the pathogenesis of conformational diseases. *J Clin Invest* **123**, 3124–3134 (2013).
191. Ioannidis, N. M. *et al.* REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am J Hum Genet* **99**, 877–885 (2016).
192. Van Dijck, E. *et al.* Rare Heterozygous PCSK1 Variants in Human Obesity: The Contribution of the p.Y181H Variant and a Literature Review. *Genes (Basel)* **13**, 1746 (2022).
193. Ozcan, U. *et al.* Endoplasmic reticulum stress links obesity, insulin action, and type 2 diabetes. *Science* **306**, 457–461 (2004).
194. Ramos-Molina, B., Molina-Vega, M., Fernández-García, J. C. & Creemers, J. W. Hyperphagia and Obesity in Prader–Willi Syndrome: PCSK1 Deficiency and Beyond? *Genes (Basel)* **9**, 288 (2018).
195. Singh, R. & Lauth, M. Emerging Roles of DYRK Kinases in Embryogenesis and Hedgehog Pathway Control. *J Dev Biol* **5**, 13 (2017).
196. GTEx Portal. <https://gtexportal.org/home/gene/DYRK1B>.
197. Abu Jhaisha, S. *et al.* DYRK1B mutations associated with metabolic syndrome impair the chaperone-dependent maturation of the kinase domain. *Sci Rep* **7**, 6420 (2017).
198. Kokkorakis, N. & Gaitanou, M. Minibrain-related kinase/dual-specificity tyrosine-regulated kinase 1B implication in stem/cancer stem cells biology. *World J Stem Cells* **12**, 1553–1575 (2020).
199. Becker, W. A wake-up call to quiescent cancer cells – potential use of DYRK1B inhibitors in cancer therapy. *The FEBS Journal* **285**, 1203–1211 (2018).
200. Zhuang, L. *et al.* DYRK1B-STAT3 Drives Cardiac Hypertrophy and Heart Failure by Impairing Mitochondrial Bioenergetics. *Circulation* **145**, 829–846 (2022).
201. Bhat, N. *et al.* Dyrk1b promotes hepatic lipogenesis by bypassing canonical insulin signaling and directly activating mTORC2 in mice. *J Clin Invest* **132**, e153724.
202. Rühlemann, M. C. *et al.* Genome-wide association study in 8,956 German individuals identifies influence of ABO histo-blood groups on gut microbiome. *Nat Genet* **53**, 147–155 (2021).
203. Gao, X. R., Huang, H. & Kim, H. Genome-wide association analyses identify 139 loci associated with macular thickness in the UK Biobank cohort. *Hum Mol Genet* **28**, 1162–1172 (2019).
204. Ashford, A. L. *et al.* Identification of DYRK1B as a substrate of ERK1/2 and characterisation of the kinase activity of DYRK1B mutants from cancer and metabolic syndrome. *Cell Mol Life Sci* **73**, 883–900 (2016).

205. Mendoza-Caamal, E. C. *et al.* Two novel variants in DYRK1B causative of AOMS3: expanding the clinical spectrum. *Orphanet J Rare Dis* **16**, 291 (2021).
206. Orenstein, N. *et al.* DYRK1B haploinsufficiency in a family with metabolic syndrome and abnormal cognition. *Clin Genet* **101**, 265–266 (2022).
207. Lee, K., Deng, X. & Friedman, E. Mirk protein kinase is a mitogen-activated protein kinase substrate that mediates survival of colon cancer cells. *Cancer Res* **60**, 3631–3637 (2000).
208. Simoneschi, D. *et al.* CRL4AMBRA1 is a master regulator of D-type cyclins. *Nature* **592**, 789–793 (2021).
209. Costanzo, M. C. *et al.* The Type 2 Diabetes Knowledge Portal: An open access genetic resource dedicated to type 2 diabetes and related traits. *Cell Metab* **35**, 695-710.e6 (2023).
210. Balduzzi, S., Rücker, G. & Schwarzer, G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health* **22**, 153–160 (2019).
211. von Groote-Bidlingmaier, F. *et al.* DYRK1 is a co-activator of FKHR (FOXO1a)-dependent glucose-6-phosphatase gene expression. *Biochem Biophys Res Commun* **300**, 764–769 (2003).
212. Lauth, M. *et al.* DYRK1B-dependent autocrine-to-paracrine shift of Hedgehog signaling by mutant RAS. *Nat Struct Mol Biol* **17**, 718–725 (2010).
213. Courcet, J.-B. *et al.* The DYRK1A gene is a cause of syndromic intellectual disability with severe microcephaly and epilepsy. *J Med Genet* **49**, 731–736 (2012).
214. Hong, S.-H. *et al.* Minibrain/Dyrk1a Regulates Food Intake through the Sir2-FOXO-sNPF/NPY Pathway in Drosophila and Mammals. *PLoS Genet* **8**, e1002857 (2012).
215. Ziki, M. D. A. & Mani, A. Metabolic Syndrome: Genetic Insights into Disease Pathogenesis. *Curr Opin Lipidol* **27**, 162–171 (2016).
216. Acekifi, C. *et al.* Pharmacologic and genetic approaches define human pancreatic  $\beta$  cell mitogenic targets of DYRK1A inhibitors. *JCI Insight* **5**, e132594.
217. Shah, B. P., Sleiman, P. M., Mc Donald, J., Moeller, I. H. & Kleyn, P. Functional characterization of all missense variants in LEPR, PCSK1, and POMC genes arising from single-nucleotide variants. *Expert Review of Endocrinology & Metabolism* **18**, 209–219 (2023).
218. Montagne, L. *et al.* CoDE-seq, an augmented whole-exome sequencing, enables the accurate detection of CNVs and mutations in Mendelian obesity and intellectual disability. *Mol Metab* **13**, 1–9 (2018).
219. Gasperini, M., Starita, L. & Shendure, J. The power of multiplexed functional analysis of genetic variants. *Nat Protoc* **11**, 1782–1787 (2016).
220. European Programme OBELISK | The European Childhood Obesity Group. <https://www.ecog-obesity.eu/european-programme-obelisk/>.
221. Gao, W. *et al.* Human GLP1R variants affecting GLP1R cell surface expression are associated with impaired glucose control and increased adiposity. *Nat Metab* (2023) doi:10.1038/s42255-023-00889-6.
222. Courbage, S. *et al.* Implication of Heterozygous Variants in Genes of the Leptin–Melanocortin Pathway in Severe Obesity. *The Journal of Clinical Endocrinology & Metabolism* **106**, 2991–3006 (2021).
223. Khera, A. V. *et al.* Polygenic Prediction of Weight and Obesity Trajectories from Birth to Adulthood. *Cell* **177**, 587-596.e9 (2019).

# ANNEXES



## **ANNEXE 1. Tableaux supplémentaires**

**Tableau S1. Les caractéristiques de population de l'étude RaDiO.** Les données sont la moyenne  $\pm$  SD ou le nombre en %. *IMC*, indice de masse corporelle ; *F*, féminin ; *M*, masculin ; *NA*, non disponible ; *DT2*, diabète de type 2.

	Adultes avec une obésité	Adultes avec un surpoids	Adultes avec un poids normal	Enfants/adolescents avec une obésité	Enfants/adolescents avec un poids normal
<b>N</b>	1524	2858	2878	1032	1028
<b>Age pendant l'investigation (années)</b>	51 $\pm$ 13	54 $\pm$ 12	48 $\pm$ 12	13 $\pm$ 2,1	18 $\pm$ 3,3
<b>IMC (kg/m<sup>2</sup>)</b>	37 $\pm$ 7,2	27 $\pm$ 1,4	22 $\pm$ 1,8	31 $\pm$ 5,4	20 $\pm$ 2,3
<b>Sexe</b>	M:512 / F:1,012	M:1,823 / F:1,035	M:1,254 / F:1,624	M:479 / F:553	M:531 / F:597
<b>DT2</b>	536 (35%)	1143 (40%)	313 (11%)	NA	NA

**Tableau S2. Liste des amorces utilisées pour le séquençage Sanger de *PCSK1* sur l'ADN des familles disponibles.**

Exon ciblé de <i>PCSK1</i>	Sens de l'amorce	Séquence de l'amorce (5'→3')
Exon 2	<i>Forward</i>	GGGCTTAATACCTGGGTGATG
	<i>Reverse</i>	TGGCTCAACCAATTCAACC
Exon 3	<i>Forward</i>	ACTGCTTCTCCTCCCCCTAA
	<i>Reverse</i>	TGATGGGAAGGGTACTGGAG
Exon 8	<i>Forward</i>	GGAAACCCTAACCTGCAAGA
	<i>Reverse</i>	GCCAGAGTCCAGTGGCTTAT
Exon 9	<i>Forward</i>	GCCAGCCTTTCAAGTGAGTAA
	<i>Reverse</i>	CAGCATTTTCGTTTGGGAATC
	<i>Reverse</i>	TGGTTAGTGACCTCGAAGTGG

**Tableau S3. Cycle de température utilisé au cours de la PCR précédant le séquençage Sanger de *PCSK1* dans les familles disponibles.** Les échantillons ont été amplifiés selon le protocole *Fast Start Taq DNA Polymerase* (Roche) avec une concentration de 2,5 mM de MgCl<sub>2</sub>.

1X	40X			1X
95°C	95°C	60°C	72°C	72°C
5 min	30 sec	30 sec	45 sec	10 min

**Tableau S4. Liste des amorces utilisées pour la réalisation des PCR de mutagenèse des 63 variants du projet *PCSK1*.**

Variants		Séquence de l'amorce <i>forward</i> (5'→3')	Séquence de l'amorce <i>reverse</i> (5'→3')
A12P	c.34G>C	GAGTCTGCAGTGCCTCTTTTCGTCCTCTTTTG	CAAAAGAGGACGAAAGGAGTGCCTGCACTGCAGACTC
L58S	c.173T>C	GGAGCTGGGCTATGACCTTTTCGGGTCAGATTG	CAATCTGACCCGAAAGGTCATAGCCCAGCTCC
F70C	c.209T>G	GTTCACTTGAAAATCACTACTTATGCAAACATAAAAACCACCCAG	CTGGGGTGGTTTTTATGTTTGCATAAGTAGTGATTTTCAAGTGAAC
R80*	c.238C>T	CACCCAGAAAGGCTTGAAGGAGTGCCTTTTC	GAAAGGCACTCCTTCAAGACCTTCTGGGGTG
R80Q	c.239G>A	CACCCAGAAAGGCTCAAAGGAGTGCCTTTTCAT	ATGAAAGGCACTCCTTTGAGACCTTCTGGGGTG
R81S	c.243G>C	CCCAGAAAGTCTCGAAGCAGTGCCTTTTCATATCAC	GTGATATGAAAGGCACTGCTTCGAGACCTTCTGGG
H85R	c.254A>G	GTCTCGAAGGAGTGCCTTTTCGTATCACTAAGAGATTATCTG	CAGATAATCTCTTAGTGATACGAAAGGCACTCCTTCGAGAC
D94G	c.281A>G	TAAGAGATTATCTGATGATGGTCTGTGATATGGGCTGAAC	GTTCCAGCCATATCACACGACCATCATCAGATAATCTCTTA
R95H	c.284G>A	GAGATTATCTGATGATGATCATGTGATATGGGCTGAACAAC	GTTGTTCCAGCCATATCACATGATCATCATCAGATAATCTC
A112T	c.334G>A	GAAAGAAGTAAACGTTCAACTCTAAGGGACTCAGCAC	GTGCTGAGTCCCTTAGAGTTGAACGTTTACTTCTTTTC
M125I	c.375G>A	CACTAAATCTCTTCAATGATCCCATATGGAATCAGCAATGG	CCATTGCTGATCCATATGGGATCATTGAAGAGATTTAGTG
L132S	c.395T>C	GTGGAATCAGCAATGGTACTCGCAAGATAACCAGG	CCTGGTATCTTGCGAGTACCATTGCTGATTCCAC
T157M	c.470C>T	ACCTGTTTGGCAAAAAGGCATTATGGGCAAAGGAGT	ACTCCTTTGCCATAATGCCTTTTTGCCAAACAGGT
T175M	c.524C>T	GATGGTTTGGAGTGAATCACATGGACATTTATGCC	GGCATAAATGTCCATGTGATTCCACTCCAAACCATC
Y181H	c.541T>C	CGGACATTTATGCCAACCATGATCCAGAGGCTAGC	GCTAGCCTCTGGATCATGGTTGGCATAAATGTCCG
Y187H	c.559T>C	CAACTATGATCCAGAGGCTAGCCATGATTTAATGATAATGACCAT	ATGGTCATTATCATAAATCATGGCTAGCCTCTGGATCATAGTTG
D193Y	c.577G>T	TCCAGAGGCTAGCTATGATTTAATGATAATTACCATGATCCATTTCC	GGAAATGGATCATGGTAATTATCATAAATCATAGCTAGCCTCTGGA
T203I	c.608C>T	TCCCCGATATGATCCATAAACGAGAACAACACG	CGTGTGTTTCTCGTTTATGGGATCATATCGGGGA
I216V	c.646A>G	CGGGACCAGATGTGCAGGAGAAGTTGCCATGCAAG	CTTGATGGCAACTTCTCCTGCACATCTGGTCCCG
A217V	c.650C>T	CAGATGTGCAGGAGAAATTGTCATGCAAGCAAATAATCAC	GTGATTATTTGCTTGCATGACAATTTCTCCTGCACATCTG
R282W	c.844C>T	AGGGGCCTGGCTGGCTAGCCAG	CTGGGCTAGCCAGCCAGGCCCT
Y290C	c.869A>G	GCTAGCCAGAAAGGCTTTTGAATGTGGTGTCAAACA	TGTTTTGACACCACATTCAAAGCCTTCTGGGCTAGC
K293R	c.878A>G		Twist
V304I	c.910G>A	GGGGTCCATCTTCATCTGGGCTTCGGG	CCCGAAGCCCAGATGAAGATGGACCCC
R312C	c.934C>T	CGGGAACGGGGGGTGTGAGGGAGATAAT	ATTATCTCCCTGACACCCCCCGTTTCCCG
D324Y	c.970G>T	TAATTGTGACTGTGATGGCTACACATACAGCATCTACAC	GTGTAGATGCTGTATGTGTAGCCATCACAGTCACAATTA
S325N	c.974G>A	TGATGGCTACACAGACAACATCTACACCATCTCCA	TGGAGATGGTGTAGATGTTGTCTGTGTAGCCATCA

I329V	c.985A>G	AGACAGCATCTACACCGTCTCCATCAGCAGTGC	GCACTGCTGATGGAGACGGTGTAGATGCTGTCT
I329M	c.987C>G	CACAGACAGCATCTACACCATGTCCATCAGCAGT	ACTGCTGATGGACATGGTGTAGATGCTGTCTGTG
S340Y	c.1019C>A	CCTCCCAGCAAGGCCTATACCCCTGGTAC	GTACCAGGGGTATAGGCCTTGCTGGGAGG
S354C	c.1061C>G	CACACTGGCCACCTGTTACAGCAGCGGAG	CTCCGCTGCTGTAACAGGTGGCCAGTGTG
Q363*	c.1087C>T	CGGAGATTACACCGACTAGAGAATCACGAGCGC	GCGCTCGTGATTCTCTAGTCGGTGAATCTCCG
T366M	c.1097C>T		Twist
T375A	c.1123A>G	CTGCACAATGACTGCGCGGAGACGCACACAG	CTGTGTGCGTCTCCGCGCAGTCATTGTGCAG
T377M	c.1130C>T	GCACAATGACTGCACGGAGATGCACACAGGCA	TGCCTGTGTGCATCTCCGTGCAGTCATTGTGC
G390S	c.1168G>A	GCACCTCTGGCTGCTAGCATCTTCGCTC	GAGCGAAGATGCTAGCAGCCAGAGGTGC
A398G	c.1193C>G	CTCTGGCCCTGGAAGGAAACCCAAATCTCAC	GTGAGATTTGGGTTTCCCTCCAGGGCCAGAG
P400S	c.1198C>T	GCCCTGGAAGCAAACCTCAAATCTCACCTGGC	GCCAGGTGAGATTTGAGTTTGCTTCCAGGGC
K428T	c.1283A>C	CAATAACCCTGGATGGAAAACGAATGGAGCAGGCTTG	CAAGCCTGCTCCATTCGTTTTCCATCCAGGGTTATTG
L449P	c.1346T>C	GCTAAATGCCAAAGCTCCGGTGGATTTAGCTGACC	GGTCAGCTAAATCCACCGGAGCTTTGGCATTAGC
V461M	c.1381G>A	AGGACCTGGAGGAGCATGCCTGAGAAGAAAG	CTTTCTTCTCAGGCATGCTCCTCCAGGTCTC
V469I	c.1405G>A	CCTGAGAAGAAAGAGTGTGTTATAAAGGACAATGACTTTGAGC	GCTCAAAGTCATTGTCCTTTATAACACACTCTTTCTTCTCAGG
V469A	c.1406T>C	TGAGAAGAAAGAGTGTGTTGCAAAGGACAATGACTTTGAGC	GCTCAAAGTCATTGTCCTTTGCAACACACTCTTTCTTCTCA
A481T	c.1441G>A	AGCCCAGAGCCCTGAAAACCTAATGGAGAAGTTATC	GATAACTTCTCCATTAGTTTTCAGGGCTCTGGGCT
I487T	c.1460T>C	GCCCTGAAAGCTAATGGAGAAGTTACTACTGAAATTCACAACAAAG	CTTGTTGGAATTTCAAGTATAACTTCTCCATTAGCTTTCAAGGC
T491I	c.1472C>T	CTAATGGAGAAGTTATCATTGAAATTCATAAGAGCTTGTGAAGGAC	GTCCTTCACAAGCTCTTATTGGAATTTCAATGATAACTTCTCCATTAG
A528T	c.1582G>A	CATGTCACACTTACTTCTACTGCTGGAAGTAGCACTG	CAGTGCTAGTTCCAGCAGTAGAAGTAAGTGTGACATG
A529T	c.1585G>A	TCACACTTACTTCTGCTACTGGAAGTAGCACTGTG	CACAGTGCTAGTTCCAGTAGCAGAAGTAAGTGTGA
R541W	c.1621C>T	GCTCTTGGCTGAAAGAGAATGGGATACATCTCCT	AGGAGATGTATCCCATTCTCTTTCAGCCAAGAGC
R541Q	c.1622G>A	TCTTGGCTGAAAGAGAACAGGATACATCTCCTAATGG	CCATTAGGAGATGTATCCTGTTCTCTTTCAGCCAAGA
H557Y	c.1669C>T	TGGGACTTCATGTCTGTTTACACATGGGGAGAGAA	TTCTCTCCCCATGTGTAACAGACATGAAGTCCCA
G582A	c.1745G>C	CTGGAAGAATTCAAAATGAAGCAAGAATTGTGAACTGGAAGCT	AGCTTCCAGTTCACAATTTCTGCTTCATTTTGAATTCTCCAG
R605C	c.1813C>T	AGCATATGAAGCAGCCTTGTGTGTACACGTCCTAC	GTAGGACGTGTACACACAAGGCTGCTTCATATGCT
T608M	c.1823C>T	GCAGCCTCGTGTGTACATGCTCTACAACACTGTTT	GAACAGTGTGTAGGACATGTACACACGAGGCTGC
G619R	c.1855G>A	CACTGTTTCAAGATGACAGAAGAAGGGTGGAGAAGAT	ATCTTCTCCACCCTTCTTCTGTCAATTCTGAACAGTG
T640A	c.1918A>G	GAACCCTAAGGAGAACGCCCTGGTGTCCAAAAG	CTTTTGGACACCAGGGCGTTCTCCTTAGGGTTC

R654Q	c.1961G>A	AGCGTAGGGGGCCAGAGGGATGAGTTG	CAACTCATCCCTCTGGCCCCCTACGCT
E660K	c.1978G>A	GGGATGAGTTGGAGAAGGGAGCCCCCTTCC	GGAAGGGGCTCCCTTCTCCAACATCCC
K692N	c.2076G>C	AAGTCCCCAAGTGCAAACCTCAACATCCCTTATG	CATAAGGGATGTTGAGGTTTGCACTTGGGGACTT
P710R	c.2129C>G	TGGAAAAGCTGAACAAACGTTCCCAGCTTAAAGACTC	GAGTCTTTAAGCTGGGAACGTTTGTTTCAGCTTTTCCA
V727D	c.2180T>A	CAGTCTGTATAATGACTATGTTGATGATTTTATAACACTAAACCTTACAAGC	GCTTGTAAAGTTTAGTGTTATAAAAATCATCAACATAGTCATTATACAGACTG
R740Q	c.2219G>A	GCACAGAGACGACCAGCTGCTTCAAGCTC	GAGCTTGAAGCAGCTGGTCGTCTCTGTGC
V746M	c.2236G>A	GGCTGCTTCAAGCTCTGATGGACATTCTGAATGAG	CTCATTCAGAATGTCCATCAGAGCTTGAAGCAGCC

**Tableau S5. Liste des amorces utilisées pour le séquençage Sanger des plasmides du projet *PCSK1*.**

Région du plasmide ciblée	Sens de l'amorce	Séquence de l'amorce (5'->3')
<i>PCSK1</i>	<i>Forward</i>	GCTCGTTTAGTGAACCGTCAG
	<i>Reverse</i>	GTTTGGAGTGGAATCACACG
	<i>Forward</i>	CCAAGCTGGACCTTCATGT
	<i>Reverse</i>	GTGCAGTCATTGTGCAGGTC
	<i>Forward</i>	CACCATCTCCATCAGCAGTG
	<i>Reverse</i>	CTCTTTCAGCCAAGAGCACA
	<i>Forward</i>	GCTATCAAGTCCCTGGAGCA
	<i>Reverse</i>	GGGATGTTGAGCTTTGCAC
	<i>Forward</i>	AACACCCTGGTGTCCAAAAG
	<i>Reverse</i>	CCGGCCGTTTAAACCTTATC

**Tableau S6. Liste des amorces utilisées lors des expériences de qPCR pour le projet *PCSK1*.**

Gène ciblé	Sens de l'amorce	Séquence de l'amorce (5'->3')	Température
<i>PCSK1</i>	<i>Forward</i>	GAACAAACACGGGACCAGAT	60°C
	<i>Reverse</i>	TTCAAAGCCTTCTGGGCTA	60°C
	<i>Forward</i>	ACCATCTCCATCAGCAGTG	60°C
	<i>Reverse</i>	TCTCCGTGCAGTCATTGTG	60°C
<i>RPLP0</i>	<i>Forward</i>	ACACCTTCCCCTTGCTGA	60°C
	<i>Reverse</i>	CTCTTCCTTGGCTTCAACCT	60°C

**Tableau S7. Cycle de température utilisé au cours des expériences de qPCR.**

Hold Stage	PCR stage		Melting curve		
<b>1X</b>	<b>40X</b>		<b>1X</b>		
50°C	95°C	60°C	95°C	60°C	95°C
2 min	15 sec	1 min	15 sec	1 min	15 sec

**Tableau S8 : Expression protéique des variants perte de fonction partielle de *PCSK1* (Groupe\_B) dans les lysats cellulaires par western blotting.** Ce tableau récapitule l'expression protéique des variants de *PCSK1* du Groupe\_B par western blotting. Taille des formes précurseurs de PC1/3 : 94kDa et 87kDa ; taille de la forme active de PC1/3 : 74/66 kDa. Trois expériences indépendantes ont été réalisées pour chaque mutation (les images des western blotting sont disponibles en **Annexe 2, Figure S2**). *N<sub>nwt</sub>*, nombre de participants avec un poids normal portant des mutations dans l'étude RaDiO ; *N<sub>ob</sub>*, nombre de participants atteints d'obésité portant des mutations dans l'étude RaDiO ; *N<sub>owt</sub>*, nombre de participants atteints de surpoids portant des mutations dans l'étude RaDiO.

	Variants de <i>PCSK1</i> perte de fonction partielle (Groupe_B)											
Variants	WT	L58S	R80Q	M125I	T157M	Y181H	K293R	R312C	S340Y	T366M	A528T	R605C
<i>N<sub>ob</sub></i> / <i>N<sub>owt</sub></i> / <i>N<sub>nwt</sub></i>	-	0/1/0	0/1/0	2/6/5	1/0/0	1/1/2	1/0/0	0/1/0	0/1/0	0/1/0	1/0/1	0/0/1
Formes précurseurs	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
Forme active	Oui	Oui	Oui	Oui	Non	Non	Non	Non	Oui	Non	Oui	Oui
Exemples												
kD												
100 —												
70 —												
55 —												

**Tableau S9. Analyse de la coségrégation des familles portant les mutations p.R81S, p.L132S, p.V304I, p.Q363\*, c.1095+1G>A et p.T375A de PCSK1 avec le surpoids et l'obésité.** Le tableau montre la coségrégation entre des variants du Groupe\_A de PCSK1 (p.R81S, p.L132S, p.V304I, p.Q363\*, c.1095+1G>A et p.T375A) et l'adiposité dans 6 familles. Les différents membres des familles ont été séquencés par séquençage Sanger. *DT2*, diabète de type 2; *WT*, wild-type.

Mutation	Exon	Statut	Parenté	Phénotype
c.243G>C / p.R81S	2	Hétérozygote	Cas index RaDiO	Obésité
		Hétérozygote	Enfant	Obésité
		WT	Enfant	Obésité
		Hétérozygote	Enfant	Obésité
		Hétérozygote	Enfant	Obésité
		WT	Petit enfant	Surpoids
c.395T>C / p.L132S	3	WT	Père	Obésité
		Hétérozygote	Cas index RaDiO	Obésité
		Hétérozygote	Mère	Poids Normal
c.910G>A / p.V304I	8	Hétérozygote	Cas index RaDiO	Poids Normal
		Hétérozygote	Cas index RaDiO	Obésité
		WT	Frère	Obésité
c.1087C>T / p.Q363*	8	Hétérozygote	Cas index RaDiO	Surpoids
		Hétérozygote	Frère	Obésité
c.1095+1G>A / p.?	8-9	Hétérozygote	Cas index RaDiO	Surpoids
		WT	Sœur	DT2 et Obésité
c.1123A>G / p.T375A	9	Hétérozygote	Cas index RaDiO	Obésité
		WT	Frère	Obésité
		WT	Père	Obésité
		Hétérozygote	Mère	Poids Normal

**Tableau S10. Liste des amorces utilisées pour le séquençage Sanger de DYRK1B sur l'ADN des familles disponibles.**

Exon ciblé de DYRK1B	Sens de l'amorce	Séquence de l'amorce (5'->3')
Exon 6	<i>Forward</i>	TTTGACAGAGACTTGGAAAGAGCAT
	<i>Reverse</i>	CCGCCCTAAGCCTCCCGTTG
Exon 8	<i>Forward</i>	TGCTCCTGGGCACACCCTACGAC
	<i>Reverse</i>	ACACACGGGGAAACAGAGGCACAAG

**Tableau S11. Cycle de température utilisé au cours de la PCR précédant le séquençage Sanger de DYRK1B dans les familles disponibles.** Les échantillons ont été amplifiés selon le protocole Fast Start Taq DNA Polymerase (Roche) avec une concentration de 2,5 mM de MgCl<sub>2</sub>.

1X	40X			1X
95°C	95°C	63°C	72°C	72°C
5 min	30 sec	30 sec	45 sec	10 min

**Tableau S12. Liste des amorces utilisées pour la réalisation des PCR de mutagenèse des 63 variants du projet *DYRK1B*.**

Variants		Séquence de l'amorce <i>forward</i> (5'->3')	Séquence de l'amorce <i>reverse</i> (5'->3')
V3I	c.7G>A	TCGCCATGGCCATCCCACCGGGC	GCCCCGGTGGGATGGCCATGGCGA
P5L	c.14C>T	GGCCGTCCCCTGGGCCATGGTC	GACCATGGCCCAGTGGGACGGCC
R27Q	c.80G>A	TATTGCCTGATGTGCAGCTACTGCCTCGGAG	CTCCGAGGCAGTAGCTGCACATCAGGCAATA
R31P	c.92G>C	GCTACTGCCTCCGAGGCTGCCCC	GGGGCAGCCTCGGAGGCAGTAGC
A40S	c.118G>T	GGCCTTCCGGGATTCAACCTCAGCCCC	GGGGCTGAGGTTGAATCCCGGAAGGCC
K57R	c.170A>G	GTGGACCTCATCAAGACCTACAGGCACATCAATGAG	CTCATTGATGTGCCTGTAGGTCTTGATGAGGTCCAC
K68Q	c.202A>C	TATACTATGCGAAGAAGCAGCGGGCGGCC	GGGCCC GCCGCTGCTTCTTCGCATAGTATA
R70Q	c.209G>A	AGAAGAAGCGGCAGGCCAGCAGGC	GCCTGCTGGGCCTGCCGCTTCTTCT
S79L	c.236C>T	GGCGCCACCCCAGGATTTGAGCAACAAGAAG	CTTCTTGTGCTCAAATCCTGGGGTGGCGCC
K86del	c.256_258del	GCAACAAGAAGGAGAAGGTCCTGAACCATGGTTA	TAACCATGGTTCAGGACCTTCTCCTTCTTGTTC
R102H	c.305G>A	GACTACATCGTGCACAGTGGCGAGCGC	GCGCTCGCCACTGTGCACGATGTAGTC
G120V	c.359G>T	CGCTCATTGGCAAAGTCTCCTTTGGCCAGGT	ACCTGGCCAAAGGAGACTTTGCCAATGAGCG
H131Y	c.391C>T	CCAGGTGGTAAAAGCCTATGATTATCAGACCCAGG	CCTGGGTCTGATAATCATAGGCTTTCACCACCTGG
R157Q	c.470G>A	CCAGATTGAGCTGCAGCTGCTGGAGCTGA	TCAGCTCCAGCAGCTGCAGCTCAATCTGG
T167M	c.500C>T	GGAGCTGATGAACCAGCATGACATGGAGATGAAGTA	TACTTCATCTCCATGTCATGCTGGTTCATCAGCTCC
M169R	c.506T>G	ACCAGCATGACACGGAGAGGAAGTACTATATAGTACAC	GTGTACTATATAGTACTTCTCCTCCTCGTGCATGCTGGT
Y172C	c.515A>G	ACACGGAGATGAAGTACTGTATAGTACACCTGAAGCG	CGCTTCAGGTGTACTATACAGTACTTCATCTCCGTGT
L176M	c.526C>A	AGATGAAGTACTATATAGTACACATGAAGCGGCACTTCATGT	ACATGAAGTGCCGCTTCATGTGTACTATATAGTACTTCATCT
H179L	c.536A>T	GTACACCTGAAGCGGCTTTCATGTTCCGGAAC	GTTCCGGAACATGAAGAGCCGCTTCAGGTGTAC
T223M	c.668C>T	CGCAGCAGCTCTGCATGGCACTGCTCTT	AAGAGCAGTGCCATGCAGAGCTGCTGCG
N249S	c.746A>G	CCGAAAACATCTTGCTGTGCAGCCCCAAGCGCAG	CTGCGCTTGGGGCTGCACAGCAAGATGTTTTCGG
D259N	c.775G>A	CGCCATCAAGATTGTGAACTTCGGCAGCTCCTG	CAGGAGCTGCCGAAGTTCACAATCTTGATGGCG
P282L	c.845C>T	CTTCTACCGCTCACTTGAGGTGCTCCTGG	CCAGGAGCACCTCAAGTGAAGCGGTAGAAG
N323D	c.967A>G	GAGGTCGACCAGATGGACCGCATTGTGGAGG	CCTCCACAATGCGGTCCATCTGGTGCACCTC
R324L	c.971G>T	GTCGACCAGATGAACCTCATTGTGGAGGTGCTG	CAGCACCTCCACAATGAGGTTTCATCTGGTGCAC
A335T	c.1003G>A	ATCCCACCGGCCACCATGCTGGACC	GGTCCAGCATGGTGGCCGGTGGGAT

R344C	c.1030C>T	CCAGGCGCCCAAGGCTTGCAAGTACTTTGA	TCAAAGTACTTGCAAGCCTTGGGCGCCTGG
R344H	c.1031G>A	GGCGCCCAAGGCTCACAAGTACTTTGAACG	CGTTCAAAGTACTTGTGAGCCTTGGGCGCC
R349W	c.1045C>T	GGCTCGCAAGTACTTTGAATGGCTGCCTGG	CCAGGCAGCCATTCAAAGTACTTGCAGCC
R349Q	c.1046G>A	CGCAAGTACTTTGAACAGCTGCCTGGGGG	CCCCCAGGCAGCTGTTCAAAGTACTTGCG
G352R	c.1054G>A	GTACTTTGAACGGCTGCCTAGGGGTGGCTG	CAGCCACCCCTAGGCAGCCGTTCAAAGTAC
G352A	c.1055G>C	ACGGCTGCCTGCGGGTGGCTGGA	TCCAGCCACCCCGAGGCAGCCGT
G353C	c.1057G>T	GGCTGCCTGGGTGTGGCTGGACC	GGTCCAGCCACACCCAGGCAGCC
R358*	c.1072C>T	GGTGGCTGGACCCTATGAAGGACGAAAGAAC	GTTCTTTTCGTCTTCATAGGGTCCAGCCACC
R358Q	c.1073G>A	GGTGGCTGGACCCTACAAAGGACGAAAGAACTC	GAGTTCTTTTCGTCTTTGTAGGGTCCAGCCACC
T360K	c.1079C>A	CTGGACCCTACGAAGGAAGAAAGAACTCAGGAAGG	CCTTCCTGAGTTCTTTCTTCCTTCGTAGGGTCCAG
G371R	c.1111G>A	ACCAGGGCCCCAGGACACGGCGG	CCGCCGTGTCCTGGGGCCCTGGT
G371R	c.1111G>C	CCAGGGCCCCCGGACACGGCG	CGCCGTGTCCGGGGGCCCTGG
A399D	c.1196C>A	GCCACAGCCCCGACGACTACCTCCG	CGGAGGTAGTCGTCCGGGGCTGTGGC
R403H	c.1208G>A	GCCGACTACCTCCACTTCCAGGACCTG	CAGGTCCTGGAAGTGGAGGTAGTCGGC
R410H	c.1229G>A	GGACCTGGTGCTGCACATGCTGGAGTATG	CATACTCCAGCATGTGCAGCACCAGGTCC
A418T	c.1252G>A	TATGAGCCCGCCACCCGCATCAGCC	GGCTGATGCGGGTGGCGGGCTCATA
G429S	c.1285G>A	GGCTCTGCAGCACAGCTTCTTCCGCCG	CGGCGGAAGAAGCTGTGCTGCAGAGCC
G429R	c.1285G>C	GCTCTGCAGCACCGCTTCTTCCGCC	GGCGGAAGAAGCGGTGCTGCAGAGC
R432H	c.1295G>A	CGGCTTCTTCCACCGCACGGCCG	CGGCCGTGCGGTGGAAGAAGCCG
P443L	c.1328C>T	CCAACACGGGCCTGGCAGGCAGCAG	CTGCTGCCTGCCAGGCCCGTGTGG
S446G	c.1336A>G	GCCCGGCAGGCGGCAGTGCCTCC	GGAGGCACTGCCGCCTGCCGGGC
S447R	c.1341T>A	GGCAGGCAGCAGAGCCTCCACCTCG	CGAGGTGGAGGCTCTGCTGCCTGCC
T450I	c.1349C>T	CAGTGCCTCCATCTCGCCCGCGC	GCGCGGGCGAGATGGAGGCACTG
A453V	c.1358C>T	CACCTCGCCCGTGCCTCGACA	TGTCGAGGGGCACGGGCGAGGTG
S462R	c.1386C>G	CCCCTCTTCCAGGACCGCCAGCTCC	GGAGCTGGCGGTCTGGAAGAGGGG
G472S	c.1414G>A	CATCTCCAGTTCTGGAAGCTCCAGTGGCTCC	GGAGCCACTGGAGCTTCCAGAAGTGGAGATG
R481W	c.1441C>T	CCTCCAGTGACAAGTGGACCTACCGCTAC	GTAGCGGTAGGTCCAGTTGCTACTGGAGG
R484C	c.1450C>T	GACAACCGGACCTACTGCTACAGCAACCGAT	ATCGGTTGCTGTAGCAGTAGTCCGTTGTC
R488Q	c.1463G>A	CCGCTACAGCAACCAATATTGTGGGGGCC	GGCCCCACAATATTGTTGCTGTAGCGG

C490F	c.1469G>T	CGCTACAGCAACCGATATTTTGGGGGCCCT	AGGGCCCCCAAATATCGGTTGCTGTAGCG
C490W	c.1470T>G	CAGCAACCGATATTGGGGGGGCCCTGG	CCAGGGCCCCCAATATCGGTTGCTG
G494E	c.1481G>A	GTGGGGGCCCTGAGCCCCCTATCAC	GTGATAGGGGGCTCAGGGCCCCAC
P559S	c.1675C>T	CCAACCTCACCATCACCCCGGAGC	GCTCCGGGGGTGATGGTGAGGTTGG
P578S	c.1732C>T	GACTGCTCCCCATCTCACCCAGCGC	GCGCTGGGTGAGATGGGAGCAGTC
A581V	c.1742C>T	ACCTCACCCAGTGCCTGCCCC	GGGGGCAGGCACTGGGTGAGGT
R600H	c.1799G>A	ATGACTGGAGGTCATCCACCCCTCCCG	CGGGAGGGGTGGATGACCTCCAGTCAT
D608V	c.1823A>T	CCGCCTCCTGATGTCCCTGCCACTCTG	CAGAGTGGCAGGGACATCAGGAGGCGG
P614T	c.1840C>A	GCCACTCTGGGACTCACCTGGGCC	GGCCCAGGTGAGTCCCCAGAGTGGC
R619C	c.1855C>T	CACCTGGGCCTCTGTGGTGTACCC	GGGGTACACCACAGAGGCCAGGTG

**Tableau S13. Liste des amorces utilisées pour le séquençage Sanger des plasmides *DYRK1B*.**

Région du plasmide ciblée	Sens de l'amorce	Séquence de l'amorce (5'->3')
<i>DYRK1B</i>	<i>Forward</i>	GCTCGTTTAGTGAACCGTCAG
	<i>Reverse</i>	CCGGCCGTTTAAACCTTATC
	<i>Forward</i>	GGATTCGAGCAACAAGAAGG
	<i>Reverse</i>	CAGAGTGGCAGGGTCATCA
	<i>Forward</i>	CACGGCACTGCTCTTTCTG
	<i>Reverse</i>	ATCGGTTGCTGTAGCGGTAG
	<i>Forward</i>	GACCCTACGAAGGACGAAAG
	<i>Reverse</i>	CAAAGTACTTGCGAGCCTTG
	<i>Forward</i>	CCAGCTCCATCTCCAGTTCT
	<i>Reverse</i>	TCATGCTGGTTCATCAGCTC

**Tableau S14. Liste des amorces utilisées lors des expériences de qPCR pour le projet *DYRK1B*.**

Gène ciblé	Sens de l'amorce	Séquence de l'amorce (5'->3')	Température
<i>DYRK1B</i>	<i>Forward</i>	CGTACAGGTTGTAGGACAGC	60°C
	<i>Reverse</i>	GAACCAGGCCAGATTGAG	60°C
<i>G6PC</i>	<i>Forward</i>	CGGAAGTGTGCTGTAGTAGT	60°C
	<i>Reverse</i>	GCATTAACTCCTTTGGGTAGC	60°C
<i>RPLP0</i>	<i>Forward</i>	ACACCTTCCCCTTGCTGA	60°C
	<i>Reverse</i>	CTCTTCCTTGCTTCAACCT	60°C

**Tableau S15. Evaluation des critères ACMG pour chaque variant rare du gène *DYRK1B* identifié dans RaDiO. *PVS1* : preuve très forte de pathogénicité 1 ; *PS3* : preuve forte de pathogénicité 3 ; *PM2* : preuve modérée de pathogénicité 2 ; *PM4* : preuve modérée de pathogénicité 4 ; *PP3* : preuve secondaire de pathogénicité 3 ; *PP5* : preuve secondaire de pathogénicité 5 ; *P/LP* : variant pathogène ou probablement pathogène à la suite d'une combinaison des critères ACMG.**

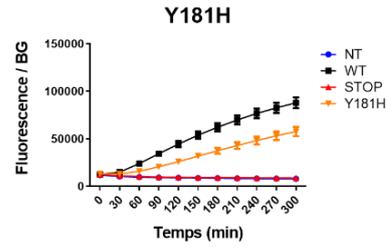
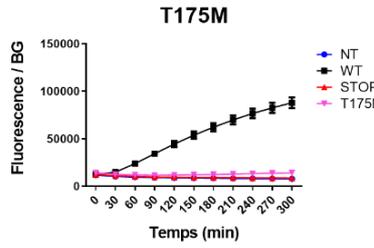
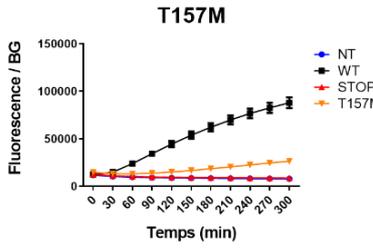
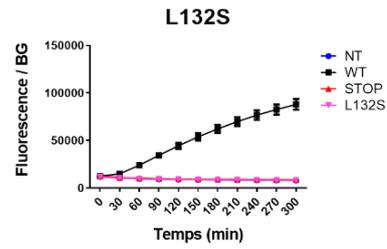
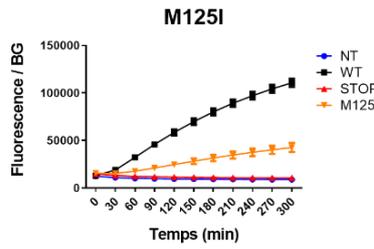
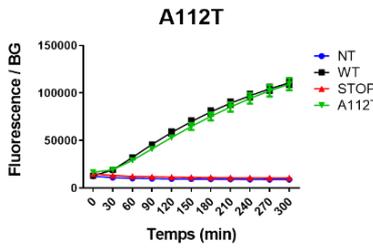
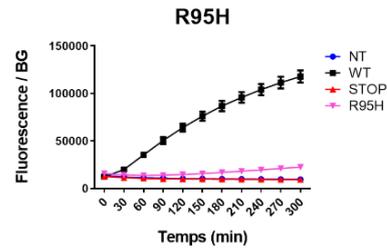
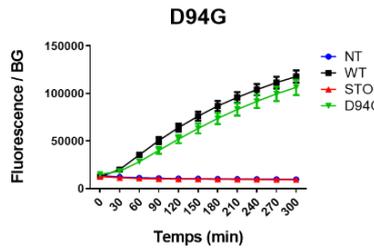
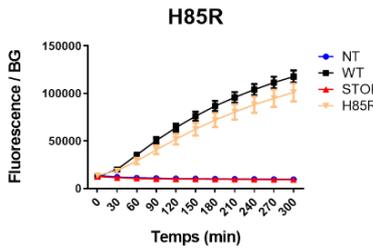
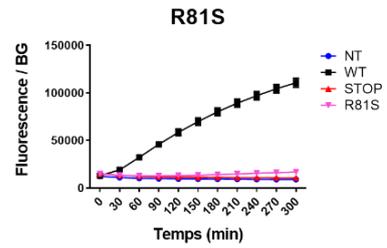
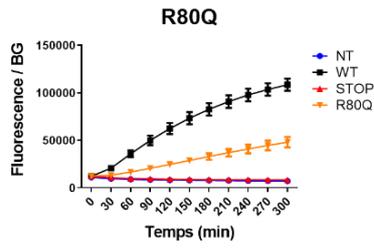
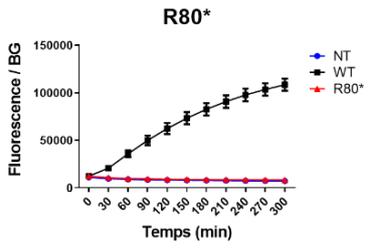
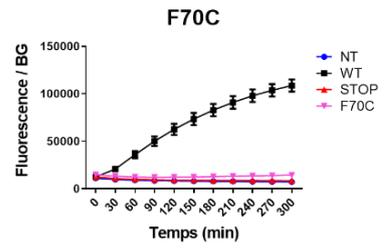
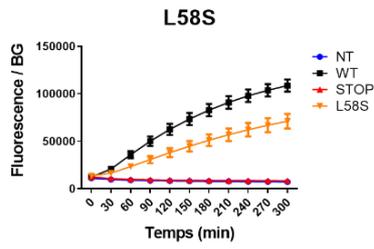
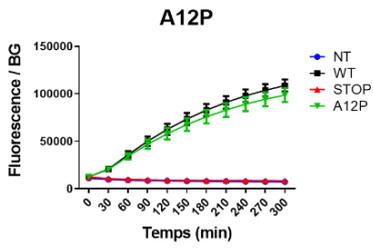
Rare variants in <i>DYRK1B</i> (NM_004714.3)	PVS1	PS3	PM2	PM5	PP3	PP5	Catégorie
c.7G>A, p.V3I							Neutre
c.14C>T, p.P5L		1	1				<b>P/LP</b>
.80G>A, p.R27Q							Neutre
c.92G>C, p.R31P			1				Neutre
c.118G>T, p.A40S			1				Neutre
c.170A>G, p.K57R			1				Neutre
c.202A>C, p.K68Q		1				1	Neutre
c.209G>A, p.R70Q		1					Neutre
c.236C>T, p.S79L			1				Neutre
c.256_258del, p.K86del		1	1				<b>P/LP</b>

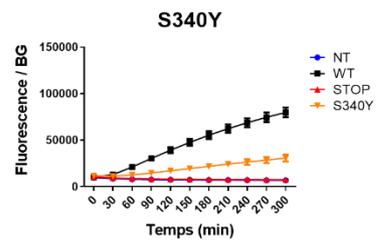
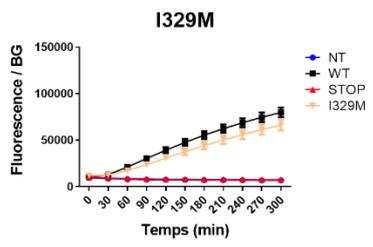
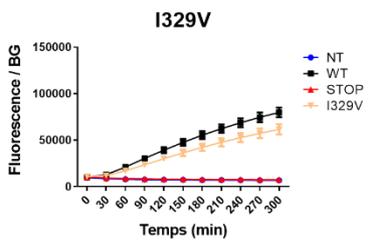
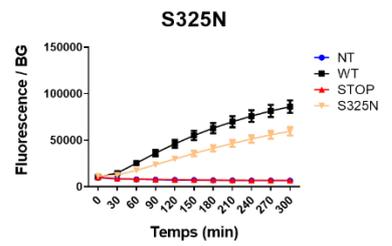
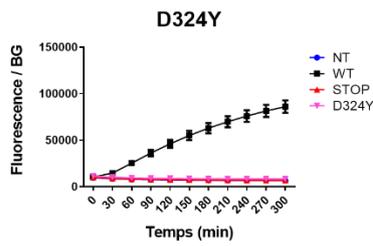
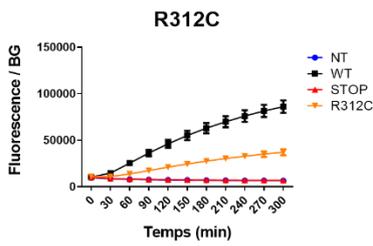
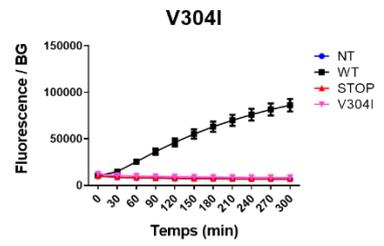
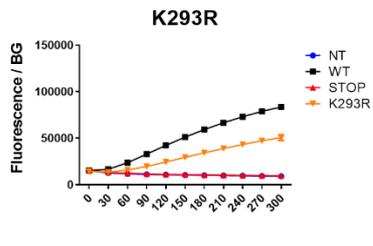
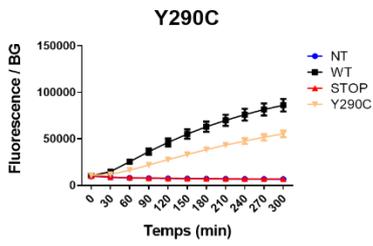
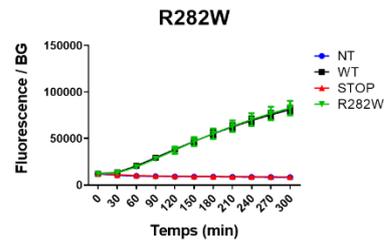
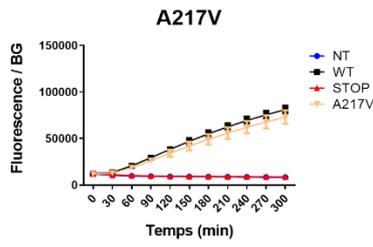
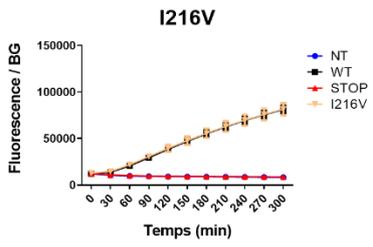
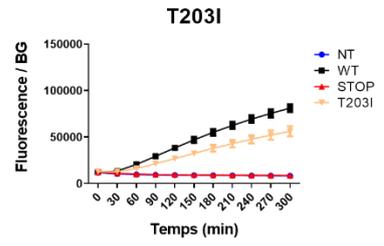
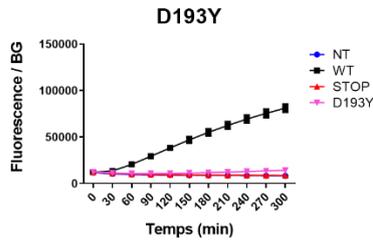
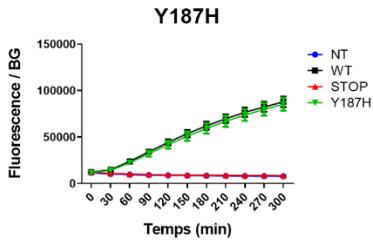
c.305G>A, p.R102H		1		1			P/LP
c.359G>T, p.G120V		1	1		1		P/LP-full
c.391C>T, p.H131Y		1	1				P/LP
c.470G>A, p.R157Q		1	1				P/LP
c.500C>T, p.T167M		1					Neutre
c.506T>G, p.M169R			1				Neutre
c.515A>G, p.Y172C			1				Neutre
c.526C>A, p.L176M		1					Neutre
c.536A>T, p.H179L		1	1		1		P/LP-full
c.668C>T, p.T223M		1					Neutre
c.746A>G, p.N249S			1				Neutre
c.775G>A, p.D259N		1	1		1		P/LP-full
c.845C>T, p.P282L		1	1		1		P/LP-full
c.967A>G, p.N323D		1	1				P/LP
c.971G>T, p.R324L		1					Neutre
c.1003G>A, p.A335T			1				Neutre
c.1030C>T, p.R344C							Neutre
c.1031G>A, p.R344H							Neutre
c.1045C>T, p.R349W		1	1				P/LP-full
c.1046G>A, p.R349Q			1				Neutre
c.1054G>A, p.G352R		1	1				P/LP
c.1055G>C, p.G352A							Neutre
c.1057G>T, p.G353C		1	1				P/LP
c.1072C>T, p.R358*	1	1	1				P/LP-full
c.1073G>A, p.R358Q		1	1				P/LP
c.1079C>A, p.T360K		1	1				P/LP
c.1111G>A, p.G371R		1	1				P/LP
c.1111G>C, p.G371R		1	1				P/LP
c.1196C>A, p.A399D			1				Neutre
c.1208G>A, p.R403H			1				Neutre
c.1229G>A, p.R410H		1	1				P/LP
c.1252G>A, p.A418T							Neutre
c.1285G>A, p.G429S							Neutre
c.1285G>C, p.G429R			1				Neutre
c.1295G>A, p.R432H			1				Neutre
c.1328C>T, p.P443L			1				Neutre
c.1336A>G, p.S446G							Neutre
c.1341T>A, p.S447R		1	1				P/LP
c.1349C>T, p.T450I			1				Neutre
c.1358C>T, p.A453V			1				Neutre
c.1386C>G, p.S462R							Neutre
c.1414G>A, p.G472S			1				Neutre
c.1441C>T, p.R481W			1				Neutre
c.1450C>T, p.R484C							Neutre
c.1463G>A, p.R488Q							Neutre
c.1469G>T, p.C490F			1				Neutre
c.1470T>G, p.C490W							Neutre
c.1481G>A, p.G494E							Neutre
c.1675C>T, p.P559S			1				Neutre
c.1732C>T, p.P578S							Neutre
c.1742C>T, p.A581V			1				Neutre
c.1799G>A, p.R600H			1				Neutre
c.1823A>T, p.D608V			1				Neutre
c.1840C>A, p.P614T							Neutre
c.1855C>T, p.R619C		1					Neutre

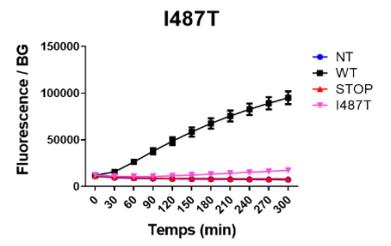
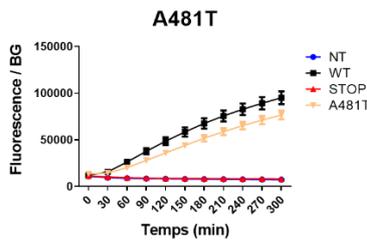
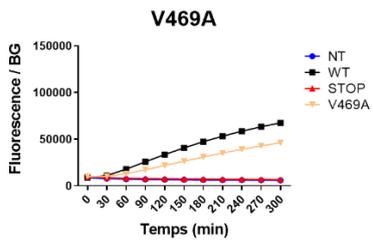
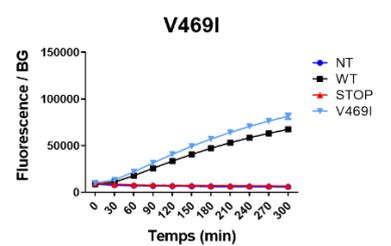
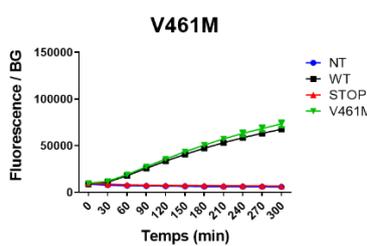
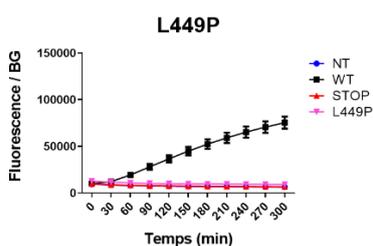
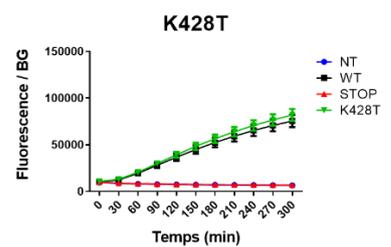
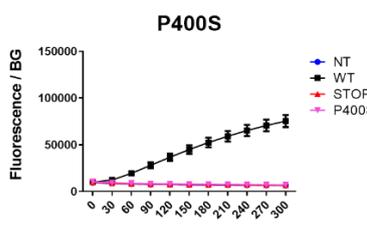
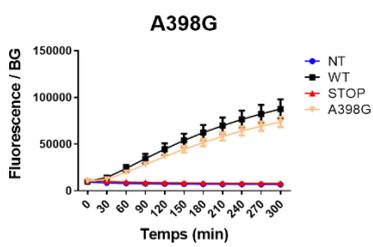
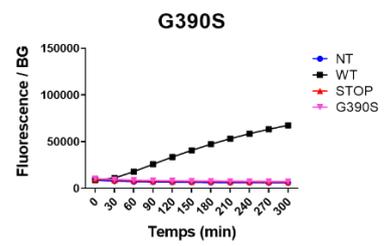
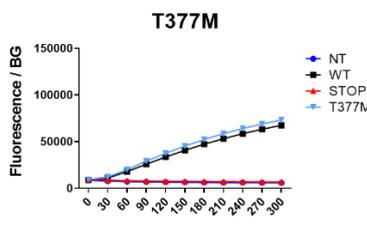
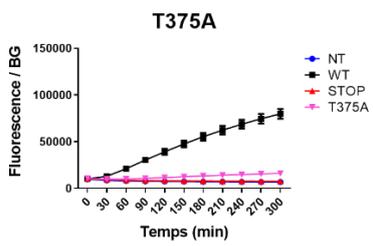
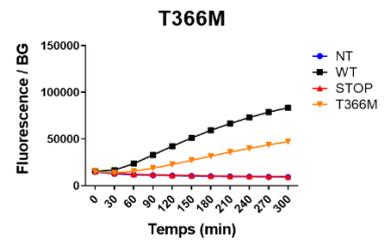
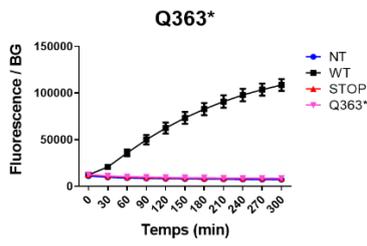
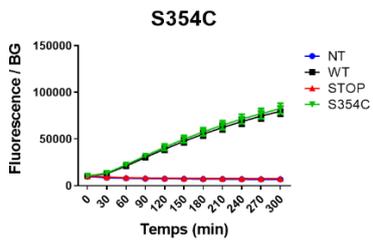
**Tableau S16 : Mutations nulles de *DYRK1B* (NM\_004714.3) détectées dans les études 52K et TOPMed.** Parmi ces variants, deux variants venant de la 52K et sept variants nuls venant de TOPMed ont été conservés pour les analyses d'association avec le risque de DT2.

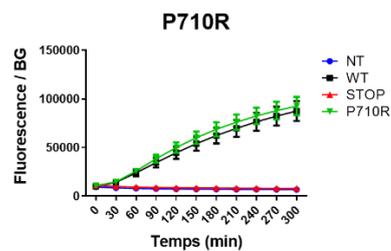
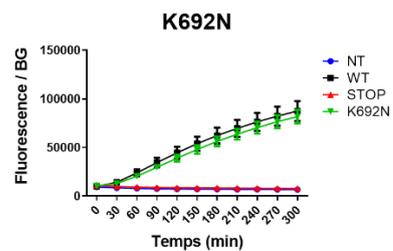
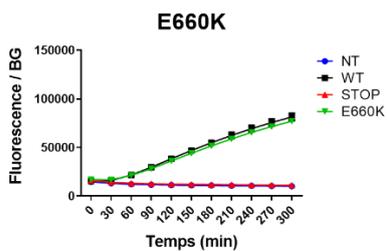
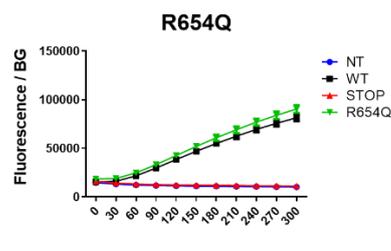
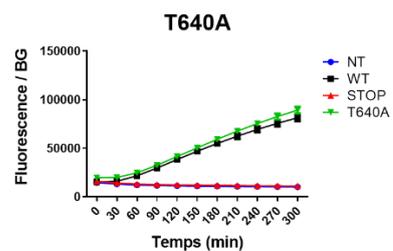
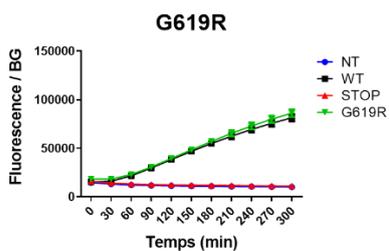
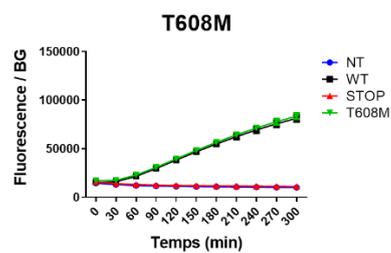
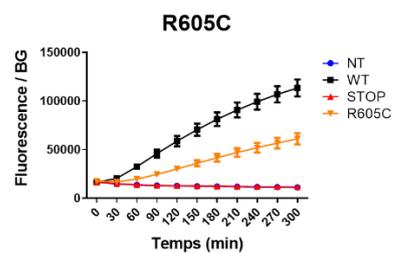
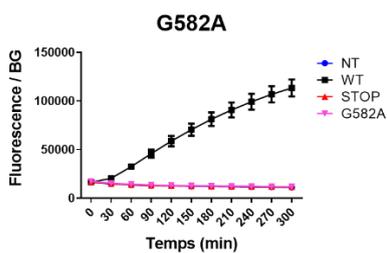
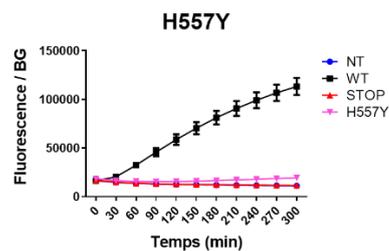
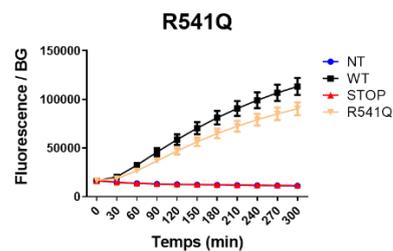
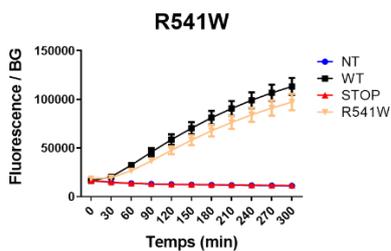
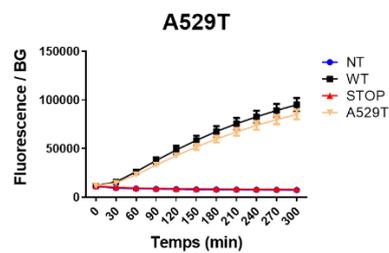
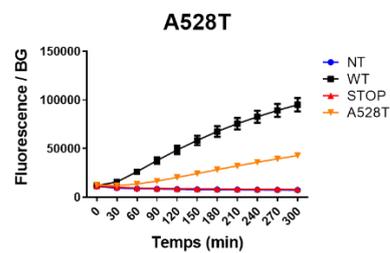
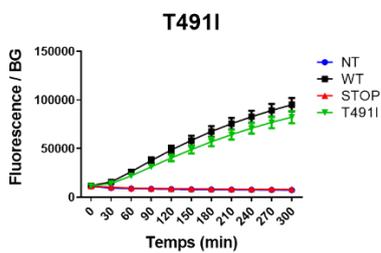
<b>Chr</b>	<b>Position (Hg38)</b>	<b>Mutation</b>	<b>52K</b>	<b>TOPMed</b>
<b>19</b>	40321205	c.184-3_184-2insGGGC		X
<b>19</b>	40318298	c.808-2A>C		X
<b>19</b>	40318281	c.823C>T, p.Gln275Ter	X	X
<b>19</b>	40316889	c.1449C>G, p.Tyr483Ter	X	X
<b>19</b>	40316876	c.1462C>T, p.Arg488Ter		X
<b>19</b>	40316713	c.1528_1531del, p.Gln511ArgfsTer52		X
<b>19</b>	40316611	c.1633del, p.Gln545SerfsTer19		X
<b>19</b>	40316611	c.1633_1634insC, p.Gln545ProfsTer30		X
<b>19</b>	40316491	c.1753_1754insC, p.Gln585ProfsTer23		X

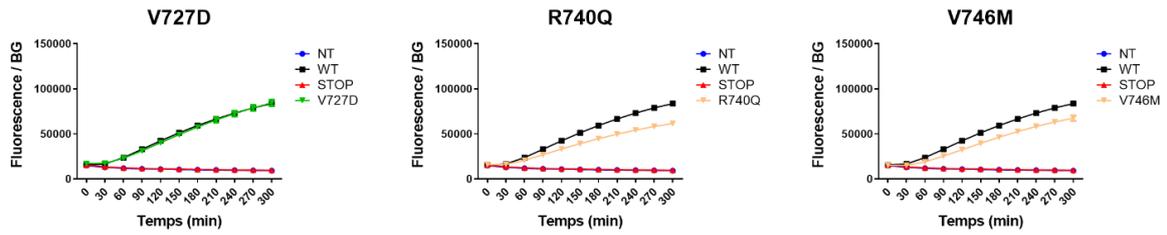
## **ANNEXE 2. Figures supplémentaires**



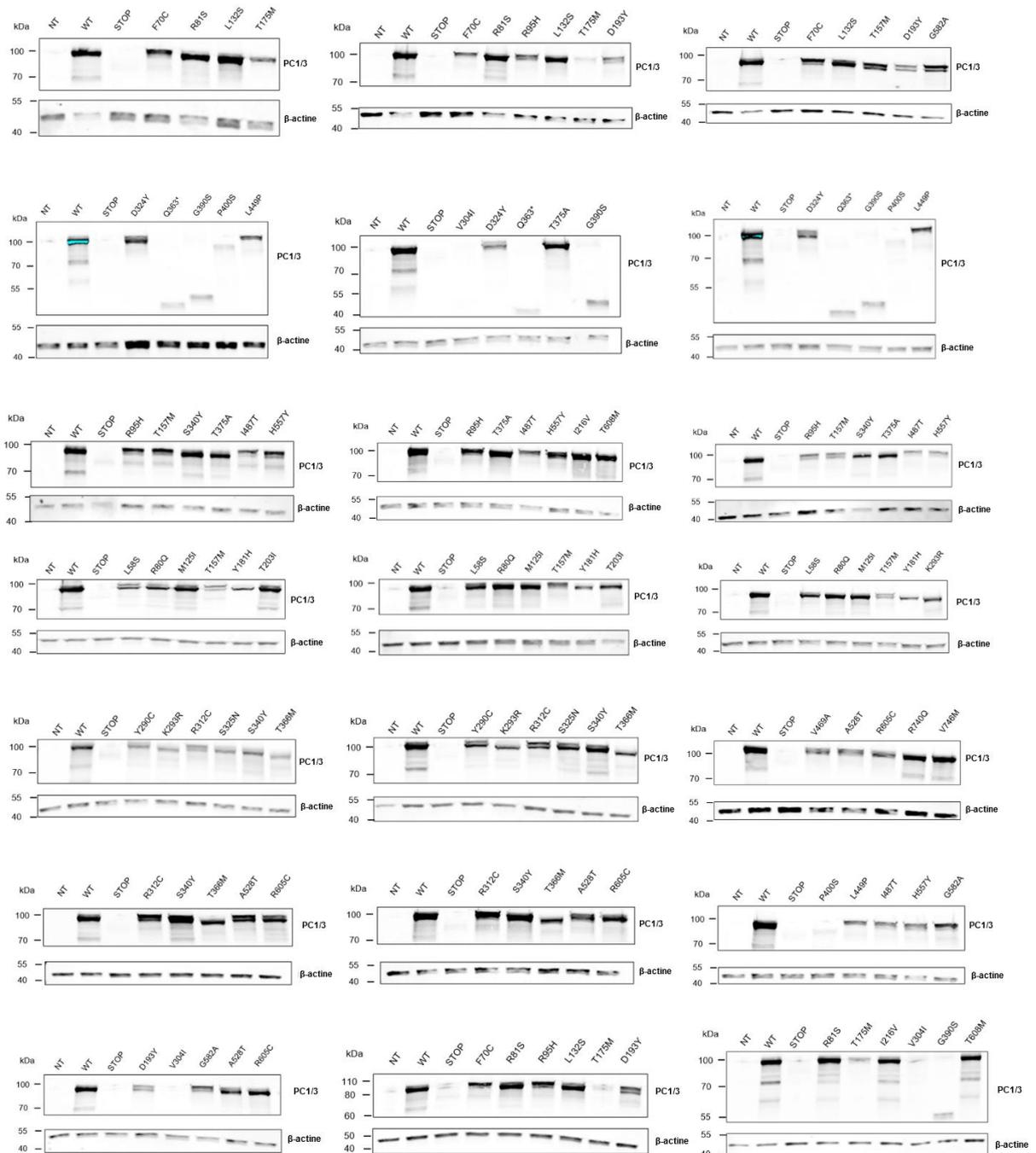




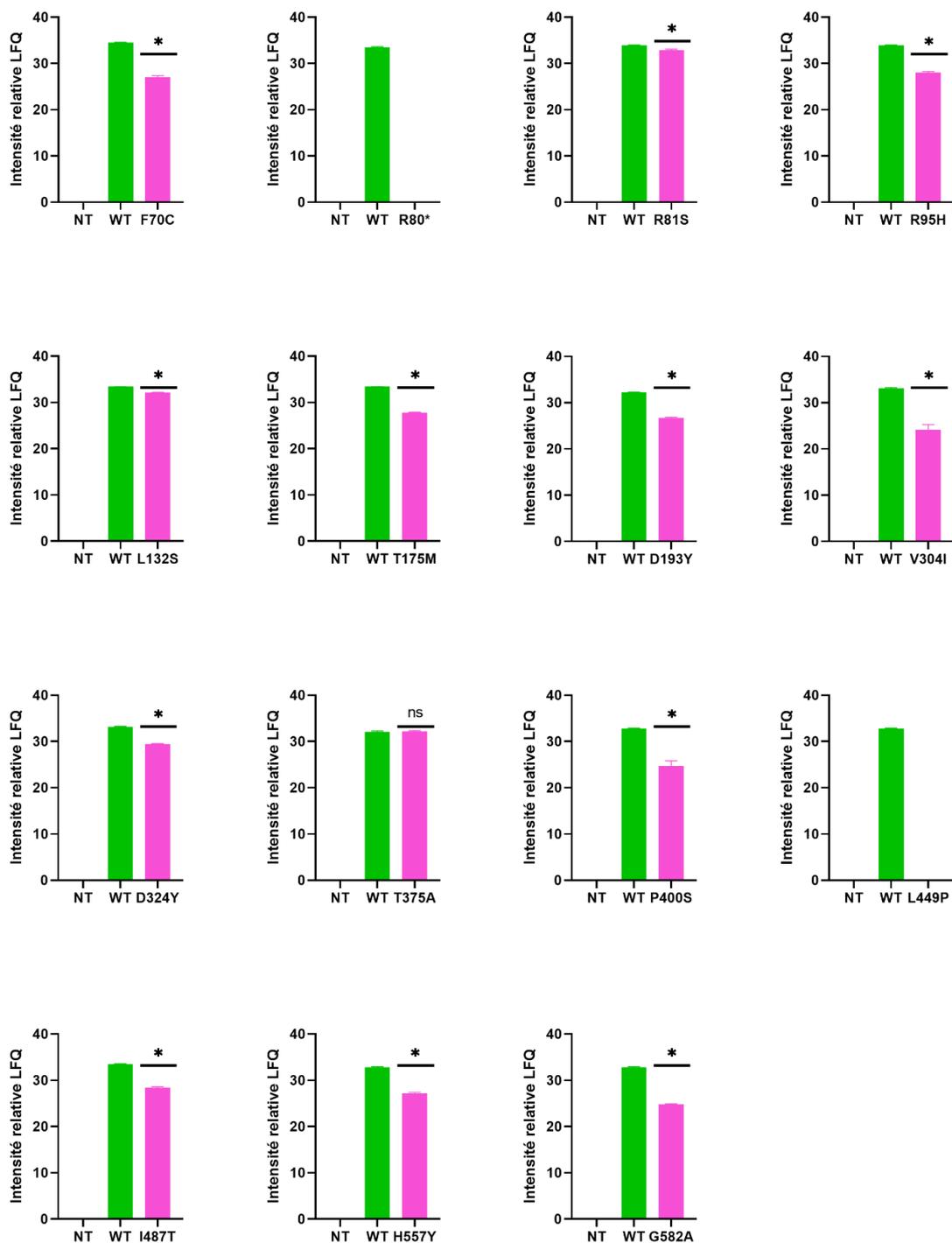




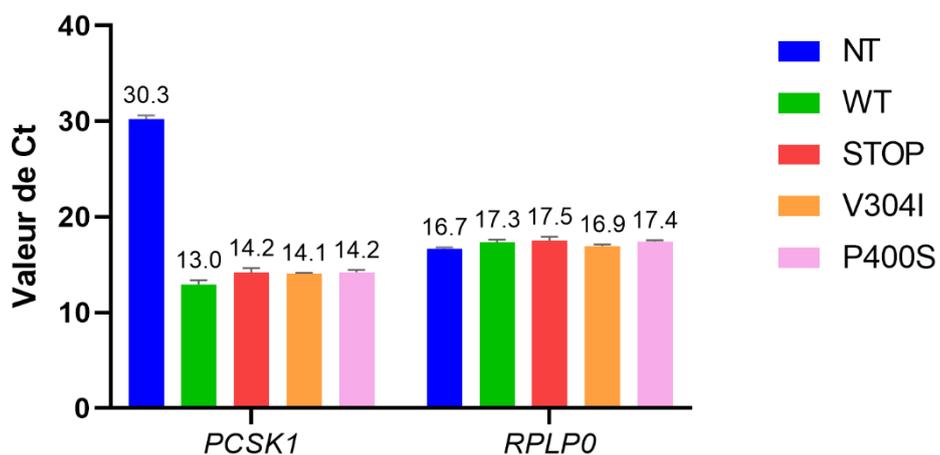
**Figure S1 : Activité enzymatique des variants non synonymes et non-sens de *PCSK1*.** Les figures montrent l'activité enzymatique de PC1/3 normalisée par l'activité de la  $\beta$ -galactosidase dans les cellules HEK293 transfectées ou non (condition non transfectée [NT]) avec le plasmide *PCSK1* wild-type ou muté. Cinq groupes ont été déterminés selon leur activité enzymatique par *machine learning* (**Figure 13**). Les variants roses sont inclus dans le Groupe\_A ; les variants oranges sont inclus dans le Groupe\_B ; les variants jaunes sont inclus dans le Groupe\_C ; les variants verts sont inclus dans le Groupe\_D et les variants bleus dans le Groupe\_E. Les conditions NT, WT et STOP sont représentées en bleu, noir et rouge respectivement. La condition STOP correspond à l'activité du variant codant p.R80\*. Les données sont la moyenne  $\pm$  SEM de quatre expériences indépendantes réalisées en triplicat technique. *NT*, non-transfecté ; *WT*, wild-type.



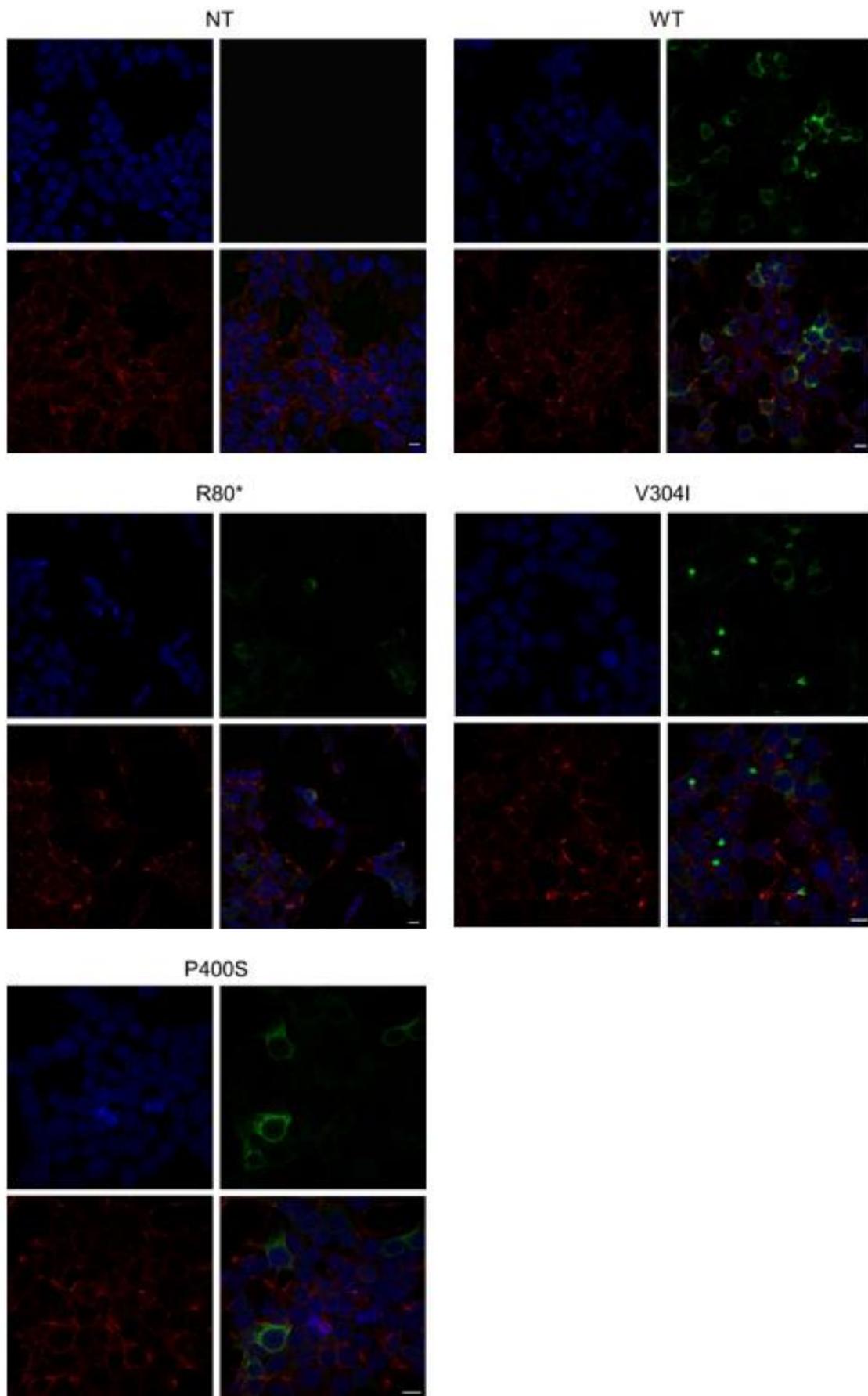
**Figure S2 : Expression protéique des variants perte de fonction de *PCSK1* (Groupe\_A et Groupe\_B) dans le lysat cellulaire par western blotting.** Expression protéique de PC1/3 dans les lysats de cellules HEK293 transfectées ou non (condition non-transfectée [NT]) avec des plasmides contenant *PCSK1* WT ou un des variants du Groupe\_A ou Groupe\_B. Les formes précurseurs font une taille de 94kDa et 87kDa alors que la forme active de PC1/3 fait une taille de 74/66 kDa. La condition STOP correspond à l'expression protéique du variant non-sens codant p.R80\*. L'expression a été confirmée par trois expériences indépendantes. *NT*, non- transfecté ; *WT*, wild-type.



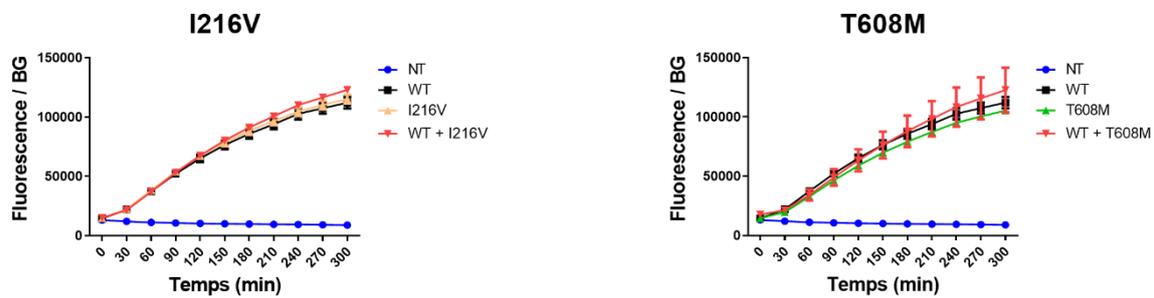
**Figure S3 : Expression protéique des variants perte de fonction complète (Groupe\_A) de *PCSK1* dans le surnageant par analyse LC-MS.** Expression protéique de PC1/3 dans les surnageants de cellules HEK293 transfectées ou non (condition non-transfectée [NT]) avec des plasmides contenant *PCSK1* WT ou un des variants du Groupe\_A. La protéine PC1/3 était considérée comme exprimée si elle était présente dans au moins deux expériences sur trois. Les données ont été transformées en log avant l'analyse. *LFQ*, label-free quantification ; *NT*, non-transfecté ; *WT*, wild-type ; \*, p-valeur significative ; *ns*, p-valeur non significative.



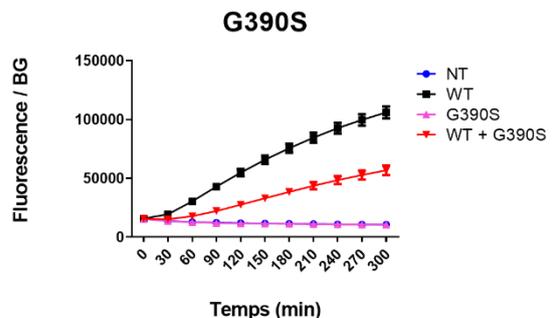
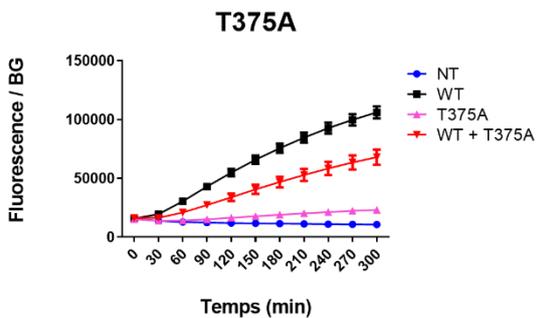
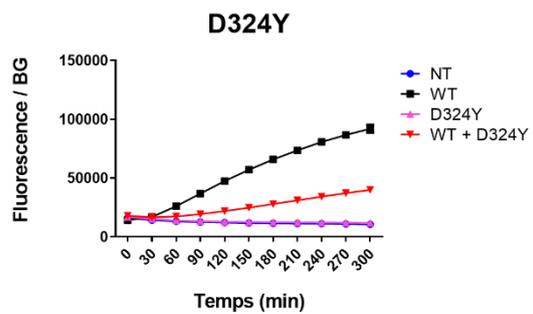
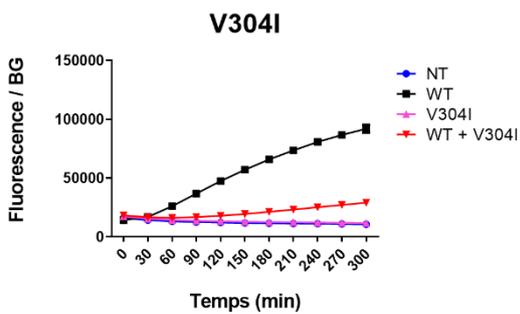
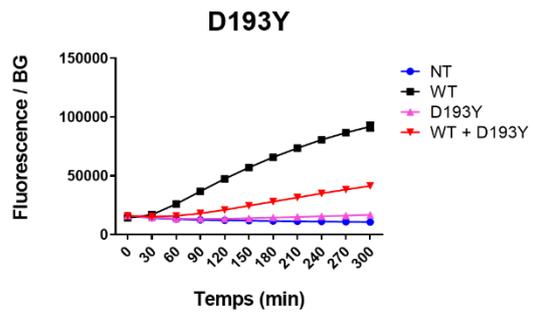
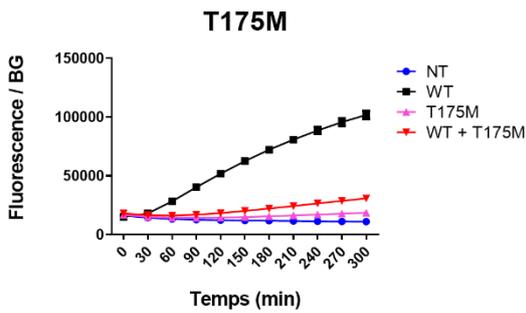
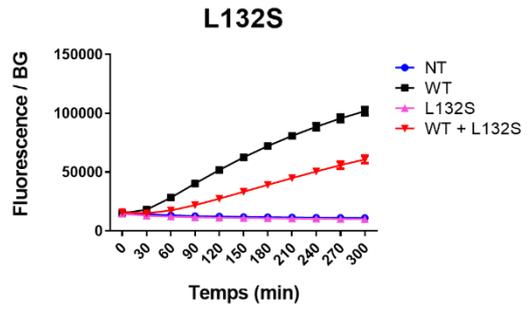
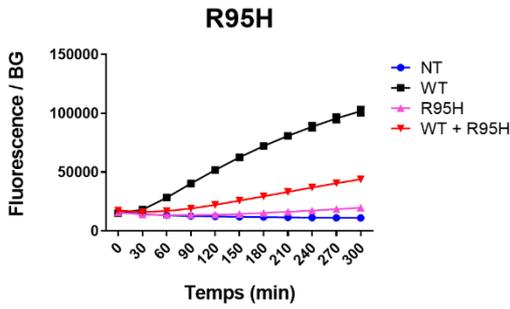
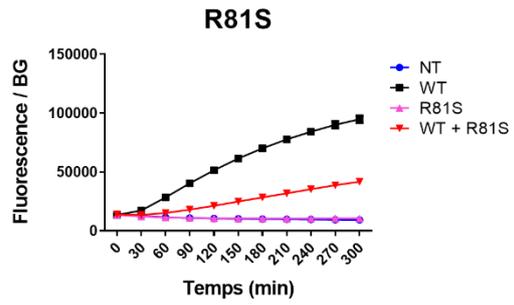
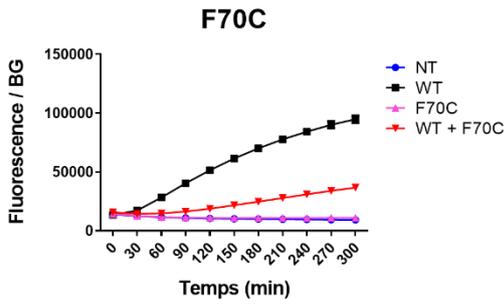
**Figure S4 : Expression génique des variants p.R80\*, p.V304I et p.P400S de *PCSK1* par RT-qPCR.** Des cellules HEK293 ont été transfectées ou non (condition non-transfectée [NT]) avec un plasmide codant *PCSK1* WT ou codant les mutation STOP, p.V304I, ou p.P400S. La condition STOP correspond à l'expression protéique du variant non-sens codant p.R80\*. La figure montre l'expression génique de *PCSK1* et *RPLP0* (utilisé comme gène de ménage) dans les cellules HEK293 transfectées par RT-qPCR. Les données sont la moyenne  $\pm$  SD de deux expériences indépendantes qui ont été réalisées en triplicat technique. *Ct*, cycle threshold; *NT*, non-transfecté ; *WT*, *wild type*.

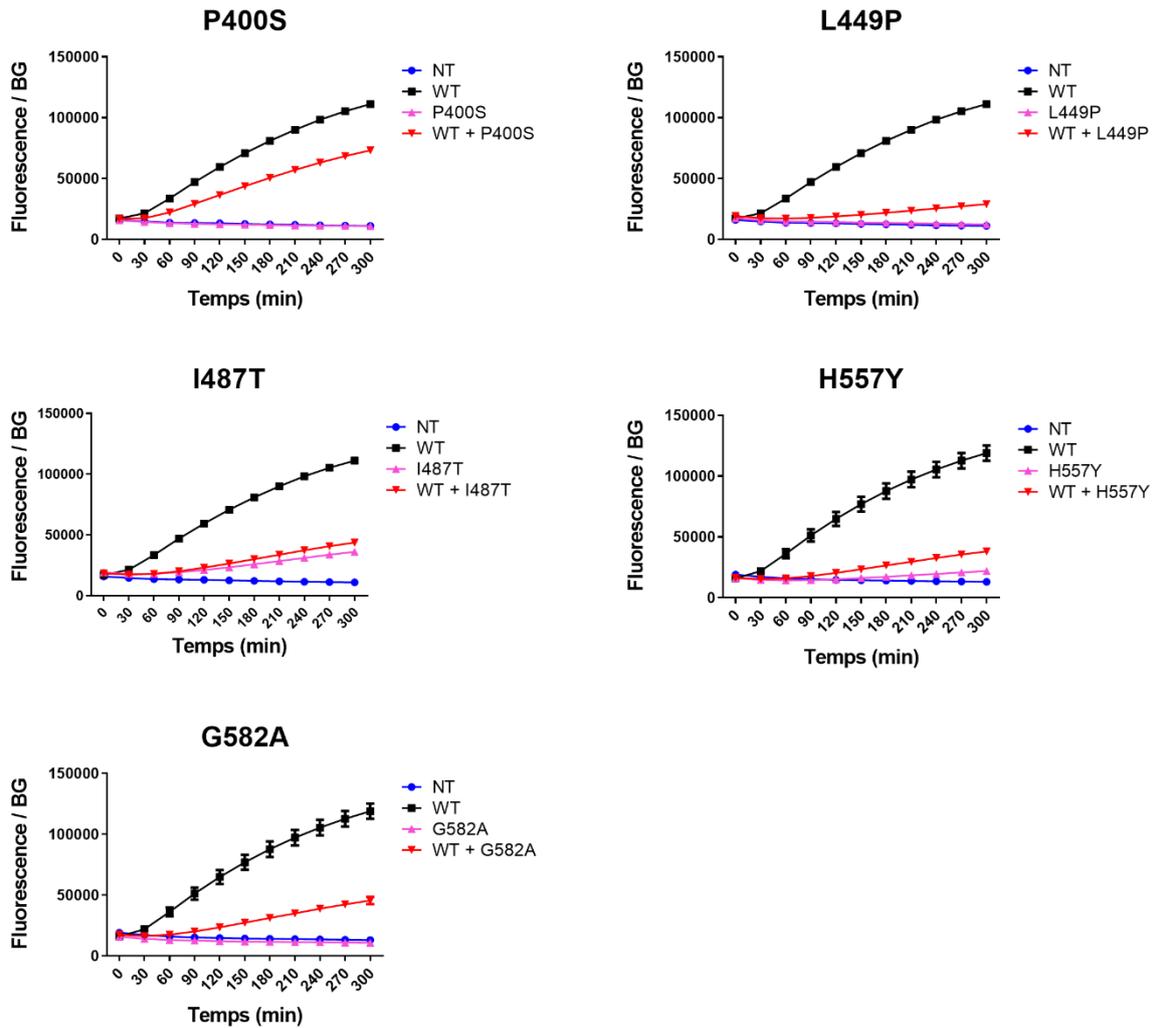


**Figure S5 : Visualisation de la protéine PC1/3 codée par les variants p.R80\*, p.V304I et p.P400S dans le lysat cellulaire par immunofluorescence.** Microscopie confocale de cellules HEK293 transfectées ou non (condition non-transfectée [NT] avec les plasmides *PCSK1* WT, p.R80\*, p.V304I, ou p.P400S. L'expression a été confirmée par trois expériences indépendantes. La barre blanche représente 10  $\mu$ m. PC1/3 est en vert, la matrice extracellulaire (F-actine) en rouge et les noyaux (DAPI) en bleu. *NT*, non-transfecté ; *WT*, wild-type.

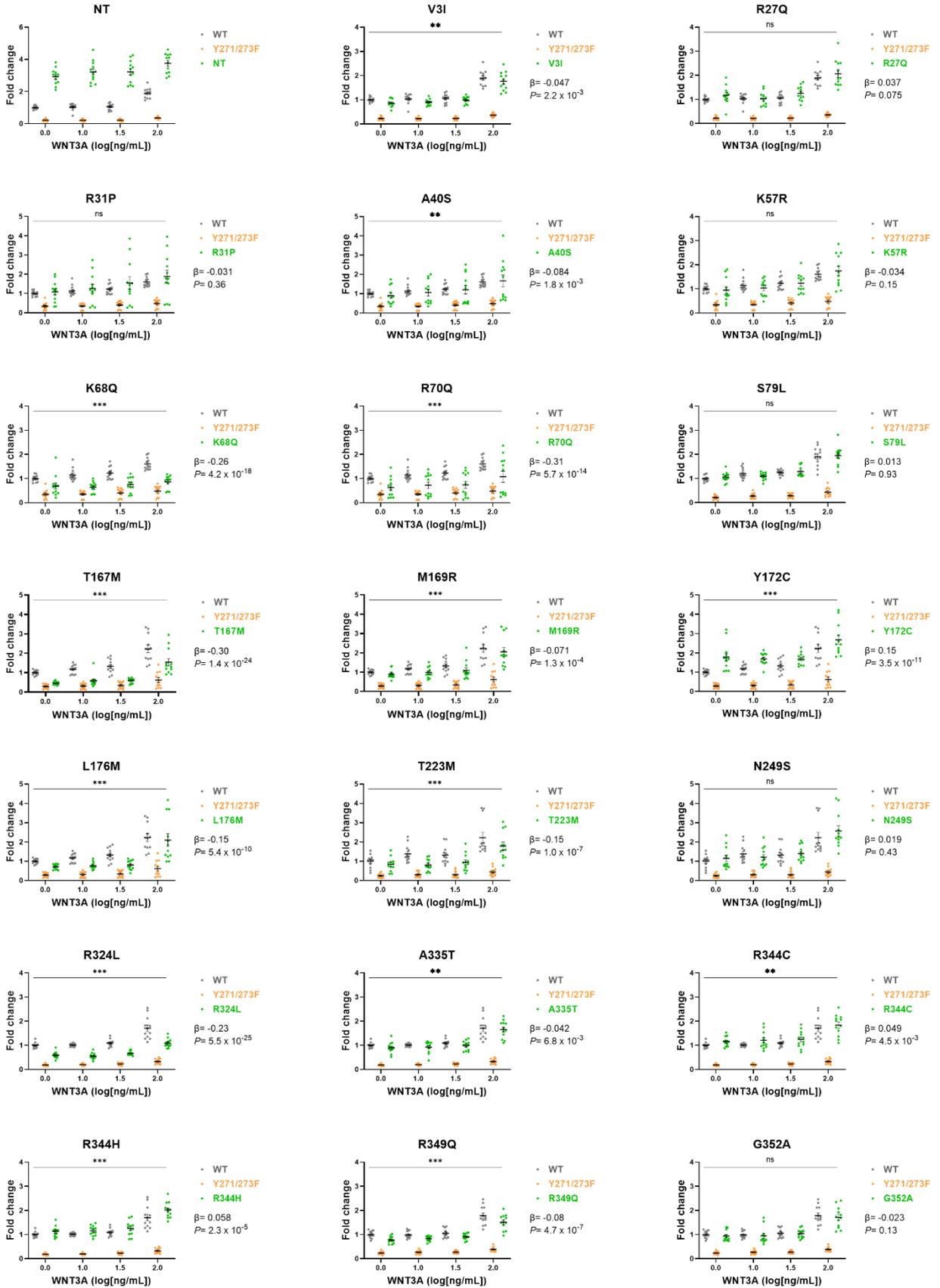


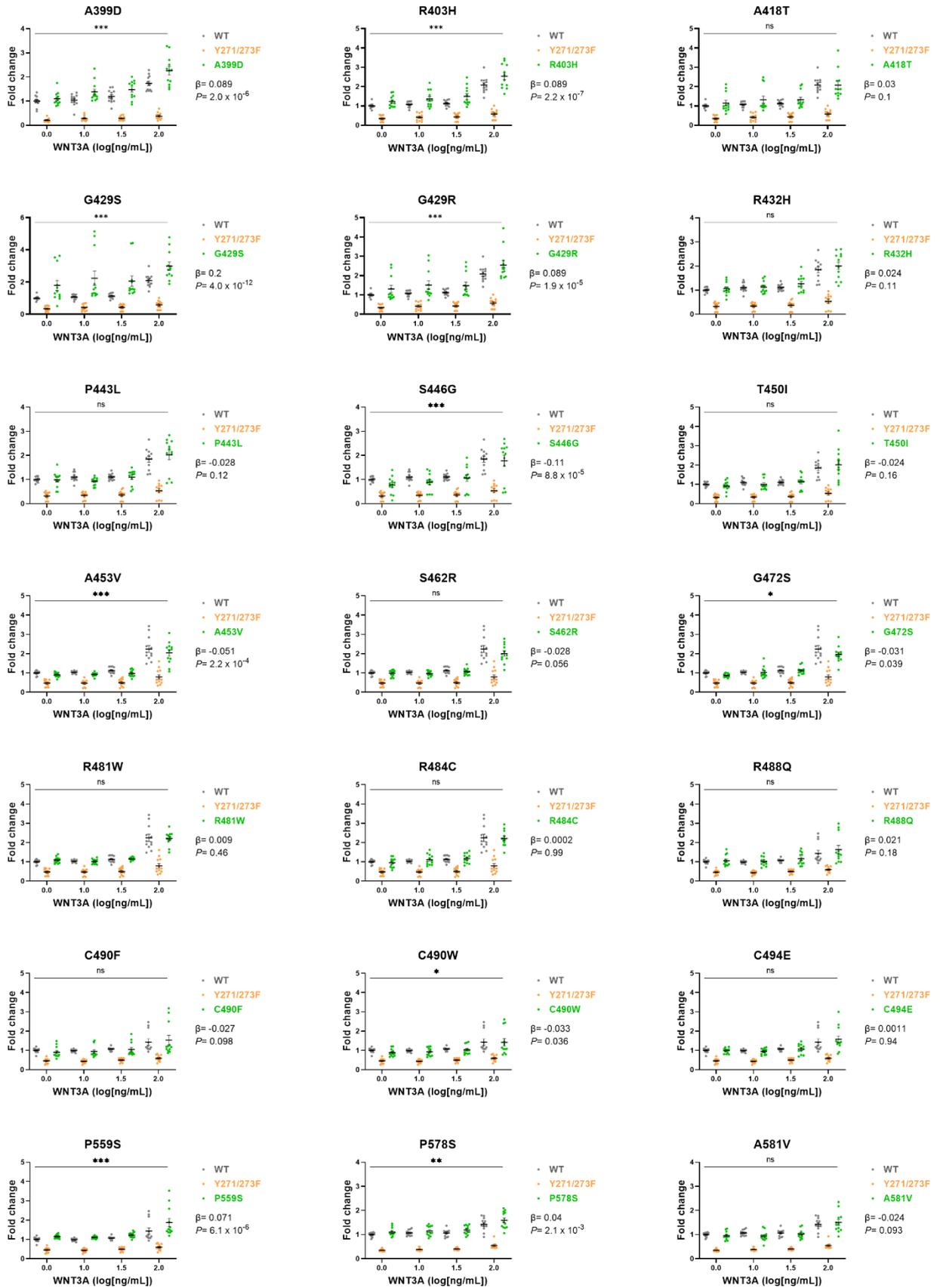
**Figure S6 : Analyse de l'effet dominant négatif des variants rares neutres p.I216V et p.T608M sur l'allèle WT par mesure de l'activité enzymatique de PC1/3.** Les figures montrent l'activité enzymatique de PC1/3 normalisée par l'activité de la  $\beta$ -galactosidase dans les cellules HEK293 cotransfectées ou non (condition non-transfectée [NT]) avec le plasmide *PCSK1* WT et avec un plasmide portant le variant rare p.I216V ou p.T608M de *PCSK1* suivant le ratio 1 : 1 (WT : mutation). Les conditions NT et WT sont représentées en bleu et en noir, respectivement. L'activité enzymatique issue de la co-transfection du WT et du variant est représentée en rouge. Les données sont la moyenne  $\pm$  SEM d'une seule expérience réalisée en triplicat technique. *NT*, non-transfecté ; *WT*, wild-type.

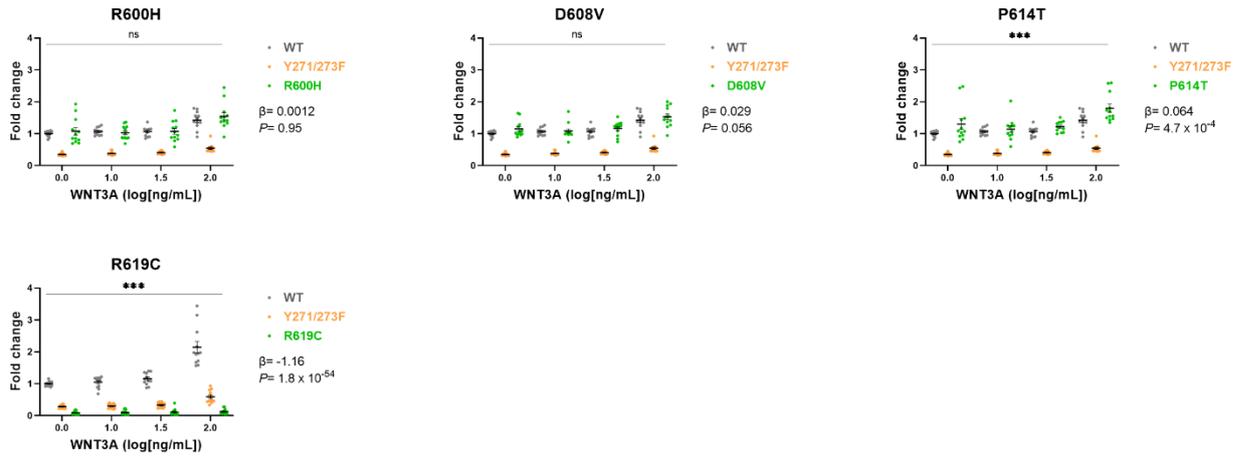




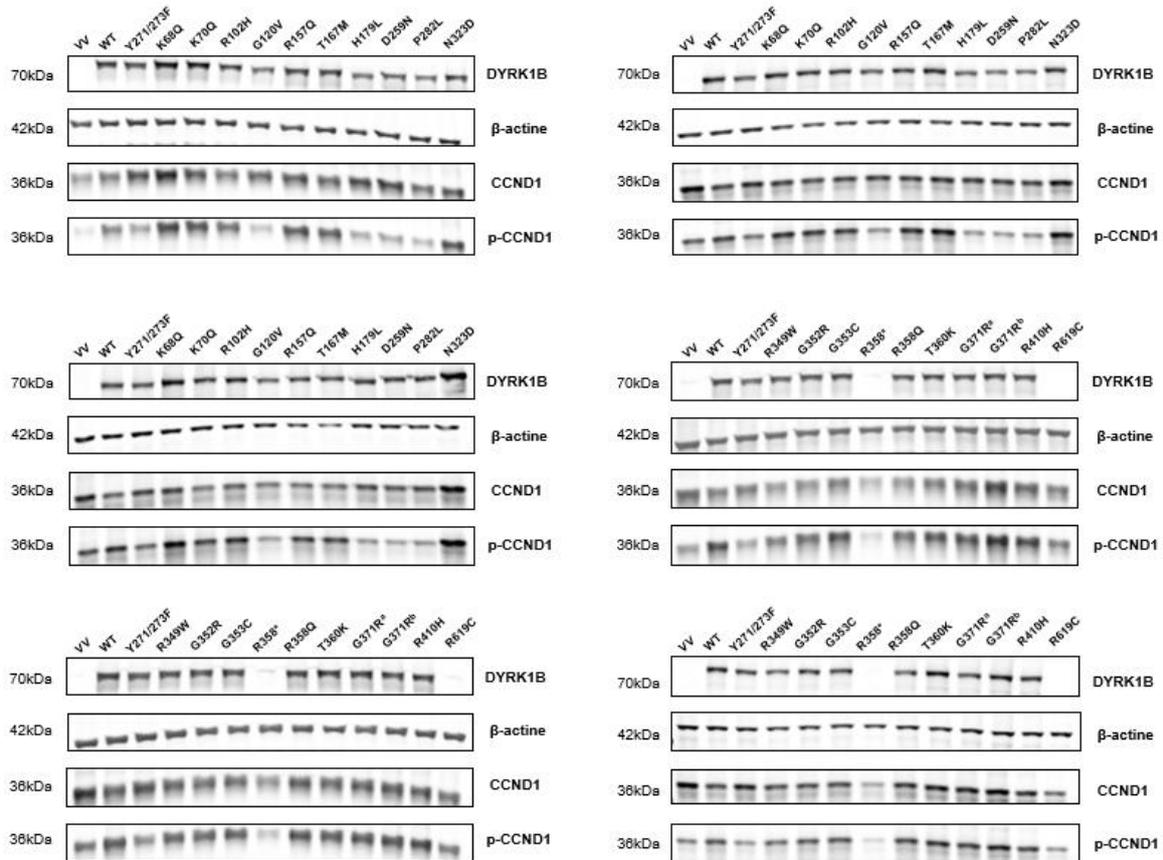
**Figure S7 : Analyse de l'effet dominant négatif des variants non synonymes perte de fonction complète (Groupe\_A) sur l'allèle WT par mesure de l'activité enzymatique de PC1/3.** Les figures montrent l'activité enzymatique de PC1/3 normalisée par l'activité de la  $\beta$ -galactosidase dans les cellules HEK293 cotransfectées ou non (condition non-transfectée [NT]) avec le plasmide *PCSK1* WT et avec un plasmide portant une mutation perte de fonction totale (Groupe\_A) de *PCSK1* suivant le ratio 1 : 1 (WT : mutation A). Les conditions NT et WT sont représentées en bleu et en noir, respectivement. La condition où est exprimée seulement un variant du Groupe\_A est représentée en rose et l'activité enzymatique issue de la cotransfection du WT et d'un des variants du Groupe\_A est représentée en rouge. Les données sont la moyenne  $\pm$  SEM de quatre expériences indépendantes réalisées en triplicat technique. *NT*, non-transfecté ; *WT*, wild-type.



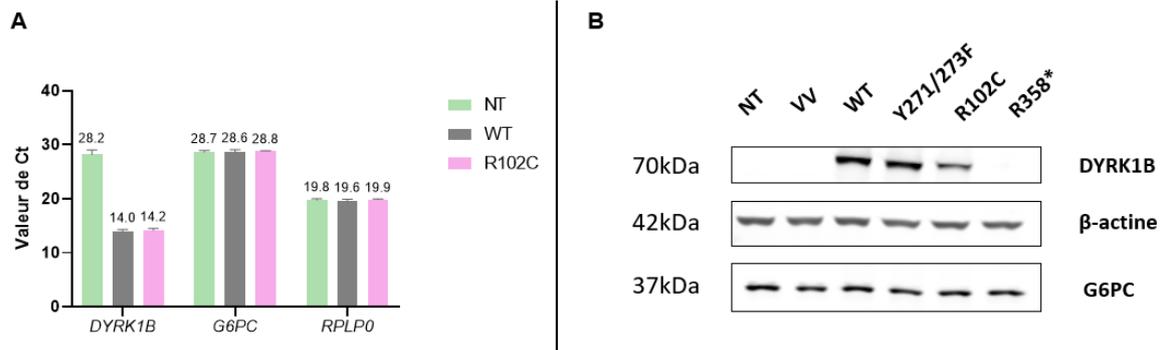




**Figure S8 : Effet des variants neutres de *DYRK1B* sur la signalisation Wnt, selon les analyses luciférase.** La figure montre les *fold changes* de l'activité luciférase normalisés par les mesures de  $\beta$ -galactosidase, réalisés dans des cellules HEK293 qui ont été transfectées ou non-transfectées (condition désignée comme non-transfectée [NT]). Cette transfection implique l'utilisation d'un plasmide exprimant *DYRK1B* WT ou un des variant de *DYRK1B*, ainsi que du plasmide TOPflash (rapporteur TCF). La réponse a été mesurée pour des concentrations croissantes de WNT3A (0, 10, 30, et 100 ng/mL), par rapport à l'activité basale observée avec *DYRK1B* WT. Les contrôles positif et négatif, c'est à dire WT et Y271/273F sont représentés respectivement en gris et orange. Les données sont la moyenne  $\pm$  SEM des *fold changes* de quatre expériences différentes réalisées en triplicat technique. L'effet de chaque variant de *DYRK1B* a été analysé en utilisant un modèle de régression linéaire (avec l'estimates  $[\beta]$  et la p-valeur sur la droite) et confirmé avec le modèle ANOVA (\* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$  versus wild-type). *NT*, non-transfecté ; *WT*, wild type.



**Figure S9 : Effets des variants P/LP-full de *DYRK1B* sur la phosphorylation de CCND1 par western blotting.** La figure montre l'expression protéique de DYRK1B, CCND1 et p-CCND1 dans des cellules HEK293 transfectées. Cette transfection implique l'utilisation d'un vecteur vide [VV] ou d'un plasmide *DYRK1B* WT ou d'un plasmide portant un variant P/LP-full de *DYRK1B* avec un plasmide codant CCND1. Quatre expériences indépendantes ont été réalisées pour chaque variant P/LP-full. *VV*, vecteur vide; *WT*, wild-type; *CCND1*, cycline D1; *p-CCND1*, CCND1 phosphorylée.



**Figure S10 : DYRK1B n'induit pas de changement d'expression de la G6PC.** (A) Des cellules HepG2 ont été transfectées ou non (condition non-transfectée [NT]) avec un plasmide exprimant *DYRK1B* WT ou la mutation p.R102C. La figure montre l'expression génique de *DYRK1B*, *G6PC* et *RPLP0* (utilisé comme gène de ménage) par RT-qPCR dans des cellules HepG2 transfectées. *RPLP0* a été utilisé comme gène de ménage. Les données sont la moyenne  $\pm$  SD de trois expériences indépendantes qui ont été réalisées en triplicat technique. (B) Des cellules HepG2 ont été transfectées ou non (condition non-transfectée [NT]) avec un plasmide vide [VV], un plasmide exprimant *DYRK1B* WT ou exprimant une mutation de *DYRK1B* (p.Y271/273F, p.R102C et p.R358\*). La figure montre l'expression protéique de DYRK1B et G6PC par western blotting dans des cellules HepG2 transfectées. Trois expériences indépendantes ont été réalisées pour chaque condition. *Ct*, cycle threshold; *NT*, non-transfecté; *VV*, vecteur vide ; *WT*, wild-type ;

**ANNEXE 3. Article “Contribution of heterozygous  
*PCSK1* variants to obesity and implications for  
precision medicine: a case-control study”**

Folon L, Baron M, Toussaint B, Vaillant E, Boissel M, Scherrer V, Loïselle H, Leloire A, Badreddine A, Balkau B, Charpentier G, Franc S, Marre M, Aboulouard S, Salzet M, Canouil M, Derhourhi M, Froguel P, Bonnefond A. Contribution of heterozygous *PCSK1* variants to obesity and implications for precision medicine: a case-control study. *The Lancet. Diabetes & Endocrinology*, 2023, 11 (3), pp.182-190.

DOI: 10.1016/S2213-8587(22)00392-8



**ANNEXE 4. Article “Pathogenic, total loss-of-function *DYRK1B* variants cause monogenic obesity associated with type 2 diabetes”**

Folon L, Baron M, Scherrer V, Toussaint B, Vaillant E, Loïselle H, Dechaume A, De Pooter F, Boutry R, Boïssel M, Diallo A, Ning L, Balkau B, Charpentier G, Franc S, Marre M, Derhourhi M, Froguel P, Bonnefond A. Pathogenic, Total Loss-of-Function *DYRK1B* Variants Cause Monogenic Obesity Associated With Type 2 Diabetes. *Diabetes Care*. 2024 Jan 3 : dc231851.

DOI : 10.2337/dc23-1851

