

**Université du Droit et de la Santé de Lille  
Ecole Doctorale Biologie - Santé de Lille**

Année universitaire 2020 – 2021

**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

# **Impact des variants génétiques rares sur l'étiologie du diabète de type 2**

**Par Sarah MEULEBROUCK**

**Sous la direction du Docteur Amélie BONNEFOND**

**Thèse soutenue le 15 décembre 2021 devant le jury composé de :**

<b>Docteur Magalie RAVIER</b>	Rapportrice
<b>Docteur Bertrand BLONDEAU</b>	Rapporteur
<b>Professeur Miriam CNOP</b>	Examinatrice, présidente du jury
<b>Professeur Raphaël SCHARFMANN</b>	Examineur
<b>Docteur Amélie BONNEFOND</b>	Directrice de thèse



<b>REMERCIEMENTS</b> .....	<b>7</b>
<b>RÉSUMÉ</b> .....	<b>8</b>
<b>ABSTRACT</b> .....	<b>10</b>
<b>LISTE DES ABRÉVIATIONS</b> .....	<b>12</b>
<b>INTRODUCTION</b> .....	<b>18</b>
<b>1 HISTOIRE DU DIABÈTE</b> .....	<b>18</b>
<b>2 DEFINITION ET DIAGNOSTIC DU DIABETE</b> .....	<b>20</b>
2.1 DÉFINITION DU DIABÈTE .....	20
2.2 DIAGNOSTIC DU DIABETE .....	21
<b>3 COMPLICATIONS DU DIABETE</b> .....	<b>22</b>
3.1 COMPLICATIONS A COURT TERME .....	22
3.2 COMPLICATIONS A LONG TERME .....	23
<b>4 EPIDÉMIOLOGIE DU DIABÈTE</b> .....	<b>25</b>
4.1 SITUATION DANS LE MONDE .....	25
4.2 SITUATION EN FRANCE .....	27
<b>5 ORGANES TOUCHÉS PAR LE DIABÈTE</b> .....	<b>28</b>
5.1 LE PANCRÉAS .....	29
5.1.1 <i>La cellule <math>\beta</math> pancréatique</i> .....	31
5.1.1.1 Synthèse de l'insuline .....	32
5.1.1.2 Sécrétion de l'insuline .....	35
5.1.1.3 Régulation de l'insulinosécrétion par le glucose .....	38
5.1.1.4 Régulation autocrine et paracrine du pancréas endocrine .....	42
5.1.1.5 Régulation de l'insulinosécrétion par les hormones incrétines .....	45
5.2 ACTION DE L'INSULINE SUR LES TISSUS CIBLES PÉRIPHÉRIQUES .....	47
5.2.1 <i>Activation du récepteur à l'insuline</i> .....	48
5.2.1.1 Activation de la voie PI3K/AKT .....	50
5.2.1.2 Activation de la voie RAS/MAPK .....	51
5.2.2 <i>Effets de l'insuline dans les tissus</i> .....	51
5.2.2.1 Effets de l'insuline dans le muscle squelettique .....	51
5.2.2.2 Effets de l'insuline dans le tissu adipeux .....	53
5.2.2.3 Effets de l'insuline dans le foie .....	54
<b>6 PRISE EN CHARGE DU DIABETE</b> .....	<b>56</b>
6.1 MODIFICATIONS DU STYLE DE VIE .....	57
6.1.1 <i>L'alimentation</i> .....	57
6.1.2 <i>L'activité physique</i> .....	57
6.2 TRAITEMENTS PHARMACOLOGIQUES .....	58
6.2.1 <i>Traitements oraux</i> .....	58
6.2.1.1 Metformine .....	58
6.2.1.2 Sulfonylurées et glinides .....	58
6.2.1.3 Inhibiteurs d' $\alpha$ -glucosidase .....	59
6.2.1.4 Thiazolidinediones .....	60
6.2.1.5 Inhibiteurs de dipeptidyl peptidase-4 .....	60
6.2.1.6 Inhibiteurs du cotransporteur sodium-glucose SGLT2 .....	61
6.2.2 <i>Traitements injectables</i> .....	61
6.2.2.1 Agonistes du récepteur au GLP-1 .....	61
6.2.2.2 Insuline .....	62

<b>7 GENETIQUE DU DIABETE DE TYPE 2.....</b>	<b>63</b>
7.1 LES OUTILS GÉNÉTIQUES.....	64
7.1.1 <i>Analyses de liaison familiale</i> .....	64
7.1.1.1 Génomique de restriction fragment length polymorphisms .....	65
7.1.1.2 Génomique des séquences microsatellites.....	66
7.1.1.3 Séquençage Sanger .....	67
7.1.2 <i>Approche gène-candidat</i> .....	68
7.1.3 <i>Etudes d'association pangénomiques ou GWAS</i> .....	69
7.1.4 <i>Séquençage de nouvelle génération ou NGS</i> .....	71
7.2 LES FORMES MONOGÉNIQUES DU DIABÈTE DE TYPE 2.....	74
7.2.1 <i>Le MODY</i> .....	74
7.2.1.1 <i>HNF4A/MODY1</i> .....	75
7.2.1.2 <i>GCK/MODY2</i> .....	76
7.2.1.3 <i>HNF1A/MODY3</i> .....	77
7.2.1.4 <i>PDX1/MODY4</i> .....	78
7.2.1.5 <i>HNF1B/MODY5</i> .....	79
7.2.1.6 <i>NEUROD1/MODY6</i> .....	79
7.2.2 <i>Le diabète néonatal</i> .....	83
7.2.2.1 Le diabète néonatal transitoire .....	83
7.2.2.2 Le diabète néonatal permanent .....	84
7.2.2.3 Gènes et régions chromosomiques associés au diabète néonatal .....	84
7.2.3 <i>Importance du diagnostic génétique du diabète monogénique</i> .....	88
7.3 LES FORMES POLYGÉNIQUES DU DIABÈTE DE TYPE 2 .....	92
7.3.1 <i>Continuum entre les formes monogéniques et polygéniques</i> .....	92
7.3.2 <i>Identification de variants fréquents associés au DT2 commun</i> .....	93
7.3.2.1 Identification de gènes associés au DT2 commun par études de liaison familiale .....	93
7.3.2.2 Identification de gènes associés au DT2 commun par approche gène-candidat.....	95
7.3.2.3 Identification de SNPs associés au DT2 commun par études GWAS .....	97
7.3.3 <i>Contribution de variants rares à la forme commune du DT2</i> .....	98
7.3.4 <i>Combinaison du séquençage NGS et de la génétique fonctionnelle</i> .....	99
7.3.4.1 Le gène <i>MTNR1B</i> .....	100
7.3.4.2 Le gène <i>PPARG</i> .....	101
7.3.4.3 Le gène <i>HNF1A</i> .....	102
<b>MON TRAVAIL DE THESE .....</b>	<b>105</b>
<b>1 LE PROJET OPRD1.....</b>	<b>106</b>
<b>2 LE PROJET GLIS3.....</b>	<b>109</b>
<b>MATERIEL ET METHODES .....</b>	<b>115</b>
<b>1 ETUDES GÉNÉTIQUES .....</b>	<b>115</b>
1.1 PARTICIPANTS AUX ÉTUDES.....	115
1.1.1 <i>Participants à l'étude RaDiO</i> .....	115
1.1.2 <i>Participants à la UK Biobank</i> .....	116
1.2 IDENTIFICATION DES VARIANTES DU GÈNE <i>OPRD1</i> .....	116
1.2.1 <i>Séquençage des gènes OPRD1 et GLIS3 et identification des variants dans l'étude RaDiO</i> .....	116
1.2.2 <i>Identification des variants du gène OPRD1 dans la UK Biobank</i> .....	119
1.3 ETUDES D'ASSOCIATION ENTRE LES VARIANTES DES GÈNES <i>OPRD1</i> ET <i>GLIS3</i> , ET DIVERS TRAITS MÉTABOLIQUES.....	120
1.3.1 <i>Etudes d'association réalisées pour l'étude RaDiO et la UK Biobank</i> .....	120
1.3.2 <i>Etudes d'association réalisées pour l'AMP T2D Knowledge Portal</i> .....	121
<b>2 ETUDES FONCTIONNELLES .....</b>	<b>121</b>
2.1 GÉNÉRATION DES PLASMIDES.....	121
2.1.1 <i>PCR de mutagenèse</i> .....	124

2.1.2	Dialyse .....	125
2.1.3	Digestion .....	125
2.1.4	Transformation.....	125
2.1.5	Minipreps et extraction des plasmides.....	125
2.1.6	Réaction de séquence et séquençage.....	127
2.1.7	Maxipreps et extraction des plasmides .....	128
2.2	MODELES CELLULAIRES UTILISES .....	129
2.2.1	Lignée cellulaire HEK293 .....	129
2.2.2	Lignée cellulaire EndoC8H1 .....	129
2.3	EXPÉRIENCES DE LUCIFÉRASE .....	130
2.3.1	Expériences de SRE-luciférase .....	131
2.3.2	Expériences de INS-luciférase.....	133
2.4	EXTRACTION D'ARN, RETROTRANSCRIPTION ET QPCR.....	134
2.4.1	Transfection des cellules HEK293 .....	135
2.4.2	Extraction d'ARN .....	135
2.4.3	Rétrotranscription .....	135
2.4.4	qPCR .....	136
2.5	WESTERN BLOTS.....	138
2.5.1	Mesure de l'expression du récepteur DOP.....	138
2.5.2	Mesure de l'expression du facteur de transcription GLIS3 .....	141
2.6	EXPERIENCES D'IMMUNOFLUORESCENCE .....	142
2.7	TESTS DE STIMULATION-SECRETION .....	143
2.7.1	Electroporation des cellules EndoC8H1 .....	143
2.7.2	Test de stimulation-sécrétion .....	144
2.7.3	Quantification des cellules électroporées.....	145
2.7.4	Dosage de l'insuline par ELISA.....	145
<b>RÉSULTATS .....</b>		<b>149</b>
<b>1 PROJET OPRD1.....</b>		<b>149</b>
1.1	IDENTIFICATION DE VARIANTS DU GÈNE <i>OPRD1</i> DANS L'ÉTUDE RADIO .....	149
1.2	PRÉDICTION <i>IN SILICO</i> DE L'EFFET FONCTIONNEL DES VARIANTS DU GÈNE <i>OPRD1</i> .....	149
1.3	ETUDES D'ASSOCIATION ENTRE LES VARIANTS RARES DU GÈNE <i>OPRD1</i> ET LES TRAITS MÉTABOLIQUES ASSOCIÉS AU DT2 ...	149
1.4	IDENTIFICATION DES VARIANTS GOF ET LOF DU GÈNE <i>OPRD1</i> .....	151
1.5	IMPACT DES VARIANTS DU GÈNE <i>OPRD1</i> SUR L'EXPRESSION ET LA LOCALISATION CELLULAIRE DU RÉCEPTEUR DOP ....	153
1.6	POSITION DES 34 VARIANTS DU GÈNE <i>OPRD1</i> DANS LA STRUCTURE DU RÉCEPTEUR DOP .....	159
1.7	ETUDES D'ASSOCIATION ENTRE L'AGRÉGATION DE VARIANTS GOF OU LOF DU GÈNE <i>OPRD1</i> ET LES TRAITS MÉTABOLIQUES ASSOCIÉS AU DT2.....	159
1.7.1	Effet des variants rares du gène <i>OPRD1</i> identifiés dans l'étude <i>RaDiO</i> .....	159
1.7.1.1	Effets des variants rares GoF .....	160
1.7.1.2	Effet des variants rares LoF .....	160
1.7.2	Effets du variant fréquent <i>p.I52V</i> dans l'AMP T2D Knowledge Portal .....	160
1.7.3	Effets des variants rares LoF identifiés dans la UK Biobank .....	161
1.8	EFFETS DE L'ACTIVATION DU RÉCEPTEUR DOP SUR LA SÉCRÉTION D'INSULINE .....	164
<b>2 PROJET GLIS3 .....</b>		<b>167</b>
2.1	IDENTIFICATION DE VARIANTS RARES DU GÈNE <i>GLIS3</i> DANS L'ÉTUDE RADIO .....	167
2.2	PRÉDICTION <i>IN SILICO</i> DE L'EFFET FONCTIONNEL DES VARIANTS DU GÈNE <i>GLIS3</i> .....	167
2.3	ETUDE D'ASSOCIATION ENTRE L'AGRÉGATION DE VARIANTS RARES DU GÈNE <i>GLIS3</i> ET LE RISQUE DE DT2 .....	169
2.4	IDENTIFICATION DES VARIANTS LOF PAR EXPÉRIENCES DE LUCIFÉRASE.....	170
2.4.1	Vérification de l'efficacité de la transfection dans les cellules HEK293.....	170
2.4.1.1	Expression du gène <i>GLIS3</i> .....	170
2.4.1.2	Expression de la protéine <i>GLIS3</i> .....	171

2.4.2	Identification des variants LoF du gène <i>GLIS3</i> par expériences de luciférase.....	171
2.5	ÉTUDE D'ASSOCIATION ENTRE L'AGRÉGATION DES VARIANTS RARES LOF DU GÈNE <i>GLIS3</i> ET LE RISQUE DE DT2.....	173
2.6	IDENTIFICATION DES VARIANTS P/LP DU GÈNE <i>GLIS3</i> SELON LES CRITÈRES ACMG.....	173
2.7	ENRICHISSEMENT EN VARIANTS P/LP DANS LE DOMAINE C-TERMINAL DE LA PROTÉINE <i>GLIS3</i> .....	176
2.8	ÉTUDE D'ASSOCIATION ENTRE L'AGRÉGATION DE VARIANTS RARES P/LP DU GÈNE <i>GLIS3</i> ET LE RISQUE DE DT2.....	177
<b>DISCUSSION ET PERSPECTIVES.....</b>		<b>179</b>
<b>1</b>	<b>PROJET <i>OPRD1</i>.....</b>	<b>179</b>
<b>2</b>	<b>PROJET <i>GLIS3</i>.....</b>	<b>184</b>
<b>CONCLUSION.....</b>		<b>189</b>
<b>BIBLIOGRAPHIE.....</b>		<b>191</b>
<b>ANNEXES.....</b>		<b>214</b>
<b>ANNEXE 1. CARTES REPRÉSENTANT LA CONSTRUCTION DES DIFFÉRENTS PLASMIDES UTILISÉS DANS LE CADRE DES PROJETS <i>OPRD1</i> ET <i>GLIS3</i>.....</b>		<b>214</b>
<b>ANNEXE 2. LISTE DES AMORCES MODIFIÉES UTILISÉES POUR LA RÉALISATION DES PCR DE MUTAGENÈSE DU PROJET <i>OPRD1</i>.....</b>		<b>217</b>
<b>ANNEXE 3. LISTE DES AMORCES MODIFIÉES UTILISÉES POUR LA RÉALISATION DES PCR DE MUTAGENÈSE DU PROJET <i>GLIS3</i>.....</b>		<b>219</b>
<b>ANNEXE 4. LISTE DES AMORCES UTILISÉES POUR LE SÉQUENÇAGE DES PLASMIDES <i>POPRD1</i>.....</b>		<b>223</b>
<b>ANNEXE 5. LISTE DES AMORCES UTILISÉES POUR LE SÉQUENÇAGE DES PLASMIDES <i>PGLIS3</i>.....</b>		<b>223</b>
<b>ANNEXE 6. ACTIVITÉ LUCIFÉRASE DES VARIANTS NEUTRES DU GÈNE <i>OPRD1</i>.....</b>		<b>224</b>
<b>ANNEXE 7. EXPRESSION DES DIFFÉRENTS MUTANTS DU RÉCEPTEUR DOP.....</b>		<b>225</b>
<b>ANNEXE 8. LOCALISATION CELLULAIRE DES DIFFÉRENTS MUTANTS DU RECEPTEUR DOP.....</b>		<b>226</b>
<b>ANNEXE 9. DONNÉES CLINIQUES DES INDIVIDUS PORTEURS OU NON D'UN VARIANT RARE GOF OU LOF DANS L'ÉTUDE RADIO.....</b>		<b>231</b>
<b>ANNEXE 10. SECRETION D'INSULINE PAR LES CELLULES ENDOCBH1 SUREXPRIMANT LE GENE <i>OPRD1</i> ET STIMULEES A BAS ET HAUT GLUCOSE, EN COMBINAISON AVEC L'AGONISTE DE DOP DII.....</b>		<b>232</b>
<b>ANNEXE 11. ACTIVITE LUCIFERASE DES VARIANTS NEUTRES DU GENE <i>GLIS3</i>.....</b>		<b>233</b>
<b>ANNEXE 12. EVALUATION DES CRITÈRES ACMG POUR CHAQUE VARIANT RARE DU GÈNE <i>GLIS3</i> IDENTIFIÉ DANS L'ÉTUDE RADIO.....</b>		<b>235</b>
<b>ANNEXE 13. DONNÉES CLINIQUES DES INDIVIDUS PORTEURS OU NON D'UN VARIANT RARE P/LP DANS L'ÉTUDE RADIO.....</b>		<b>237</b>
<b>ANNEXE 14. ARTICLE « MIRROR EFFECT OF DELTA OPIOID RECEPTOR MUTATIONS ON TYPE 2 DIABETES AND OBESITY” SOUMIS POUR PUBLICATION.....</b>		<b>238</b>

## Remerciements

Je voudrais tout d'abord remercier le Dr Magalie RAVIER et le Dr Bertrand BLONDEAU de m'avoir suivie durant ces 3 années au cours de mes comités de suivi individuel, et d'avoir accepté de juger mon travail de thèse. Je remercie également le Pr Miriam CNOP ainsi que le Pr Raphaël SCHARFMANN d'avoir accepté d'être les examinateurs de ce travail.

Je tiens à remercier le Pr Philippe FROGUEL de m'avoir permis d'intégrer l'UMR 1283/8199 afin d'y réaliser ce projet de thèse.

Un grand merci au Dr Amélie BONNEFOND, ma directrice de thèse, de m'avoir permis d'intégrer son équipe et de m'avoir fait découvrir le monde de la génétique. Merci de m'avoir encadrée pendant ces 3 années et d'avoir partagé ton savoir.

Je tiens aussi à remercier l'équipe NGS, l'équipe de bioinformatique et l'équipe de statistique pour toutes les analyses réalisées, et sans qui ce projet n'aurait pas pu voir le jour.

Un grand merci aux différents membres de l'équipe fonctionnelle pour leur aide, et plus particulièrement à Hélène pour toute la mutagenèse ; Raphy, Gurvan, Julie et Audrey pour les différentes expériences fonctionnelles du projet *OPRD1* ; Margot et Victoria pour leur contribution aux nombreuses luciférases du projet *GLIS3* ; Morgane et Marlène pour leurs précieux conseils concernant la culture des cellules  $\beta$  pancréatiques.

Plus particulièrement, j'aimerais remercier Hélène, Margot, Lise, Victoria, et Judith qui ont toujours été présentes, dans les bons comme dans les mauvais moments. Merci pour ces pauses repas, ces (longues) pauses improvisées, pour les fous rires partagés.

Merci à mes parents et ma sœur d'avoir été présents lorsque ça n'allait pas, et d'avoir toujours cru en moi.

Finalement, je voudrais remercier du fond du cœur mon amour, Matthias, d'avoir toujours été présent pour moi au cours de ces dernières années. Merci pour ton courage et ta patience pour avoir su me supporter au quotidien, malgré mon caractère pas toujours facile.... Tu as toujours été là pour moi, pour me reconforter ou me consoler lors de mes périodes de déprime. Je ne saurai jamais assez te remercier.

## Résumé

Le DT2 est une maladie métabolique chronique incluant des formes monogéniques (MODY et diabète néonatal), dans lesquelles l'atteinte d'un seul gène est suffisante pour provoquer la maladie, et polygéniques, provoquées par la combinaison de plusieurs facteurs génétiques et environnementaux, et correspondant à la majorité des cas. Alors que l'investigation de l'étiologie génétique du DT2 monogénique a permis d'identifier plusieurs variants rares impliqués dans son développement, l'investigation du DT2 polygénique, principalement par études d'association pangénomiques (GWAS), a permis d'identifier surtout des variants fréquents situés dans des régions non-codantes du génome et n'ayant qu'un faible impact sur le risque de développer la maladie, expliquant faiblement son héritabilité. Des variants rares présents dans les régions codantes de certains gènes pourraient être associés au DT2 commun avec un effet plus important.

Par séquençage de nouvelle génération (NGS) mené dans le cadre du projet RaDiO, 31 variants rares et 3 variants fréquents hétérozygotes du gène *OPRD1*, codant le récepteur  $\delta$  aux opioïdes (DOP), ont été identifiés chez 6971 individus. En parallèle, 105 variants rares hétérozygotes du gène *GLIS3*, un facteur de transcription jouant un rôle majeur dans la cellule  $\beta$  pancréatique et impliqué dans développement du diabète néonatal, ont été identifiés chez 5471 individus.

L'investigation des variants du gène *OPRD1* par luciférase, suite à leur stimulation avec les agonistes de DOP deltorphin II et DPDPE, a permis d'identifier 7 variants gain de fonction (GoF), comprenant le variant I52V fréquent chez les Africains, ainsi que 12 variants rares perte de fonction (LoF). En parallèle, des expériences de western blot et d'immunofluorescence ont démontré que les variants LoF P14R et G36E altèrent l'expression et la localisation du récepteur DOP, qui n'est plus situé à la membrane mais est retenu dans la cellule. Des études d'association réalisées entre les clusters de variants rares GoF et LoF et différents traits métaboliques associés au DT2 chez les 6971 individus issus de l'étude RaDiO, puis répliquées chez 34 812 individus de l'Accelerated Medicine Partnership T2D Knowledge Portal et 187 242 individus de la UK Biobank ont mis en évidence des effets opposés de ces variants : alors que les variants GoF sont associés à une augmentation du risque de DT2 mais à une diminution de l'adiposité, les variants LoF sont associés à une diminution du risque de DT2 mais une

augmentation de l'adiposité. Finalement, étant donné qu'*OPRD1* est exprimé dans les cellules  $\beta$  pancréatiques humaines, son effet sur la sécrétion d'insuline a été étudié : la stimulation de cellules  $\beta$  humaines EndoC $\beta$ H1 surexprimant *OPRD1* avec du glucose, couplé à l'agoniste de DOP deltorphin II, inhibe la sécrétion d'insuline par les cellules  $\beta$ .

L'investigation des variants rares du gène *GLIS3* par luciférase a permis d'identifier 49 variants LoF. 18 d'entre eux ont été définis comme pathogènes ou probablement pathogènes selon les critères de l'American College of Medical Genetics and Genomics (ACMG). Les études d'association réalisées au sein de la cohorte issue de l'étude RaDiO ont mis en évidence une association entre ce cluster de variants et une augmentation du risque de DT2.

La combinaison du séquençage NGS et de la génétique fonctionnelle a ainsi permis d'identifier 2 nouveaux gènes porteurs de variants GoF ou LoF associés au DT2 polygénique : le gène *OPRD1* qui présente des effets opposés sur le DT2 et l'adiposité et inhibe la sécrétion d'insuline, et le gène *GLIS3* qui en plus d'être associé au diabète néonatal est également associé au DT2 polygénique. Nous supposons que l'identification de nouveaux gènes associés au DT2 polygénique selon cette stratégie permettrait d'approfondir la compréhension de l'héritabilité de cette maladie, mais également d'évoluer vers une médecine de précision par adaptation du traitement des patients en fonction des gènes atteints.

## **Abstract**

Type 2 diabetes (T2D) is a chronic metabolic disorder that includes monogenic forms (e.g. MODY and neonatal diabetes), in which a single mutation in only one gene is sufficient to cause the disease, and polygenic forms, in which the disease is caused by an interaction between multiple genetic and environmental factors, corresponding to the majority of cases. Whereas the investigation of the genetic etiology of monogenic forms has enabled the identification of many rare genetic variants involved in this disorder, the investigation of the polygenic forms (mainly through genome-wide association studies [GWAS]) led to the identification of frequent variants that are located in non-coding regions of the genome, and presenting with a small effect size on the disease risk, weakly explaining the heritability of the pathology. Rare variants, located in coding regions of some genes, could be associated with common T2D with a stronger effect size.

By Next Generation Sequencing (NGS) performed in the context of the RaDiO project, 31 rare variants and 3 frequent variants that are heterozygous have been identified in the coding regions of the *OPRD1* gene, encoding the  $\delta$  opioid receptor (DOP) in a cohort of 6,971 individuals. Moreover, 105 rare heterozygous variants of the *GLIS3* gene, that is a key regulator in the pancreatic  $\beta$  cell involved in the development of neonatal diabetes, have been identified in a cohort of 5,471 individuals.

By investigating the genetic variants of *OPRD1* through luciferase assays following DOP activation by the 2 DOP agonists deltorphin II and DPDPE, 7 gain-of-function (GoF) (among them the variant I52V that is frequent in the African population) and 12 rare loss-of-function (LoF) variants have been identified. In parallel, western blotting and immunofluorescence experiments showed that the rare LoF variants P14R and G36E decrease the expression and the localization of DOP, that is not located at the membrane like the wild-type DOP but retained inside the cells. Association studies performed between the clusters of rare GoF and LoF variants and various metabolic traits associated with T2D, followed by a replication in 34,812 individuals from the Accelerated Medicine Partnership T2D Knowledge Portal and 187,242 individuals from the UK Biobank highlighted opposite effects of these variants : whereas GoF variants are associated with an increase in T2D risk but with a decrease in adiposity, LoF variants are associated with a decreased T2D risk, but an increased adiposity.

Finally, as *OPRD1* is expressed in human pancreatic  $\beta$  cells, we studied its effect on insulin secretion. The stimulation of EndoC $\beta$ H1 cells (that are a model of human  $\beta$  cell) overexpressing *OPRD1* with glucose, coupled with the DOP agonist deltorphin II, inhibits insulin secretion from  $\beta$  cells.

The investigation of rare heterozygous variants of *GLIS3* through luciferase assays enabled the identification of 49 LoF variants. 18 of them have been defined as pathogenic or likely pathogenic following the application of the American College of Medical Genetics and Genomics (ACMG). Association studies conducted between this cluster of 18 variants and T2D in the individuals from the RaDiO study showed that these variants are associated with an increased risk to develop T2D.

By combining NGS and functional genetic experiments, 2 new genes carrying rare GoF or LoF variants have been found to be associated with common T2D during my thesis: *OPRD1*, presenting with opposite effects on T2D risk and adiposity and inhibiting insulin secretion, and *GLIS3*, that is associated with common T2D in addition to be involved in neonatal diabetes. The identification of new genes associated with common forms of T2D following this strategy would lead to a better understanding of its heritability, but could also contribute to the development of precision medicine by adapting the treatment depending of the genes that are altered.

## Liste des abréviations

ABCC8	ATP binding cassette subfamily C member 8
ACC	Acétyl-CoA carboxylase
ACMG	American College of Medical Genetics and Genomics
ADA	American Diabetes Association
ADN	Acide désoxyribonucléique
ADNc	ADN complémentaire
ADP	Adénosine diphosphate
AMPc	Adénosine monophosphate cyclique
AMPK	AMP-activated protein kinase
APOA1	Apolipoprotéine A1
APOA2	Apolipoprotéine A2
ARN	Acide ribonucléique
ARNm	ARN messenger
ATP	Adénosine triphosphate
BCA	Acide bicinchoninique
BSA	Albumine de sérum bovin
BUB2	Mitotic check point protein BUB2
CAMK2	Calcium/calmoduline kinase 2
CCND2	Cycline D2
CEPH	Centre d'étude du polymorphisme humain
C-FOS	FBJ murine osteosarcoma viral oncogene homolog
CMIP	C-Maf-Inducing Protein
CoA	Coenzyme A
Ct	Cycle threshold
DAG	Diacylglycérol
ddNTP	didésoxynucléotide
DESIR	Données épidémiologiques sur le syndrome d'insulinorésistance
DII	[D-Ala2]-Deltorhin II
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Diméthylsulfoxyde
dNTP	Désoxyribonucléotide
DPDPE	[D-Pen2,D-Pen5]Enkephalin
DPP-4	Dipeptidyl peptidase-4
DT1	Diabète de type 1
DT2	Diabète de type 2
DTT	Dithiothréitol
EGTA	Acide egtazique
ELISA	Enzyme-linked immunosorbent assay
EPAC2	Exchange protein directly activated by cAMP 2
ERK	Extracellular signal-regulated kinase
ELK-1	ETS-like 1
FAS	Fatty acid synthase
FOXO	Forkhead box O
G-6-Pase	Glucose-6-phosphatase

GABA	Acide $\gamma$ -aminobutyrique
GAD65	Isoforme de 65 kDa de la glutamate décarboxylase
GAPDH	Glycéraldéhyde 3-phosphate déshydrogénase
GCGR	Récepteur au glucagon
GCK	Glucokinase
Gi	Protéine G inhibitrice
GIP	Glucose-dependent insulintropic polypeptide
GIPR	Récepteur au GIP
GIRK	G-protein inward rectifying K <sup>+</sup> channel
GLIS1	GLI-Similar family zinc finger 1
GLIS2	GLI-Similar family zinc finger 2
GLIS3	GLI-Similar family zinc finger 3
GLISBS	GLIS binding sites
GLP-1	Glucagon-like peptide 1
GLP1R	Récepteur au GLP-1
GLUT1	Transporteur de glucose 1
GLUT2	Transporteur de glucose 2
GLUT4	Transporteur de glucose 4
gnomAD	Genome Agregation Database Browser
GoF	Gain de fonction
GPCR	Récepteur couplé aux protéines G
GRB10	Growth-factor receptor bound protein 10
GRB14	Growth-factor receptor bound protein 14
GRK	G-protein coupled receptor kinase
Gs	Protéine G stimulatrice
GSK3	Glycogène synthase kinase 3
GWAS	Genome-wide association studies, études d'association pangénomiques
HbA1c	Hémoglobine glyquée
HDL	High density lipoproteins, lipoprotéines à haute densité
HEK293	Human Embryonic Kidney 293
HNF1A/HNF-1 $\alpha$	Hepatocyte nuclear factor 1 $\alpha$
HNF1B/HNF-1 $\beta$	Hepatocyte nuclear factor 1 $\beta$
HNF4A/HNF-4 $\alpha$	Hepatocyte nuclear factor 4 $\alpha$
HRP	Horseradish peroxidase, peroxydase de raifort
HSL	Lipase hormonosensible
hTERT	Transcriptase inverse de la télomérase humaine
HYMAI	Hydatiform mole-associated and imprinted
IA-2	Islet antigen 2
IA-2 $\beta$	Islet antigen 2 $\beta$
IFG	Impaired fasting glucose, anomalie de la glycémie à jeun
IGT	Impaired glucose tolerance, intolérance au glucose
INS	Préproinsuline
INSR	Récepteur à l'insuline
IP3	Inositol-1,4,5-trisphosphate
IRE	Elément de réponse à l'insuline
IRS	Insulin receptor substrate

K <sup>+</sup> <sub>ATP</sub>	Canal potassique ATP-dépendant
KCNJ11	Potassium inwardly rectifying channel subfamily J member 11
Kir6.2	Inward-rectifier potassium channel 6.2
LB	Lysogeny broth
LDL	Low density lipoproteins, lipoprotéines à faible densité
LEPR	Récepteur à la leptine
LOD	Logarithm of odds, logarithme des probabilités
LoF	Perte de fonction
MAF	Minor allele frequency
MAFA	MAF BZIP transcription factor A
MAPK	Mitogen-activated protein kinase
MC3R	Récepteur à la mélanocortine 3
MC4R	Récepteur à la mélanocortine 4
MEK	Mitogen-activated protein kinase kinase
MTNR1B/MT2	Récepteur à la mélatonine 1B/2
MUNC13-1	Mammalian uncoordinated 13-1
MUNC18-1	Mammalian uncoordinated 18-1
NEUROD1	Neuronal differentiation 1
NGN3	Neurogénine 3
NGS	Next-generation sequencing, séquençage de nouvelle génération
NSF	N-ethylmaleimide sensitive fusion
NT	Non-transfecté
OGTT	Oral glucose tolerance test, test d'hyperglycémie par voie orale
OMS	Organisation Mondiale de la Santé
OPRD1/DOP	Récepteur δ aux opioïdes
OPRK1/KOP	Récepteur κ aux opioïdes
OPRM1/MOP	Récepteur μ aux opioïdes
OR	Odds ratio
P90RSK	p90 ribosomal S6 kinase
PBS	Phosphate buffered saline
PC2	Prohormone convertase 2
PC3	Prohormone convertase 3
PCR	Polymerase chain reaction
PDE3B	phosphodiesterase 3B
PDK-1	Phosphoinositide-dependent kinase-1
PDX1	Pancreatic and duodenal homeobox 1
PENK	Proenképhaline
PEPCK	Phosphoenolpyruvate carboxykinase
PGC-1	Peroxisome-proliferator-activated receptor-gamma coactivator-1
PH	Pleckstrin homology
PI3K	phosphoinositide 3-kinase
PIP2	Phosphatidylinositol-4,5-bisphosphate
PIP3	Phosphatidylinositol-3,4,5-trisphosphate
PKA	Protéine kinase A
PKB/AKT	Protéine kinase B
PKC	Protéine kinase C

PLA2	Phospholipase A2
PLC	Phospholipase C
POMC	Proopiomélanocortine
PP	Polypeptide pancréatique
PP1	Protéine phosphatase 1
PPAR $\alpha$	Peroxisome proliferator-activated receptor $\alpha$
PPAR $\gamma$	Peroxisome proliferator-activated receptor $\gamma$
PPYR1	Récepteur 1 au polypeptide pancréatique
PTB	Phosphotyrosine binding
RaDiO	Rare variants involved in diabetes and obesity
RAF	Rapidly accelerated fibrosarcoma
RAS	Rat sarcoma virus
REVEL	Rare exome variant ensemble learner
RFLP	Restriction fragment length polymorphism
RPKM	Reads per kilobase million
RT	Rétrotranscription
RyR	Récepteur à la ryanodine
SCID	Déficit immunitaire combiné sévère
SDS	Sodium dodécyl sulfate
SGLT2	Cotransporteur sodium-glucose de type 2
SH2	Src homology 2
SH2B1	Src homology 2 b adaptator protein 1
SH2B2/APS	Src homology 2 b adaptator protein 2
SHC	SRC homology 2 domain containing adaptor protein
SNAP-25	Synaptosomal nerve-associated protein 25
SNAP- $\alpha$	Synaptosomal nerve-associated protein $\alpha$
SNARE	Soluble N-ethylmaleimide sensitive factor attachment protein receptor
SNC	Système nerveux central
SNP	Single-nucleotide polymorphism, polymorphisme d'un seul nucléotide
SOC	Super optimal broth
SOS	Son of sevenless
SRE	Serum Response Element
SREBP-1	Sterol regulatory element binding protein 1
SST	Short sequence repeat
SST-14	Isoforme 14 de la somatostatine
SSTR	Récepteur à la somatostatine
SUR1	Sulfonylurea receptor-1
SV40LT	Simian vacuolating virus 40 large T antigen
TBC1D1	TRE-2/BUB2/CDC16 domain family member 1
TBC1D4/AS160	TRE-2/BUB2/CDC16 domain family member 4
TBS	Tris buffered saline
TCF7L2	Transcription factor 7-like 2
TDAH	Trouble du déficit de l'attention avec hyperactivité
TEMED	Tetramethylethylenediamine
TMB	3, 3', 5, 5'-tétraméthylbenzidine
TPM	Transcripts per million

t-SNARE	Target SNARE
TZD	Thiazolidinediones
USP6/TRE-2	Ubiquitin specific protease 6
VAMP	Vesicle-associated membrane protein
v-SNARE	Vesicular SNARE
WES	Whole-exome sequencing
WGS	Whole-genome sequencing
WT	Wild-type
ZAC	Zinc activated ion channel
ZNT8	Zinc transporter 8

# **INTRODUCTION**

# Introduction

## 1 Histoire du diabète

La première description du diabète date de l'antiquité. Un papyrus égyptien datant de 1500 avant J.C., découvert par l'égyptologue allemand Georg Ebers (1837-1898), décrivait déjà une polyurie chez certains patients atteints d'une « maladie sucrée »<sup>1</sup>. Ce fut Arétée de Cappadoce qui fut le premier, au 2<sup>ème</sup> siècle, à employer le mot « diabète » signifiant « siphon » en latin, étant donné le flux important d'urine excrétée par les patients. Il décrit également chez eux une soif qui ne s'assouvit pas, la « fonte » de la chair et des membres, ainsi qu'une faible survie<sup>1</sup>. En 400 – 500, le physicien indien Sushruta et le chirurgien indien Charaka parlèrent de « madhumeha », signifiant « urine miel » chez les patients, en référence à son goût sucré. Ce furent les premiers à distinguer deux types de diabète, considérés aujourd'hui comme le diabète de type 1 (DT1) et le diabète de type 2 (DT2)<sup>2</sup>. Le goût sucré de l'urine des patients diabétiques fut également décrit dans l'Encyclopédie du philosophe et médecin persan Avicenne (980-1037), qui décrivit chez ces patients un appétit anormal et une gangrène diabétique<sup>1,2</sup>. Le terme « mellitus », venant du latin et signifiant « sucré comme le miel » fut quant à lui utilisé pour la première fois par le chirurgien britannique John Rollo (décédé en 1809) en 1798, afin de différencier le diabète sucré du diabète insipide (« insipidus ») où l'urine n'est pas sucrée<sup>2</sup>.

La première mention du diabète en Europe a été faite par Thomas Willis (1621-1675), dans son ouvrage de 1674, où il décrit l'urine des patients diabétiques comme étant très sucrée comme du sucre ou du miel, sans pour autant considérer la présence de sucre dans celle-ci<sup>1</sup>. L'hyperglycémie présente chez les patients fut décrite pour la première fois en 1776 par Matthew Dobson (1735-1784). Il observa un goût sucré à la fois dans le sérum et l'urine de patients, et remarqua que l'évaporation de l'urine laisse un dépôt blanc sentant le sucre. Il en conclut que le sucre sécrété par les reins n'est pas formé *de novo* dans les organes sécréteurs, mais est déjà existant dans le sérum sanguin<sup>1</sup>.

Au 19<sup>ème</sup> siècle, Claude Bernard (1813-1878) décrivit que le glucose n'est pas présent dans le sang uniquement après un repas ou en conditions pathologiques comme le diabète, mais qu'il est également présent chez les individus normoglycémiques, même à jeun. Sa découverte du glycogène dans le foie le laissa suggérer que le sucre absorbé par les intestins

est converti en glycogène dans le foie, et que celui-ci peut être dégradé en glucose et libéré dans le sang en conditions de jeun<sup>1</sup>. En 1889, le médecin allemand Oskar Minkowski (1858-1931) et son collègue, le physicien Joseph von Mering (1849-1908) démontrèrent que l'ablation du pancréas chez le chien entraîne un diabète sévère, et émirent l'hypothèse qu'une sécrétion interne à cet organe pourrait être impliquée dans le métabolisme du glucose, mettant en avant l'implication de cet organe dans la pathologie<sup>1</sup>. C'est en 1893 que le médecin légiste Edouard Laguesse (1861-1927) suggéra que cette sécrétion pourrait être produite par les îlots pancréatiques, appelé « îlots de Langerhans » en référence à Paul Langerhans, qui fut le premier à les avoir observés<sup>1</sup>. Cette sécrétion fut renommée « insuline », venant du Latin et signifiant « île », d'abord par le physiologiste belge Jean de Meyer en 1909, puis par le physiologiste britannique John James Rickard Macleod dans les années 1920<sup>1</sup>.

Ce fut dans le laboratoire de Macleod à Toronto que l'insuline et son action furent découvertes. Les résultats de Frederick Banting (1891-1941), chirurgien orthopédique et Charles Best (1899-1978), étudiant en médecine à l'époque, furent publiés en 1922. Ils décrivent que l'injection d'extraits de pancréas de bœuf, préparés par James Bertram Collip (1892-1965) chez un jeune patient diabétique permet d'améliorer la glycémie ainsi que les symptômes cliniques. Grâce à leurs travaux sur l'insuline, Banting et Macleod obtinrent en 1923 le prix Nobel de Physiologie et de Médecine, qu'ils partagèrent avec leurs collègues Best et Collip<sup>1</sup>. Cependant, c'est uniquement en 1938 que Franck Young (1908-1988) suggéra que l'insuline est produite par les cellules  $\beta$  des îlots de Langerhans, ce qui fut confirmé par l'Américain Paul Lacy (1924-2005) en 1959<sup>1</sup>. Sa séquence en acides aminés fut découverte en 1955 par le biochimiste anglais Frederic Sanger (1918-2013), et sa structure tridimensionnelle par la chimiste britannique Dorothy Hodgkin (1910-1994), qui leur valut à chacun l'obtention d'un prix Nobel de Chimie<sup>1</sup>.

Dans les années 1930, Wilhelm Falta (1875-1950) et Harold Himsworth (1905-1993) décrivent que certains individus diabétiques sont plus sensibles à l'effet hypoglycémiant de l'insuline, et que d'autres, au contraire, sont insensibles ou résistants à son action. Les individus insulinosensibles ont été décrits comme minces, et nécessitant un traitement à l'insuline pour prévenir l'acidocétose, tandis que les individus insulino-résistants ont été décrits comme étant généralement plus vieux, obèses, et ne présentant pas d'acidocétose<sup>1</sup> : ce furent les prémices de la distinction entre le DT1 (diabète insulino-dépendant) et le DT2 (diabète

non-insulinodépendant), qui aura officiellement lieu suite à la classification des différents types de diabète réalisée par le National Diabetes Data Group en 1979<sup>3</sup>.

## 2 Définition et diagnostic du diabète

### 2.1 Définition du diabète

Le diabète est défini comme un trouble métabolique chronique ayant des étiologies multiples, et est caractérisé par une hyperglycémie et des troubles du métabolisme des carbohydrates, des graisses et des protéines, provoqués par des défauts de la sécrétion et/ou de l'action de l'insuline, hormone hypoglycémiante produite et sécrétée par les cellules  $\beta$  pancréatiques<sup>4</sup>.

Typiquement, le diabète peut être catégorisé en plusieurs types selon son étiologie :

- Le DT1, qui correspond à 5-10% des cas, et apparaît chez les individus plutôt jeunes. Il est provoqué par une destruction des cellules  $\beta$  pancréatiques sécrétrices d'insuline, menant à un déficit de cette hormone. Chez la majorité de ces individus, une réaction auto-immunitaire est présente et participe à la destruction des cellules  $\beta$  pancréatiques, comme en témoigne la présence d'auto-anticorps dirigés contre les cellules des îlots pancréatiques, notamment contre l'isoforme 65 kDa de la glutamate décarboxylase (GAD65), l'insuline, les tyrosine phosphatases islet antigen 2 (IA-2) et islet antigen 2 $\beta$  (IA-2 $\beta$ ), ou le transporteur de zinc 8 (ZNT8)<sup>5</sup>.
- Le DT2, qui correspond à plus de 90% des cas, et apparaît généralement chez les personnes plus âgées. Il est causé par une diminution relative de l'insulinosécrétion, couplée à une insulino-résistance<sup>5</sup>. La suite de cet écrit se focalisera uniquement sur ce type de diabète, et notamment sur son étiologie génétique.
- Le diabète gestationnel. Ce diabète est diagnostiqué au cours du 2<sup>ème</sup> ou du 3<sup>ème</sup> trimestre de grossesse, chez les femmes qui ne présentaient pas de signe de diabète auparavant<sup>5</sup>. Cependant, sa présence durant la grossesse augmente le risque d'être diabétique plus tard dans la vie<sup>5</sup>, et peut également entraîner des complications chez l'enfant, parmi lesquelles on peut citer une malformation congénitale, une hypoglycémie néonatale, une naissance prématurée, une macrosomie ou encore un décès intra-utérin du fœtus<sup>6</sup>.

- Les autres types de diabète, tels que celui associé aux maladies du pancréas exocrine (pancréatite, fibrose kystique), ou le diabète induit par les médicaments et les substances chimiques (glucocorticoïdes, antirétroviraux)<sup>6</sup>.

Le diabète peut se manifester par différents symptômes cliniques, comme la soif, la polyurie, la perte de poids, ou encore des troubles de la vision. Cependant, dans le cas du DT2, ces symptômes sont généralement faibles ou absents, et le diabète peut se développer et s'aggraver durant des années avant que les patients n'aillent consulter<sup>7</sup>.

## 2.2 Diagnostic du diabète

Les critères de diagnostic du diabète ont été établis par l'American Diabetes Association (ADA) et par l'Organisation Mondiale de la Santé (OMS), et sont multiples (figure 1)<sup>5,8</sup>.

Un diabète peut être diagnostiqué suite à une mesure de la glycémie à jeun égale ou supérieure à 126 mg/dL, où la période de jeun correspond à l'absence d'ingestion de calories pendant au moins 8 heures. Le diagnostic du diabète peut également être établi grâce à un test de tolérance oral au glucose, qui consiste à ingérer 75 g de glucose, dilués dans de l'eau, puis à mesurer la glycémie 2 heures plus tard. Si celle-ci est supérieure à 200 mg/dL, le patient est considéré comme étant diabétique. En outre, le taux d'hémoglobine glyquée (HbA1c) peut également être utilisée comme outil de diagnostic. Celui-ci correspond à la fraction glyquée de l'hémoglobine, et sert ainsi d'indicateur de la glycémie sur les 3 mois qui précèdent (correspondant à la durée de vie des globules rouges, qui est de 120 jours). Ainsi, si un patient présente un taux de HbA1c égal ou supérieur à 6,5%, on peut affirmer qu'il est diabétique. Ces 3 méthodes sont considérées comme équivalentes pour détecter la pathologie<sup>5</sup>. Finalement, la mesure aléatoire d'une glycémie égale ou supérieure à 200 mg/dL chez un patient présentant une hyperglycémie ou une crise d'hyperglycémie permet également d'établir un diagnostic<sup>5</sup>.

Cependant, en dehors de signes flagrants de diabète, il est nécessaire de confirmer le diagnostic, soit en testant à nouveau le patient pour le même critère de diagnostic, soit en réalisant un nouveau test, basé sur un autre critère<sup>5</sup>.

Chez certains patients, les mesures de la glycémie sont trop faibles pour considérer un diabète à proprement parler, mais sont plus élevées que les valeurs normales : on parle alors de prédiabète. Le prédiabète contribue à l'augmentation du risque de développement d'un diabète ou d'une maladie cardiovasculaire. Il est généralement associé à une obésité, une dyslipidémie couplée à hypertriglycéridémie et/ou à un faible taux de cholestérol transporté par les lipoprotéines à haute densité (high density lipoproteins [HDL]), et une hypertension. Suivant les critères de l'ADA, on parle d'anomalie de la glycémie à jeun (impaired fasting glucose [IFG]) lorsque la mesure de la glycémie à jeun est comprise entre 100 et 125 mg/dL, et d'intolérance au glucose (impaired glucose tolerance [IGT]) lorsque, 2h après un test d'hyperglycémie provoquée par voie orale (oral glucose tolerance test [OGTT]), la glycémie se situe entre 140 et 200 mg/dL. L'IFG et l'IGT permettent de déterminer si un patient est prédiabétique. Le taux de HbA1c, lorsqu'il est compris entre 5,7 et 6,4%, peut également être considéré comme un indicateur de prédiabète<sup>5</sup>. Ces critères sont repris à la figure 1.

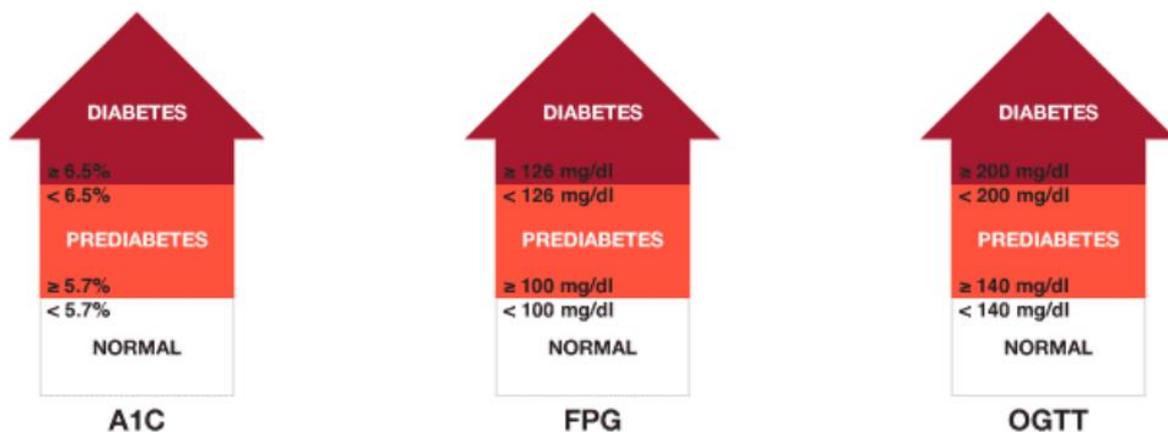


Figure 1. Diagnostic du diabète et du prédiabète selon les critères établis par l'ADA. **A1C** : taux d'hémoglobine glyquée, **FPG** : glycémie à jeun, **OGTT** : test d'hyperglycémie provoquée par voie orale. Source : site web de l'American Diabetes Association (<https://www.diabetes.org/a1c/diagnosis>).

### 3 Complications du diabète

#### 3.1 Complications à court terme

A court terme, le diabète peut entraîner des complications associées à un déficit relatif ou absolu en insuline, pouvant provoquer une déshydratation et un trouble de l'état mental.

- L'acidocétose diabétique est surtout observée chez les patients atteints de DT1, généralement plus jeunes<sup>9</sup>. Chez ces patients, suite à un déficit en insuline, les hormones antagonistes comme le glucagon (hormone hyperglycémiant) sont

favorisées. Ce déséquilibre entraîne une augmentation de la concentration en glucose libre dans le sang par glycogénolyse et néoglucogénèse hépatiques, une réduction de son utilisation par les tissus périphériques, ainsi qu'une élévation des acides gras libres, issus de la lipolyse du tissu adipeux. Sous l'action du glucagon et l'augmentation du rapport glucagon/insuline, les acides gras sont oxydés en corps cétoniques (acétoacétate et  $\beta$ -hydroxybutyrate) dans le foie, qui vont alors contribuer à l'acidose métabolique. A l'examen physique, on observe souvent chez ces patients des signes de déshydratation, une tachycardie, et une hypotension. L'acidose métabolique provoque chez eux des douleurs abdominales, des nausées et des vomissements, ainsi qu'une léthargie et une dépression du système nerveux central, qui peut aboutir au coma<sup>9</sup>.

- L'état d'hyperosmolarité hyperglycémique est quant à lui plutôt observé chez les personnes plus âgées atteintes de DT2<sup>9</sup>. Chez ces patients, l'hyperglycémie induite par le déficit en insuline entraîne une diurèse osmotique, qui n'est pas compensée par la prise de boissons, menant à une diminution du volume sanguin. Comme en cas d'acidocétose diabétique, ces patients présentent généralement des signes de déshydratation et d'hypotension. On retrouve chez eux plusieurs symptômes neurologiques dont une altération de l'état mental, une léthargie, une perte de sensibilité, et des crises d'épilepsie, menant parfois à un coma<sup>9</sup>.

### 3.2 Complications à long terme

A long terme, le diabète peut entraîner plusieurs complications qui participent à la mortalité et la morbidité de cette maladie.

Après des années de diabète, on observe, chez les patients, une altération des petits vaisseaux sanguins<sup>10</sup>. Ces altérations provoquent des complications microvasculaires, entravant le bon fonctionnement de plusieurs organes.

- La rétinopathie diabétique est la cause principale de cécité chez les adultes âgés entre 20 et 74 ans<sup>11</sup>. Elle n'apparaît généralement qu'après 15 ans de diabète<sup>12</sup>. Elle est caractérisée par des microanévrismes et la dégénération des capillaires rétiens, une néovascularisation excessive, et des modifications de la perméabilité vasculaire<sup>13</sup>. La plupart des patients sont atteints d'une phase précoce, appelée rétinopathie non-proliférante, n'altérant pas la vision. Cependant, chez 50% des patients, elle évolue

vers une rétinopathie proliférante, pouvant mener à terme à une perte totale de la vue<sup>10</sup>.

- La néphropathie diabétique représente la majorité des insuffisances rénales terminales dans le monde occidental<sup>14</sup>. Les signes de néphropathie n'apparaissent généralement pas avant une dizaine d'années de diabète, et ne sont observés qu'après 35 ans de maladie chez la plupart des patients<sup>10</sup>. Cette complication est provoquée par une altération de la structure du rein, qui perd peu à peu sa capacité à retenir les protéines, menant à une protéinurie, suivie d'une diminution du taux de filtration glomérulaire du rein, aboutissant au final à une maladie rénale chronique<sup>13</sup>.
- La neuropathie, qui se développe chez plus de la moitié des patients<sup>15</sup>, peut apparaître à l'importe quel moment du diabète<sup>12</sup>. Généralement, les patients diabétiques présentent une neuropathie sensorielle périphérique, qui les empêche de ressentir les lésions dans la région distale de leurs membres<sup>16</sup>. La neuropathie, couplée à une mauvaise perfusion du tissu, peut entraîner l'apparition d'ulcères du pied et aboutir à une amputation du membre<sup>10</sup>. Cependant, des atteintes motrices ou du système nerveux autonome, menant à des dysfonctions gastriques, vasomotrices ou encore sexuelles sont également possibles<sup>10</sup>. La neuropathie peut également entraîner des douleurs du type hyperalgésie, paresthésie, et allodynie chez 40 à 50% des sujets atteints<sup>17</sup>.

Les complications microvasculaires ne sont pas les seules à être présentes chez les patients diabétiques. Ainsi, on peut également observer chez eux des complications macrovasculaires, qui sont les principales causes de décès chez les patients. Celles-ci sont provoquées par un rétrécissement progressif des parois artérielles, appelé athérosclérose<sup>18</sup>. Ces complications correspondent à 44% et 52% des décès chez les patients atteints de DT1 et de DT2, respectivement<sup>19</sup>. Elles peuvent être de plusieurs types, telles que les coronaropathies, les accidents vasculaires cérébraux, ou encore les maladies cardiovasculaires<sup>20</sup>.

Le diagnostic du diabète est donc très important afin de permettre une prise en charge correcte du patient.

## 4 Epidémiologie du diabète

### 4.1 Situation dans le monde

Le diabète est un problème mondial : cette maladie affecte 463 millions de personnes adultes (décrites dans cette section comme étant âgées entre 20 et 79 ans) dans le monde, et on estime qu'environ 80% d'entre elles vivent dans des pays à faibles revenus ou à revenus intermédiaires. Le nombre de personnes diabétiques devrait atteindre 578 millions en 2030, et 700 millions en 2045, soit une augmentation de 51% des cas en 25 ans<sup>21</sup>. La prédiction de l'évolution de cette maladie en fonction des continents est représentée à la figure 2. Il a été estimé qu'en 2019, 4,2 millions d'adultes sont morts de causes liées à cette pathologie<sup>22</sup>. De plus, plus de la moitié des personnes vivant avec un diabète ne sont pas diagnostiquées<sup>21</sup>.

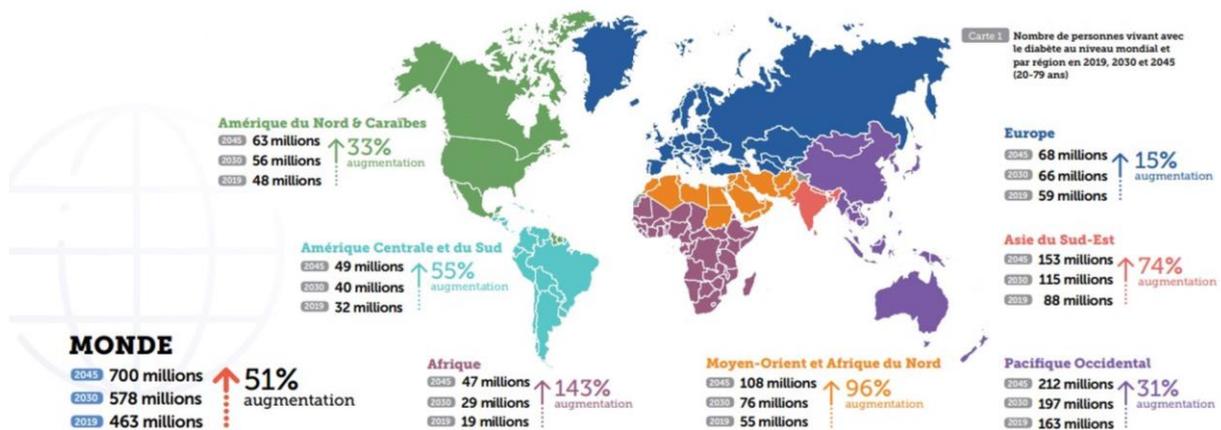


Figure 2. Représentation du nombre de de cas de diabète dans le monde, et prédiction de son évolution dans les prochaines années. Source : Fédération Internationale du Diabète, « Atlas du diabète de la FID », 9<sup>ème</sup> édition, 2019.

Divers facteurs ont un impact sur le risque de développer un diabète :

- La prévalence du diabète augmente avec l'âge. Ainsi, la tranche d'âge la moins touchée correspond aux personnes âgées de 20 à 24 ans, avec seulement 1,4% d'entre eux atteints, et la tranche d'âge la plus affectée correspond aux personnes âgées de 75 à 79 ans, pour lesquelles la prévalence a été estimée à 19,9% en 2019, et devrait croître dans les années qui suivent. En 2019, il a été estimé que le nombre de personnes âgées de 65 à 99 ans et atteintes du diabète est de 135,6 millions, et atteindra vraisemblablement 276,2 millions de personnes en 2045<sup>22</sup>.

- La prévalence du diabète varie également en fonction du sexe. Chez les femmes adultes, elle est légèrement plus faible (9%) que chez les hommes adultes (9,6%), correspondant à 17,2 millions d'hommes en plus que les femmes en 2019<sup>22</sup>.
- Le lieu de vie a également un impact important. La prévalence d'individus diabétiques vivant en zone urbaine (10,8%) est également plus élevée que la prévalence en zone rurale (7,2%), et cette prévalence devrait fortement augmenter chez les individus vivant en milieu urbain durant les années qui viennent, pour atteindre 12,5% en 2045<sup>22</sup>.
- La prévalence du diabète varie également en fonction de la région géographique considérée. Ainsi, le Moyen-Orient et l'Afrique du nord sont les régions ayant la plus forte prévalence (12,2% en 2019), alors que le reste du continent africain correspond à la région ayant la prévalence la plus faible (4,7%). Ces différences peuvent résulter d'une urbanisation plus faible, d'une malnutrition, ou d'un plus faible taux d'obésité et de surpoids dans la région africaine<sup>22</sup>. Concernant le nombre de cas par pays, ce sont, par ordre décroissant, la Chine, l'Inde et les Etats-Unis qui sont les plus touchés. Cependant, ils ne correspondent pas aux pays ayant la plus forte prévalence, qui sont les îles Marshall (30,5%), Kiribati (22,5%), et le Soudan (22,1%)<sup>21</sup>. En Europe, 8,9% de la population est atteinte de diabète, soit 53,9 millions d'adultes. Sur notre continent, en 2019, ce sont 465 900 adultes qui sont décédés du diabète ou de ses complications, et 31,4% d'entre eux étaient âgés de moins de 60 ans<sup>22</sup>.

Les coûts directs et indirects liés au diabète sont conséquents. Les coûts directs correspondent aux dépenses de santé liées au diabète réalisées par les patients, l'état, ou des organismes privés et publics. On estime qu'en 2019, les dépenses directes mondiales auraient atteint les 760 milliards de dollars, soit une augmentation de 4,5% par rapport à 2017. Les coûts indirects représentent quant à eux la diminution de la population active, la mortalité (diminution de la productivité présente et future), l'absentéisme (nombre de jours de travail manqués à cause de problèmes de santé) et le présentéisme (diminution de la production lors de la présence sur le lieu de travail à cause des conditions de santé). En 2015, il a été estimé que ces coûts correspondaient à 34,7% des dépenses mondiales liées au diabète<sup>23</sup>. On estime

que chaque année, aux Etats-Unis, les décès prématurés liés au diabète coûtent 19,9 milliards de dollars, et 90 milliards de dollars sont perdus dans les coûts indirects liés au diabète<sup>24</sup>.

## 4.2 Situation en France

En France, un rapport épidémiologique établi en 2016 rapporte que plus de 3,3 millions de personnes (soit 5% de la population) sont traitées pharmacologiquement pour un diabète, toutes formes confondues<sup>25</sup>. En 2019, les dépenses liées au diabète en France ont été estimées à 16,9 milliards de dollars<sup>22</sup>.

Au niveau géographique, le taux de prévalence le plus important a été observé dans les départements d'Outre-mer, où il est 1,5 à 2 fois plus élevé que la prévalence nationale. En métropole, en comparaison avec la prévalence nationale, c'est la Seine-Saint-Denis qui possède le taux de prévalence le plus fort (1,5), suivi du Val d'Oise et du Pas-de Calais (1,3), puis de l'Aisne, du Nord et des Ardennes (1,2). Les départements bretons sont quant à eux les moins touchés, avec un taux de prévalence variant de 0,6 à 0,7, en fonction des régions<sup>26</sup> (figure 3).

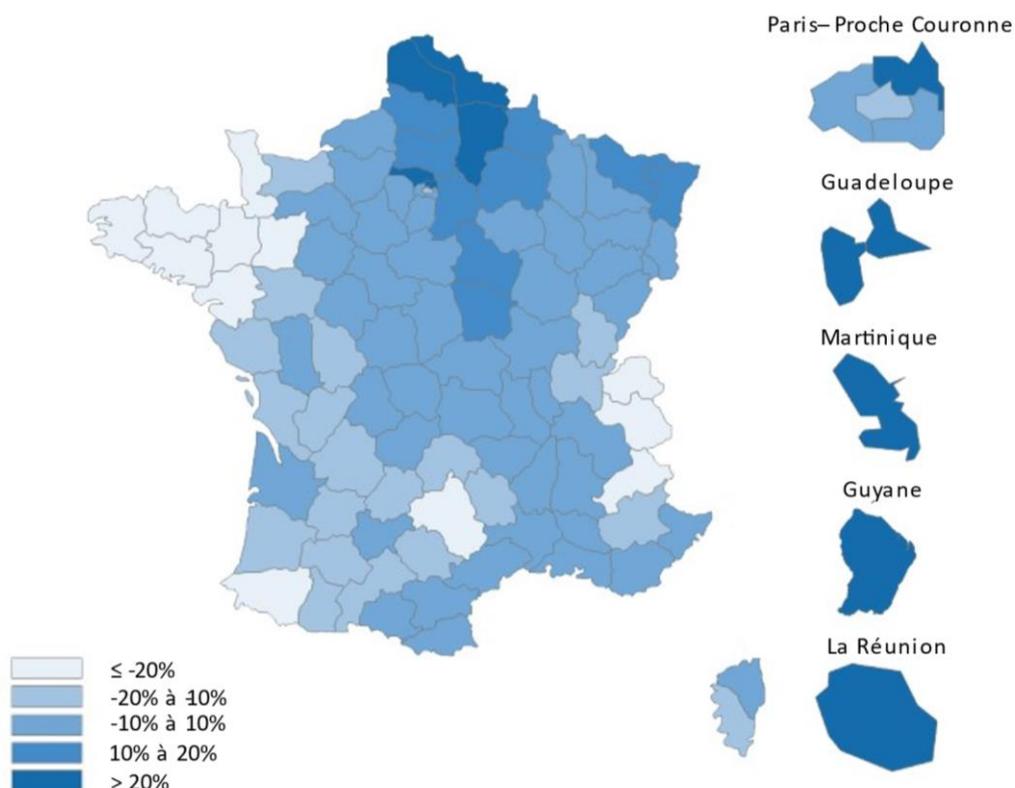


Figure 3. Variation de la prévalence du diabète traité pharmacologiquement par rapport à la prévalence nationale en fonction des départements. Source : adapté de Santé Publique France, « Le poids du diabète en France. Synthèse épidémiologique », 2016.

Comme observé au niveau mondial, les hommes sont légèrement plus touchés (1,8 millions) que les femmes (1,5 millions), et la prévalence augmente en parallèle de l'âge : 1 homme sur 5, entre 70 et 85 ans, et 1 femme sur 7, entre 75 et 85 ans, sont traités pour un diabète. Le diabète est plus fréquent chez les personnes défavorisées, qui sont également plus fréquemment hospitalisées pour des complications associées à la maladie<sup>25</sup>.

Les complications du diabète ne sont pas négligeables. En France en 2016, sur 100 000 personnes traitées pharmacologiquement pour un diabète, 244 ont été hospitalisées suite à un infarctus du myocarde, 569 pour un accident vasculaire cérébral, 805 pour une plaie du pied, 255 ont subi une amputation d'un membre inférieur, et 133 ont eu une greffe rénale/ont été mises sous dialyse, avec un âge moyen de ces patients supérieur à 69 ans<sup>25</sup>.

## 5 Organes touchés par le diabète

Le DT2 impacte la fonction de plusieurs organes importants dans la régulation du métabolisme énergétique. L'organe principalement touché est le pancréas, contenant les cellules  $\beta$  pancréatiques sécrétrices d'insuline. Effet, le DT2 est associé à une détérioration progressive de ces cellules, qui peut être provoquée par des facteurs génétiques (sur lesquels nous reviendrons plus loin) ou environnementaux (glucotoxicité, lipotoxicité, sécrétion de cytokines), altérant de ce fait la sécrétion d'insuline<sup>27</sup>.

En outre, cette pathologie est généralement associée à une insulino-résistance des organes cibles de l'insuline, parmi lesquels le muscle squelettique, le tissu adipeux et le foie, qui ont un rôle majeur dans l'homéostasie énergétique. L'insulino-résistance est définie comme une incapacité de l'insuline sécrétée à exercer ses effets biologiques sur les organes cibles, à des concentrations physiologiques normales. Ces organes ont donc besoin d'être stimulés avec de plus grandes concentrations d'insuline afin d'obtenir une réponse biologique normale<sup>28</sup>.

La détérioration des cellules  $\beta$  pancréatiques, réduisant la sécrétion d'insuline, et la présence d'une insulino-résistance des tissus cibles, répondant plus faiblement à cette hormone, sont deux processus liés. Concrètement, suite à l'insulino-résistance des tissus cibles, les cellules  $\beta$  pancréatiques sécrètent dans un premier temps de plus grandes quantités d'insuline, afin de maintenir la glycémie à des niveaux corrects. Cependant, au cours du temps, la capacité des cellules  $\beta$  pancréatiques à sécréter de l'insuline décroît. A terme, ces cellules

ne sont plus capables de sécréter suffisamment d'insuline pour maintenir une homéostasie glucidique. On observe alors l'apparition d'une hyperglycémie, ainsi qu'une altération de la réponse des organes cibles à l'insuline, altérant de ce fait le métabolisme énergétique chez les patients<sup>27,28</sup>.

Ces différents organes, impliqués dans la régulation du métabolisme énergétique, sont décrits en détails ci-dessous.

## 5.1 Le pancréas

Le pancréas est l'organe central de la physiopathologie du DT2. C'est un organe rétropéritonéal, situé dans la partie supérieure de la cavité abdominale. Chez l'homme, il mesure entre 15 et 25 cm de long, et pèse environ 100 à 150 g. Il est fortement vascularisé par plusieurs artères, notamment les artères pancréatico-duodénales supérieure et inférieure, et l'artère hépatique, mais est également bien innervé par les systèmes sympathique et parasympathique. Sa structure est divisée en 3 parties : la tête, qui s'ancre dans la première boucle du duodénum, le corps, et la queue du pancréas qui est à proximité de la rate (figure 4). Le pancréas est en réalité un organe mixte composé de 2 glandes : une glande exocrine, et une glande endocrine<sup>29</sup>.

La glande exocrine du pancréas correspond à environ 80% de son volume, et est composée de cellules pancréatiques acinaires et de cellules ductales. Les cellules acinaires sont majoritaires dans le pancréas et sont sécrétrices d'enzymes digestives, à savoir les  $\alpha$  amylases qui digèrent les carbohydrates, les lipases qui dégradent les lipides, et les protéases qui digèrent les protéines<sup>30</sup>. Les cellules ductales, quant à elles, représentent seulement 10% des cellules pancréatiques<sup>31</sup> et sécrètent du bicarbonate de sodium  $\text{HCO}_3^-$ , dont le pH alcalin est nécessaire à la neutralisation du pH acide du chyme gastrique dans le duodénum afin d'y activer les enzymes digestives, tout en protégeant les muqueuses de lésions liées à une trop forte acidité<sup>30</sup>. Structurellement, les cellules acinaires sont de forme pyramidale, et forment des grappes de cellules, appelées « acini » où elles sont orientées de façon à ce que leur membrane apicale soit dirigée vers le centre de l'acinus. Ces acini se regroupent entre eux afin de former des lobules dans le pancréas. La lumière au centre de chaque acinus contient un canalicule intercellulaire, formé par des cellules ductales, qui permet de recueillir les enzymes sécrétées par les cellules acinaires. Les canalicules de chaque acinus convergent ensuite en

canaux intralobulaires, puis en canaux extralobulaires, tous formés de cellules ductales. Le suc pancréatique ainsi formé est ensuite dirigé par les canaux extralobulaires vers le canal pancréatique principal de Wirsung qui se jette dans l'ampoule de Vater, où est également déversée la bile produite par le foie. Ce suc est alors excrété dans le duodénum, où il peut participer à la digestion<sup>29</sup> (figure 4).

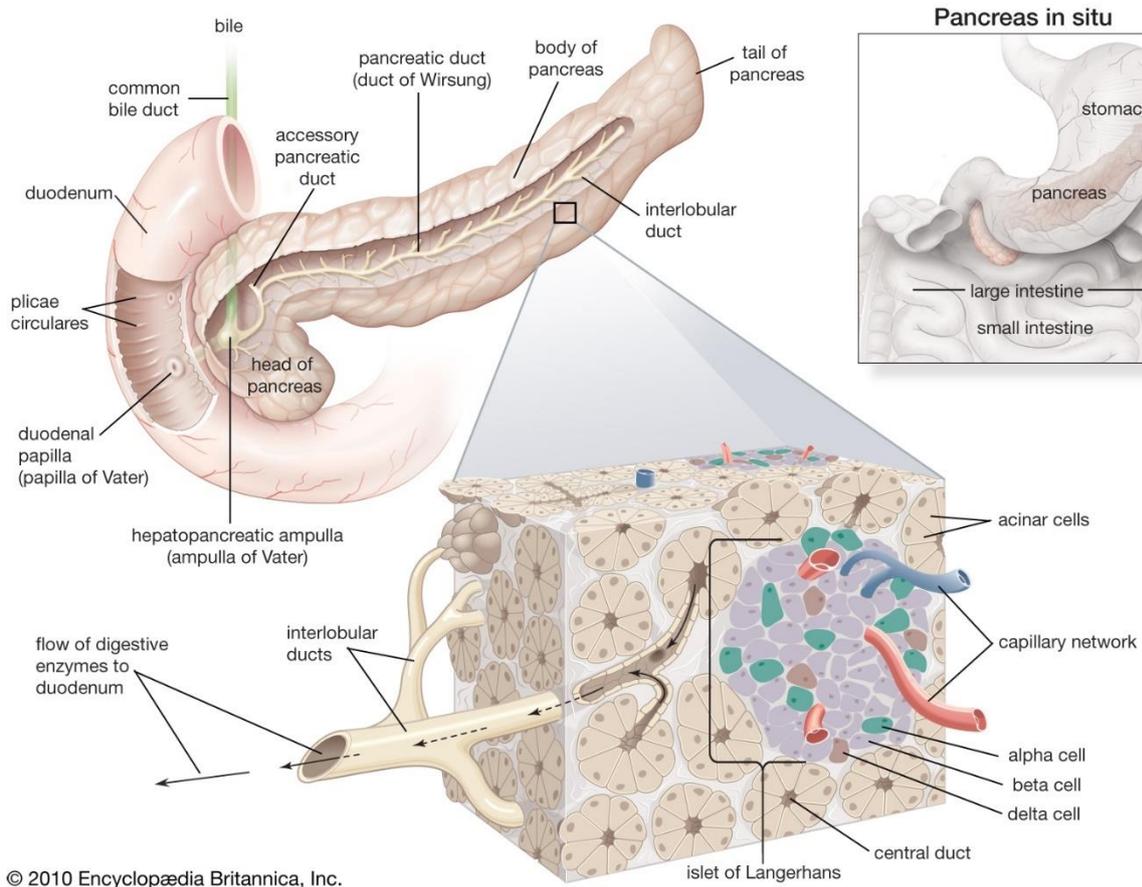


Figure 4. Structure et organisation du pancréas. Source : Encyclopædia Britannica, 2010 (<https://www.britannica.com/science/organ-biology#/media/1/623962/68636>), consulté le 1<sup>er</sup> août 2021.

Le pancréas endocrine est composé de cellules endocrines regroupées ensemble afin de former des îlots pancréatiques ou îlots de Langerhans (figure 4). Ces îlots, de forme sphérique, correspondent à environ 2% de la masse du pancréas, et sont fortement vascularisés : environ 15% de la vascularisation du pancréas est dédiée aux îlots, chacun irrigué par un réseau de capillaires dense en son sein<sup>32</sup>. Les îlots sont aussi richement innervés par les terminaisons de fibres nerveuses autonomes sympathiques et parasymphatiques, permettant de réguler la sécrétion d'insuline et de glucagon<sup>30,32</sup>.

Les îlots de Langerhans sont composés de 5 types cellulaires différents<sup>29,32</sup> :

- Les cellules  $\alpha$  pancréatiques, produisant le glucagon, hormone hyperglycémiant. Elles composent 20 à 30% des cellules de l'îlot pancréatique.
- Les cellules  $\beta$  pancréatiques, produisant l'insuline. C'est le type cellulaire majoritaire dans les îlots, correspondant à 60% des cellules le composant. Le fonctionnement de ces cellules sera expliqué en détails plus loin.
- Les cellules  $\delta$  pancréatiques, qui produisent la somatostatine. Environ 10% des cellules des îlots font partie de ce type cellulaire.
- Les cellules PP, produisant le polypeptide pancréatique PP. Elles correspondent à moins de 5% des cellules des îlots.
- Les cellules  $\epsilon$ , qui produisent la ghreline. Elles représentent uniquement 1% des cellules de l'îlot.

La disposition et la proportion de ces différents types cellulaires dans les îlots diffèrent en fonction des espèces. Ainsi, chez les rongeurs, les cellules  $\beta$  pancréatiques sont plus nombreuses, et sont situées au centre de l'îlot, composant le « corps » de celui-ci, entouré des cellules  $\alpha$  pancréatiques, moins nombreuses, et  $\delta$  pancréatiques, qui correspondent au « manteau ». Chez l'humain, les cellules  $\beta$  pancréatiques sont moins nombreuses, mais les cellules  $\alpha$  pancréatiques sont présentes en plus grand nombre, et les différents types cellulaires sont beaucoup plus dispersés<sup>33</sup> (figure 5).

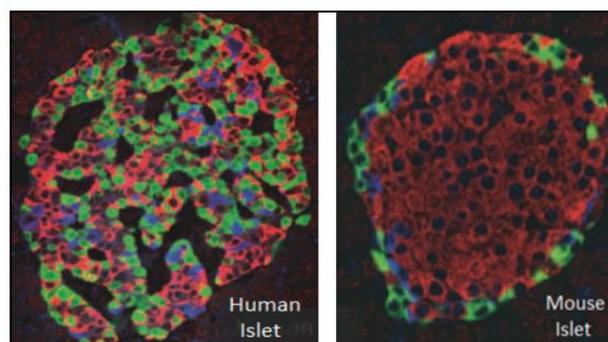


Figure 5. Comparaison de l'architecture des îlots de Langerhans chez l'humain (à gauche) et chez la souris (à droite). L'insuline est marquée en rouge, le glucagon en vert, et la somatostatine en bleu. Source : Levetan and Pierce, « Distinctions Between the Islets of Mice and Men: Implications for New Therapies for Type 1 and Type 2 Diabetes », *Endocrine Practice*, 2012.

### 5.1.1 La cellule $\beta$ pancréatique

La cellule  $\beta$  pancréatique, de par son exclusivité à sécréter de l'insuline, est le type cellulaire au cœur de la régulation de l'homéostasie glucidique. Suite à un stimulus, comme la présence de nutriments dans le sang à la suite d'un repas, ces cellules  $\beta$  pancréatiques sont

sollicitées afin d'augmenter la production et la sécrétion d'insuline. 10% des protéines composant ces cellules sont des molécules d'insuline, soit environ 10 pg d'insuline par cellule<sup>32</sup>.

#### 5.1.1.1 Synthèse de l'insuline

L'insuline est l'hormone principale de la régulation de la glycémie. Il s'agit d'un polypeptide de 51 acides aminés et d'une masse de 5,8 kDa. Un monomère d'insuline est composé d'une chaîne A de 21 acides aminés et d'une chaîne B de 30 acides aminés, et contient 3 ponts disulfure permettant le repliement correct de la protéine : 2 de ces ponts relient les chaînes A et B entre elles, tandis que le 3<sup>ème</sup> pont disulfure est interne à la chaîne A<sup>34</sup>. Son précurseur, la préproinsuline, est un polypeptide de 110 acides aminés et de 11,5 kDa, codé par le gène *INS* situé sur le bras court du chromosome 11<sup>32</sup>. Ce gène, d'une longueur de 1355 paires de bases, est composé de 3 exons, codant respectivement le peptide signal, la chaîne B de la préproinsuline ainsi qu'une partie du peptide C, et finalement la seconde partie du peptide C ainsi que la chaîne A. Une fois transcrit et épissé, l'acide ribonucléique (ARN) messenger (ARNm) obtenu est traduit pour donner naissance à la préproinsuline<sup>32</sup>.

La biosynthèse de l'insuline est stimulée par une augmentation de la concentration en glucose intracellulaire, suite à un repas, par exemple. Il existe une relation sigmoïdale entre la concentration en glucose et la biogenèse de l'insuline, qui a pour seuil une concentration en glucose de 2 à 4 mmol/L<sup>32</sup>. Ce processus, expliqué ci-dessous, est illustré aux figures 6 et 7.

La traduction de l'ARNm par les ribosomes du réticulum endoplasmique rugueux permet de synthétiser la préproinsuline, qui est dirigée dans les citernes du réticulum endoplasmique grâce à son peptide signal. Ce peptide signal est alors clivé par une signal peptidase pour générer la proinsuline de 9 kDa, composée uniquement des chaînes A et B reliées entre elles par le peptide C. La proinsuline subit alors son repliement, ainsi que la formation des 3 ponts disulfure dont la stabilité est favorisée par le peptide C<sup>32</sup>.

La proinsuline est ensuite transportée du réticulum endoplasmique vers le cis-Golgi grâce à des vésicules de transport, puis est transloquée vers le trans-Golgi où elle est empaquetée dans des vésicules de sécrétion immatures<sup>35</sup>. Durant la maturation de ces granules, leur acidification (pH 5,5) et la forte concentration en ions Ca<sup>2+</sup> induisent des modifications enzymatiques importantes de la proinsuline<sup>36</sup>. Dans un premier temps, les

endopeptidases prohormone convertases 2 (PC2) et 3 (PC3) clivent la protéine à l'extrémité C-terminale des résidus dibasiques Lys<sup>64</sup>-Arg<sup>65</sup> à la jonction de la chaîne A et du peptide C et Arg<sup>31</sup>-Arg<sup>32</sup> à la jonction de la chaîne B et du peptide C, respectivement, puis la carboxypeptidase H clive ces résidus basiques, libérant de cette façon le peptide C et l'insuline<sup>32</sup>. L'insuline et le peptide C ainsi obtenus sont stockés ensemble dans les granules de sécrétion. Chaque cellule  $\beta$  pancréatique contient environ 10 000 granules de sécrétion, contenant chacune 250 000 molécules d'insuline<sup>36</sup>. Par microscopie électronique, les granules de sécrétion d'insuline ressortent comme des granules de 200 à 300 nm à cœur dense<sup>37</sup>. Le cœur de la granule est composé de cristaux insolubles d'insuline hexamérique, formés de 3 dimères d'insuline, et interagissant avec 2 ions zinc Zn<sup>2+</sup> au centre afin de stabiliser la formation du complexe, alors que le manteau de la granule contient le peptide C et d'autres molécules nécessaires à la vésicule<sup>34,38</sup>. Suite à une stimulation, l'insuline et le peptide C sont libérés par exocytose régulée, à un rapport 1 :1. Il est à noter que l'insuline sécrétée est composée à plus de 95% d'insuline mature, et plus faiblement de proinsuline<sup>32</sup>.

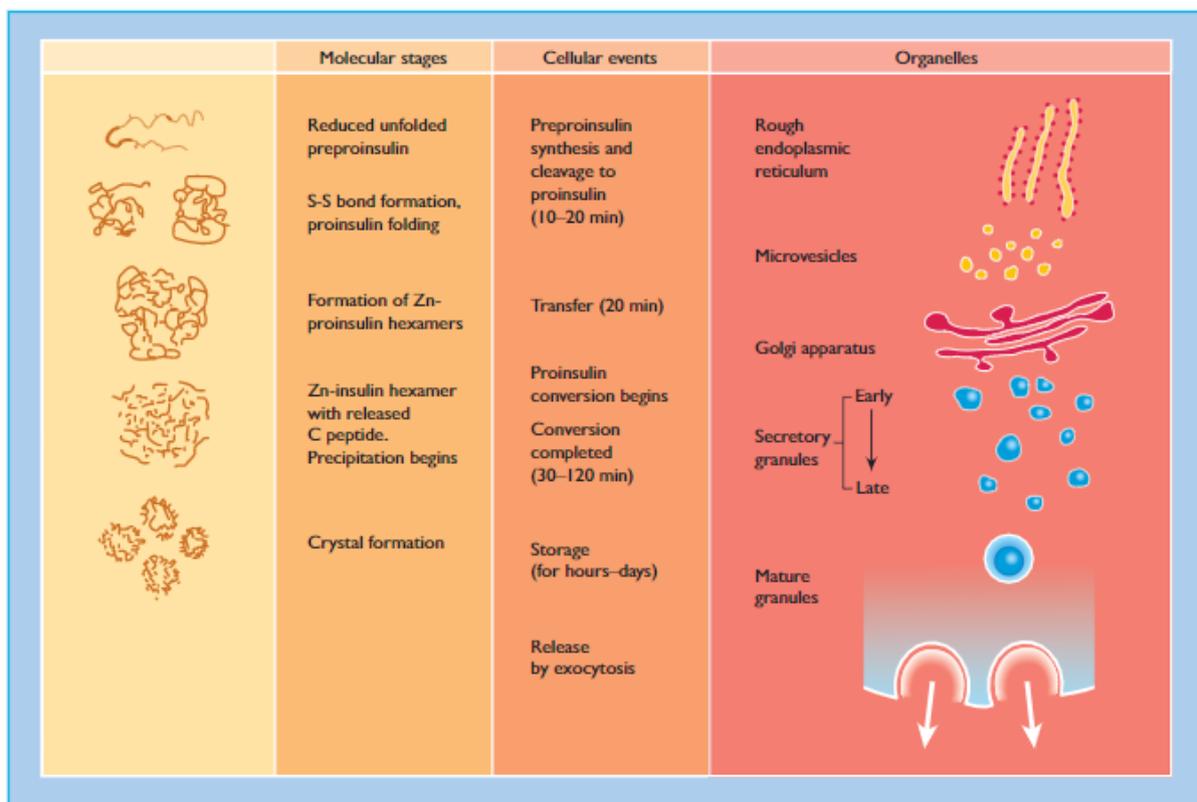


Figure 6. Processus de la biosynthèse de l'insuline. Source : Jones and Persaud, « Islet Function and Insulin Secretion », *Textbook of Diabetes*, 4<sup>ème</sup> édition, 2010.

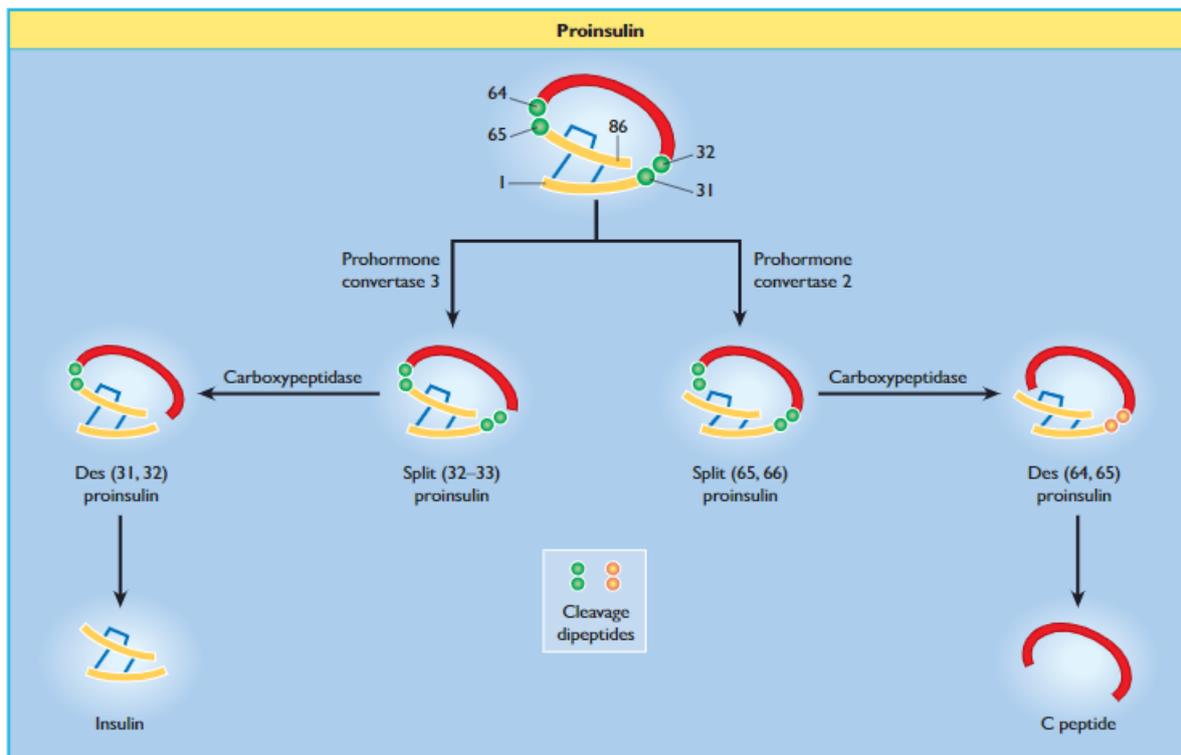


Figure 7. Processing de la proinsuline en insuline. Source : Jones and Persaud, « Islet Function and Insulin Secretion », Textbook of Diabetes, 4<sup>ème</sup> édition, 2010.

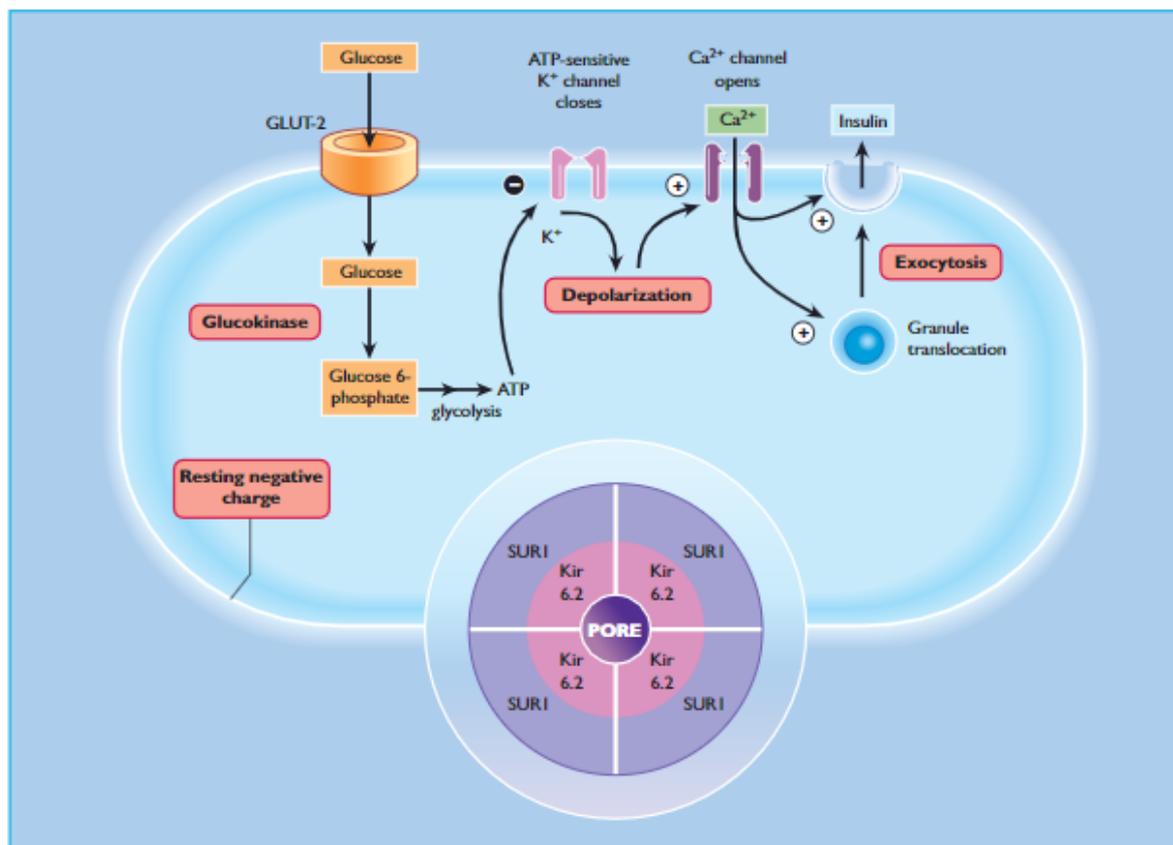


Figure 8. Représentation du mécanisme de sécrétion d'insuline induite par le glucose. Source : Jones and Persaud, « Islet Function and Insulin Secretion », Textbook of Diabetes, 4<sup>ème</sup> édition, 2010.

### 5.1.1.2 Sécrétion de l'insuline

#### 5.1.1.2.1 Stimulation de la sécrétion d'insuline

La stimulation des cellules  $\beta$  pancréatiques par le glucose entraîne l'exocytose de l'insuline, dont le processus est représenté à la figure 8. Après un repas, la concentration plasmatique en glucose augmente drastiquement, et ce glucose est internalisé par diffusion facilitée dans les cellules  $\beta$  pancréatiques par le transporteur de glucose 2 (GLUT2), dont la faible affinité pour le glucose ( $K_m = 15\text{-}20$  mM) permet un influx proportionnel à sa concentration extracellulaire<sup>35,39</sup>. Le glucose est alors phosphorylé en glucose-6-phosphate par la glucokinase (GCK), qui catalyse la première réaction de la glycolyse<sup>32</sup>. Cette enzyme de la famille des hexokinases possède une faible affinité pour le glucose ( $K_m = 6$  mM), et n'est pas inhibée par son produit, le glucose-6-phosphate. La capacité de cette enzyme à coupler la sécrétion d'insuline à la concentration extracellulaire en glucose lui a valu le nom de « senseur de glucose », et est l'étape limitante de la glycolyse<sup>32,40</sup>.

La métabolisation du glucose par la voie de la glycolyse se finalise par la production de pyruvate. Celui-ci est ensuite oxydé par le cycle de Krebs dans la mitochondrie, et génère ainsi de l'adénosine triphosphate (ATP), une molécule riche en énergie<sup>34</sup>. L'élévation de la concentration en ATP cytosolique, couplée à la diminution de la concentration en adénosine diphosphate (ADP), mène à la fermeture des canaux potassiques ATP-dépendants ( $K^+_{ATP}$ ) membranaires<sup>32</sup>. Le canal  $K^+_{ATP}$  permet coupler la métabolisation du glucose à l'activité électrique de la cellule  $\beta$  pancréatique, et ainsi à la sécrétion d'insuline<sup>32,38</sup>. Il est composé de 8 sous-unités réunies en un complexe hétéro-octamérique : 4 sous-unités inward-rectifier potassium channel 6.2 (Kir6.2), faisant partie de la famille des canaux à flux entrant dominant (inward rectifier channel), qui forment le pore laissant passer les ions  $K^+$  et capables de lier l'ATP (inhibiteur du canal), entourées de 4 sous-unités régulatrices sulonylurea receptor-1 (SUR1), de la famille des transporteurs à ATP binding cassette, capables de lier l'ADP (activateur du canal)<sup>32</sup>. L'ATP produit par la métabolisation du glucose, en se liant aux sous-unités Kir6.2, entraîne la fermeture du canal  $K^+_{ATP}$ , empêchant de ce fait la fuite d'ions  $K^+$  en extracellulaire et provoquant ainsi une dépolarisation de la cellule  $\beta$  pancréatique<sup>32</sup>. La dépolarisation induite par la fermeture des canaux  $K^+_{ATP}$  active à son tour les canaux calciques  $Ca^{2+}$  voltage-dépendants de type L<sup>32</sup>. Ces canaux permettent alors l'influx d'ions  $Ca^{2+}$  depuis

le milieu extracellulaire vers le cytosol, où ils participent à l'exocytose des granules d'insuline<sup>32,35</sup>.

#### *5.1.1.2.2 Transport des granules de sécrétion*

En amont de l'exocytose, les granules de sécrétion sont transportées du cytoplasme vers la périphérie de la cellule, proche de la membrane plasmique. Pour cela, diverses structures cellulaires participent au déplacement des granules, notamment le cytosquelette. Celui-ci est composé de microtubules ainsi que de microfilaments d'actine. L'interaction des microtubules du cytosquelette et des microfilaments d'actine avec les protéines motrices kinésine et myosine Va, respectivement, sur lesquelles les granules d'insuline viennent se lier, permet ainsi le transport de l'insuline jusqu'à la membrane plasmique<sup>41</sup>.

#### *5.1.1.2.3 Processus d'exocytose*

Lors d'une stimulation par le glucose, l'élévation de la concentration intracellulaire en ions  $Ca^{2+}$  dans la cellule  $\beta$  pancréatique induit l'exocytose de l'insuline. Ce processus est similaire à l'exocytose des vésicules contenant les neurotransmetteurs dans les jonctions synaptiques des connexions neuronales, et est rendu possible grâce à l'action des protéines de la famille soluble N-ethylmaleimide sensitive factor attachment protein receptor (SNARE)<sup>32</sup>. Ces protéines possèdent des domaines en hélice  $\alpha$ , nommés motifs SNARE, capables de former des superhélices en interagissant ensemble. Elles peuvent être divisées en protéines v-SNARE (pour vesicle), localisées sur les granules de sécrétion, et en protéines t-SNARE (pour target), localisées sur la membrane plasmique<sup>35</sup>. Dans la cellule  $\beta$  pancréatique, les protéines SNARE qui participent à l'exocytose de l'insuline sont la protéine v-SNARE synaptobrevine transmembranaire, possédant 2 motifs SNARE, et les protéines t-SNARE syntaxine-1/4 transmembranaire et synaptosomal nerve-associated protein 25 (SNAP-25) ancrée aux lipides, possédant chacune un seul motif SNARE<sup>35,42</sup>. Cependant, les protéines SNARE ne sont pas suffisantes à elles seules pour induire l'exocytose des granules de sécrétion. Elles nécessitent l'aide de protéines accessoires, comme les protéines mammalian uncoordinated-18-1 (MUNC18-1) et 13-1 (MUNC13-1), ou encore la synaptotagmine<sup>35,42</sup>.

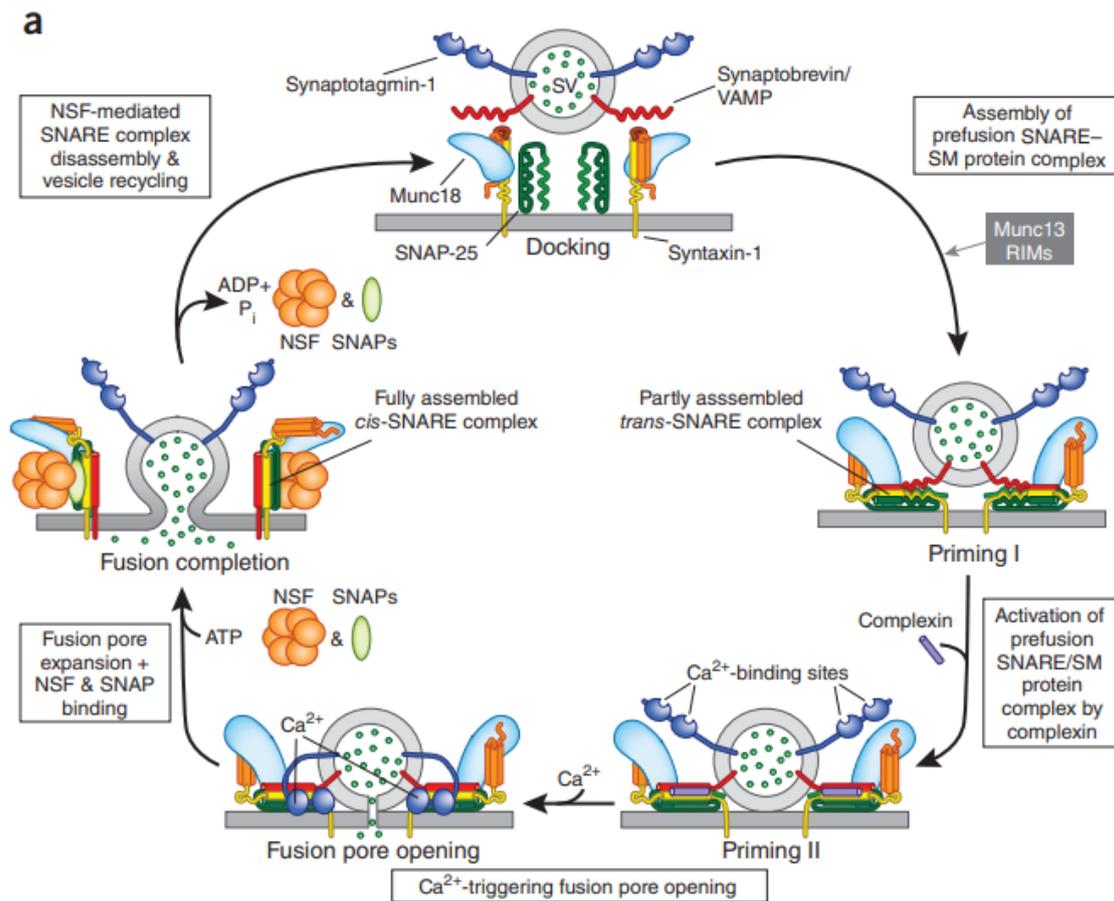


Figure 9. Processus d'exocytose. Source : Sudhof, « A molecular machine for neurotransmitter release: synaptotagmin and beyond », Nature Medicine, 2013.

Le processus d'exocytose expliqué ci-après est illustré par la figure 9. A l'étape d'arrimage (docking), la protéine MUNC18-1 est liée à la syntaxine, maintenant celle-ci en position fermée et empêchant ainsi la formation du complexe SNARE<sup>35,42</sup>. Cependant, lors de l'étape d'amorçage (priming), la protéine MUNC13-1, en interagissant avec MUNC18-1, va permettre le basculement de la syntaxine de l'état fermé vers l'état ouvert<sup>35</sup>. Une fois en conformation ouverte, la syntaxine se lie à la protéine SNAP-25 et à la synaptobrevine, formant ainsi un complexe superhélicoïdal composé des 4 hélices  $\alpha$  des protéines SNARE, appelé complexe trans-SNARE. Le resserrage de ce complexe SNARE génère alors de l'énergie qui permet de rapprocher la membrane de la granule et la membrane plasmique<sup>35</sup>. A cette étape, la fusion des membranes est inhibée par l'action de la complexine. Cette petite protéine agit comme un répresseur de la fusion membranaire, par insertion de son hélice  $\alpha$  antiparallèle dans le sillon entre les motifs SNARE de la syntaxine et de la synaptobrevine du complexe SNARE<sup>42,43</sup>. De cette façon, elle maintient le complexe SNARE dans un état métastable, empêchant donc une sécrétion spontanée<sup>35</sup>. Lors de l'influx massif d'ions  $Ca^{2+}$ , la

synaptotagmine, une protéine senseur au  $\text{Ca}^{2+}$ , est activée, et entre en compétition avec la complexine pour la liaison au complexe SNARE, tout en se liant également aux phospholipides membranaires, facilitant ainsi la formation du pore de fusion<sup>35,43</sup>. Après la fusion membranaire, le complexe SNARE se retrouve sur une seule membrane<sup>44</sup>. Il est désassemblé à l'aide de l'ATPase N-ethylmaleimide sensitive fusion (NSF) et de la protéine SNAP- $\alpha$ , permettant ainsi le recyclage des protéines SNARE pour les exocytoses futures<sup>42,45</sup>.

#### 5.1.1.3 Régulation de l'insulinosécrétion par le glucose

Il existe de nombreuses molécules qui sont capables d'induire la sécrétion d'insuline par les cellules  $\beta$  pancréatiques : ce sont des initiateurs de l'insulinosécrétion. Il s'agit principalement du glucose, mais également des acides aminés et des acides gras. Suite à un repas, ces nutriments sont absorbés par le tractus gastrointestinal, et l'élévation de leur concentration circulante va être captée par les cellules  $\beta$  pancréatiques. De l'insuline va être libérée pour stimuler leur capture par différents organes cibles, où ils seront métabolisés et stockés. La sécrétion d'insuline induite par ces initiateurs peut être régulée par d'autres molécules, qui peuvent être stimulatrices ou parfois inhibitrices de l'insulinosécrétion : ils sont appelés potentiateurs de l'insulinosécrétion, et sont composés principalement d'hormones et de neurotransmetteurs. Ces molécules n'ont pas, ou très peu d'effet sur l'insulinosécrétion en état de normoglycémie, empêchant ainsi l'apparition d'une hypoglycémie chez les individus<sup>32</sup>. Dans le cadre de cette thèse, nous nous intéresserons uniquement à la régulation de la sécrétion d'insuline par le glucose.

Le glucose est le principal sécrétagogue issu de l'alimentation, et son effet a déjà été partiellement décrit plus haut. En réalité, son effet sur l'insulinosécrétion est plus complexe, et implique différents mécanismes. Le glucose est capable d'activer deux voies différentes permettant de stimuler l'insulinosécrétion : la première est appelée voie d'initiation (triggering pathway), et la deuxième voie d'amplification (amplifying pathway)<sup>46</sup>. La voie d'amplification dépend de la voie d'initiation : elle ne peut avoir un effet qu'en présence d'une dépolarisation membranaire induisant un influx massif d'ions  $\text{Ca}^{2+}$  dans la cellule. Cette voie permet donc d'amplifier le stimulus induit par la voie d'initiation<sup>46</sup>.

#### 5.1.1.3.1 La voie d'initiation

La voie d'initiation de l'insulinosécrétion correspond à celle décrite précédemment dans la section 5.1.1.2.1 « Stimulation de la sécrétion de l'insuline », et illustrée à la figure 8 : une fois que le glucose est entré dans les cellules  $\beta$  pancréatiques, sa métabolisation entraîne la génération d'ATP, qui induit la fermeture des canaux  $K^+_{ATP}$ , entraînant alors une dépolarisation membranaire responsable de l'ouverture des canaux  $Ca^{2+}$ . Celle-ci permet l'influx d'ions  $Ca^{2+}$ , et finalement l'exocytose de l'insuline. Elle nécessite donc la présence d'une forte concentration extracellulaire en glucose pour pouvoir être activée, afin d'atteindre le seuil nécessaire pour l'inactivation des canaux  $K^+_{ATP}$  et de l'ouverture des canaux  $Ca^{2+}$  voltage-dépendants<sup>46</sup>. Une élévation du contenu cytosolique en  $Ca^{2+}$  est suffisante à elle seule pour induire la sécrétion d'insuline<sup>47</sup>.

#### 5.1.1.3.2 La voie d'amplification

La voie d'amplification a été initialement découverte en 1992, et semble indépendante de l'inactivation des canaux  $K^+_{ATP}$  : il a été observé que dans des cellules déjà dépolarisées par la fermeture de ces canaux, le glucose était tout de même capable d'augmenter l'insulinosécrétion<sup>48</sup>. Cette voie est activée par différentes molécules régulatrices, dont certaines sont générées au cours du métabolisme du glucose dans la mitochondrie (figure 10). En effet, la métabolisation du glucose par la glycolyse génère comme produit final du pyruvate, qui peut alors être métabolisé par le complexe pyruvate déshydrogénase pour générer de l'acétyl-Coenzyme A (acétyl-CoA), ou être métabolisé par la pyruvate carboxylase pour générer de l'oxaloacétate. Ces deux molécules, acétyl-CoA et oxaloacétate, peuvent alors prendre part au cycle de Krebs dans la mitochondrie, et ainsi favoriser la production d'ATP<sup>49</sup>.

#### 5.1.1.3.3 Rôle des ions $Ca^{2+}$

Les ions  $Ca^{2+}$  jouent un rôle important dans la stimulation de la sécrétion d'insuline par différents mécanismes, repris en partie à la figure 11. Tout d'abord, le glucose peut activer la phospholipase C (PLC) de manière  $Ca^{2+}$ -dépendante<sup>32</sup>. Cette enzyme génère, à partir du phosphatidylinositol-4,5-bisphosphate (PIP2) de l'inositol-1,4,5-trisphosphate (IP3) et du diacylglycérol (DAG), deux messagers secondaires importants. L'IP3, en agissant sur ses récepteurs situés sur la membrane du réticulum endoplasmique, induit la libération d'ions  $Ca^{2+}$  à partir de celui-ci, permettant l'augmentation du contenu intracellulaire en  $Ca^{2+}$  et ainsi l'amélioration de l'insulinosécrétion<sup>32</sup>. Le DAG, quant à lui, active la protéine kinase C (PKC)

sensible au DAG, qui est impliquée dans la sécrétion d'insuline stimulée par le glucose<sup>50</sup>. Les ions  $Ca^{2+}$  peuvent également activer certaines kinases dépendantes du  $Ca^{2+}$ , dont l'activation est impliquée dans la sécrétion d'insuline<sup>32</sup>. On peut notamment citer la calcium/calmoduline kinase 2 (CAMK2), impliquée dans l'élévation de la concentration intracellulaire en ions  $Ca^{2+}$  par phosphorylation des récepteurs à la ryanodine (RyR) situés à la surface du réticulum endoplasmique<sup>51</sup> et dans l'arrimage et l'amorçage des granules d'exocytose<sup>52</sup>.

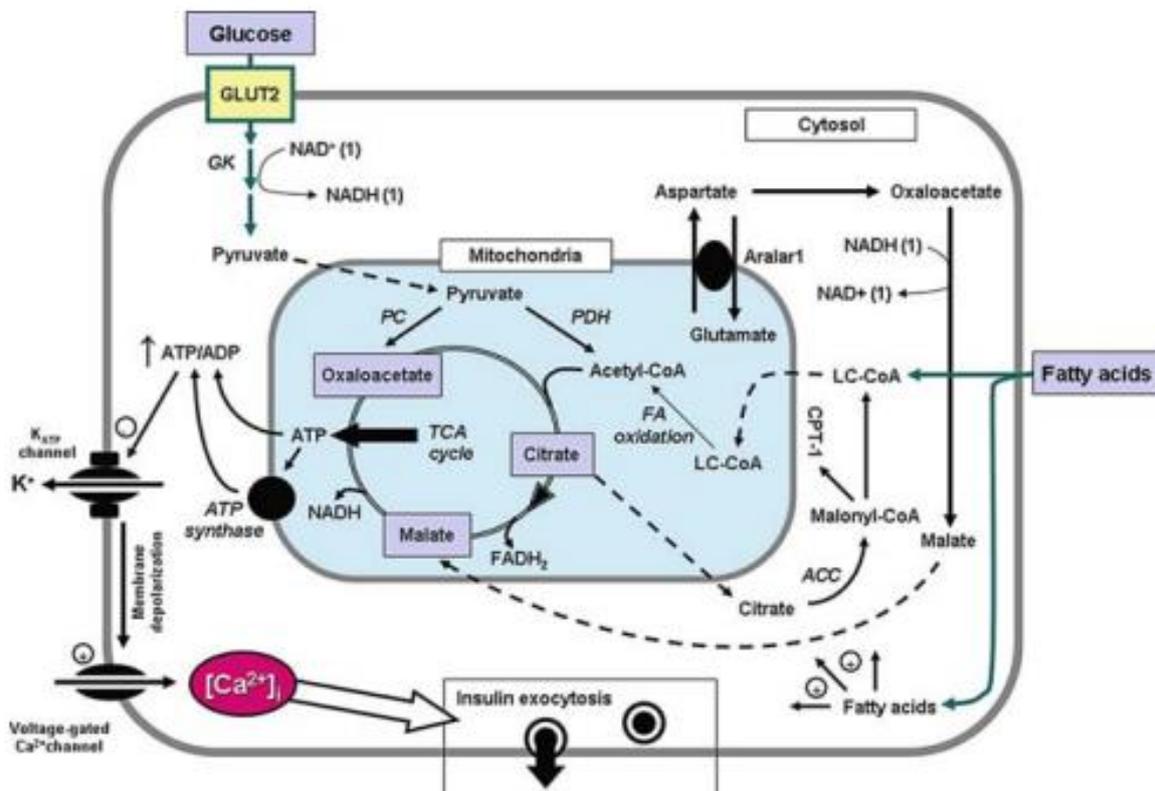


Figure 10. Représentation de la voie d'amplification de la sécrétion d'insuline par le glucose. **GK** : glucokinase, **PC** : pyruvate carboxylase, **PDH** : pyruvate déshydrogénase, **FA** : acides gras, **TCA** : acides tricarboxyliques, **LC-CoA** : acides gras à longue chaîne couplé au coenzyme A, **CPT-1** : carnitine palmitoyltransférase 1, **ACC** : acétyl-CoA carboxylase. Source : Newsholme et al., « Nutrient Regulation of Insulin Secretion and  $\beta$  Cell Functional Integrity », *Advances in Experimental Medicine and Biology*, 2010.

#### 5.1.1.3.4 Rôle de l'AMPC

L'adénosine monophosphate cyclique (AMPC) est une molécule dont la production est stimulée par le glucose, et qui est également impliquée dans la régulation de l'insulinosécrétion<sup>53</sup>, suivant les mécanismes repris à la figure 11. Cette élévation d'AMPC est provoquée par l'action de l'enzyme adénylate cyclase, dont l'activité est régulée par la concentration intracellulaire en  $Ca^{2+}$ <sup>53</sup>. L'AMPC est capable d'activer la protéine kinase A (PKA) ainsi que la protéine exchange protein directly activated by cAMP 2 (EPAC2), un facteur d'échange de guanine pour la GTPase Ras-related protein 1 RAP1<sup>53</sup>. Via ces 2 molécules,

l'AMPc est capable de réguler la dépolarisation membranaire ainsi que l'influx en ions  $\text{Ca}^{2+}$ . D'une part, la phosphorylation de la sous-unité SUR1 des canaux  $\text{K}^+_{\text{ATP}}$  par la PKA permet l'inactivation de ces canaux<sup>54</sup>, et la phosphorylation de la sous-unité Cav1.2 des canaux  $\text{Ca}^{2+}$  voltage-dépendants augmente leur activité, favorisant ainsi l'insulinosécrétion<sup>55</sup>. D'autre part, la protéine EPAC2 semble également inhiber l'activité des canaux  $\text{K}^+_{\text{ATP}}$  par son action sur la sous-unité SUR1<sup>56</sup>.

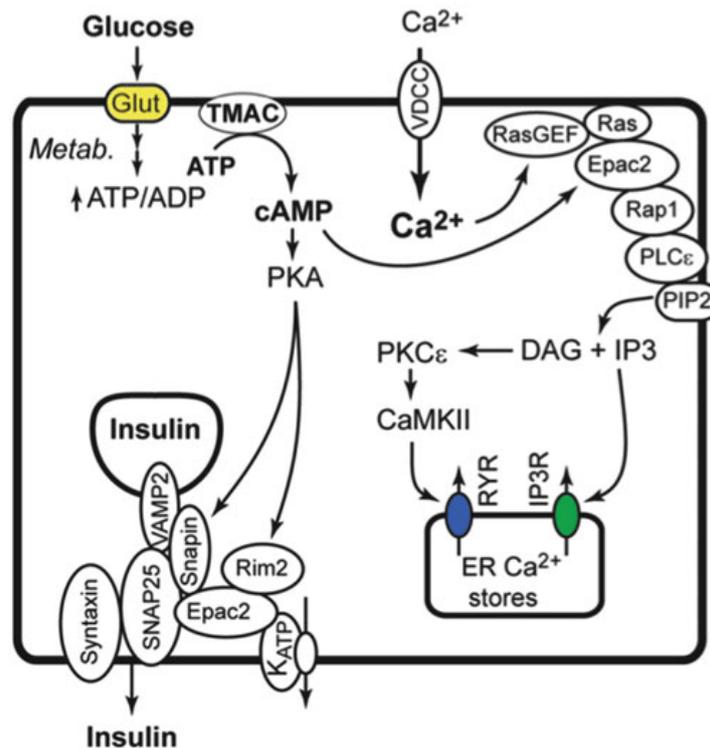


Figure 11. Mécanismes de l'insulinosécrétion médiée par les ions  $\text{Ca}^{2+}$  et par l'AMPc. **Glut** : transporteur de glucose, **TMAC** : adénylate cyclase transmembranaire, **ATP** : adénosine monophosphate, **ADP** : adénosine diphosphate, **cAMP** : adénosine monophosphate cyclique, **PKA** : protéine kinase A, **VDCC** : canal calcique voltage-dépendant, **RasGEF** : GTPase Ras, **PLCε** : phospholipase C ε, **PIP2** : phosphatidylinositol-4,5-bisphosphate, **IP3** : inositol-1,4,5-trisphosphate, **DAG** : diacylglycérol, **PKCε** : protéine kinase C ε, **CAMKII** : calcium/calmoduline kinase 2, **RYR** : récepteur à la ryanodine, **IP3R** : récepteur à l'IP3, **ER** : réticulum endoplasmique, **K<sub>ATP</sub>** : canal potassique ATP-dépendant, **SNAP25** : synaptosomal nerve-associated protein 25, **VAMP2** : vesicle-associated membrane protein. Source : Holz et al., « Molecular Basis of cAMP Signaling in Pancreatic Beta Cells », *Islets of Langerhans*, 2<sup>ème</sup> édition, 2015.

La voie de l'AMPc permet d'augmenter la concentration intracellulaire en ions  $\text{Ca}^{2+}$ . La protéine PKA est impliquée dans ce processus par le mécanisme de « calcium induced calcium release » : par activation des récepteurs à l'IP3 situés sur la membrane du réticulum endoplasmique, elle permet la libération du calcium à partir des citernes du réticulum endoplasmique vers le cytosol, participant ainsi à l'amplification de la réponse insulinaire<sup>57</sup>. La protéine EPAC2, quant à elle, active la PLCε en aval, également capable d'augmenter la

concentration intracellulaire en  $\text{Ca}^{2+}$  par activation des récepteurs à l'IP3 du réticulum endoplasmique<sup>58</sup>.

Finalement, l'AMPC peut également agir sur le processus d'exocytose, notamment en augmentant indirectement la sensibilité des granules d'insuline aux ions  $\text{Ca}^{2+}$  via la PKA<sup>59</sup>, et par liaison de la protéine EPAC2 avec la protéine SNAP-25 du complexe SNARE, facilitant l'exocytose<sup>60</sup>.

#### 5.1.1.4 Régulation autocrine et paracrine du pancréas endocrine

En dehors des nutriments, de nombreuses molécules, sécrétées par les différents types cellulaires composant les îlots pancréatiques, sont capables de moduler la sécrétion d'insuline en se liant à des récepteurs spécifiques de la cellule  $\beta$  pancréatique. Notamment, les diverses hormones sécrétées par ces différents types cellulaires sont capables de moduler la fonction des cellules voisines, de manière autocrine ou paracrine. Les interactions entre cellules sont reprises à la figure 12.

##### 5.1.1.4.1 Régulation autocrine et paracrine par les cellules $\beta$ pancréatiques

Les cellules  $\beta$  pancréatiques expriment à leur surface le récepteur à l'insuline (INSR), et contiennent la machinerie nécessaire à sa signalisation, permettant ainsi un contrôle feedback autocrine de l'insulinosécrétion<sup>61</sup>. L'action de l'insuline permet principalement de réguler l'expression de différents gènes nécessaires à la fonction des cellules  $\beta$  pancréatiques, comme celui de la préproinsuline ou de la GCK<sup>62</sup>. Ainsi, des études ont montré qu'une surexpression du récepteur à l'insuline dans une lignée murine de cellules  $\beta$  pancréatiques entraînait une élévation de l'ARNm de l'insuline et du contenu en insuline, tandis que l'expression d'un mutant non-fonctionnel du récepteur altérait l'insulinosécrétion<sup>63</sup>. D'autre part, l'action autocrine de l'insuline participe également au contrôle de la masse des cellules  $\beta$  pancréatiques via leur prolifération et leur apoptose<sup>61</sup>.

Hormis leur effet autocrine, les cellules  $\beta$  pancréatiques sont aussi capables d'agir sur les autres types cellulaires des îlots pancréatiques de manière paracrine. L'insuline sécrétée par ces cellules est capable d'inhiber la sécrétion de glucagon en agissant sur les cellules  $\alpha$  pancréatiques<sup>64</sup>. Cet effet est lié à l'activation du récepteur à l'insuline, étant donné que son knock-out spécifique aux cellules  $\alpha$  chez des souris augmente la sécrétion de glucagon<sup>65</sup>. Plus en détail, la liaison de l'insuline à son récepteur sur la cellule  $\alpha$  pancréatique et la signalisation

en aval inhibent la sécrétion de glucagon par différents mécanismes : par l'activation des canaux  $K^+_{ATP}$ <sup>66</sup>, induisant une hyperpolarisation membranaire, par l'activation de récepteurs à l'acide  $\gamma$ -aminobutyrique  $GABA_A$  constitués de canaux chloriques  $Cl^-$ <sup>67</sup>, dépolarisant ainsi la membrane, ou encore par la diminution de la concentration intracellulaire en  $AMP_c$ <sup>68</sup>, impliqué dans l'exocytose des granules de sécrétion.

L'insuline est également impliquée dans la régulation paracrine de la sécrétion de somatostatine par les cellules  $\delta$  pancréatiques. Plusieurs études sur des pancréas d'animaux, menées dans les années 1980, ont montré que l'insuline inhibe la sécrétion de somatostatine<sup>69,70</sup>, mais d'autres, plus récentes, ont montré que l'insuline n'avait pas d'effet sur sa sécrétion<sup>71</sup>. En réalité, l'insuline pourrait ne pas avoir d'effet direct sur celle-ci. D'autres molécules, sécrétées en même temps que l'insuline, pourraient être impliquées dans la régulation de la sécrétion de somatostatine<sup>72,73</sup>. D'autre part, les cellules  $\beta$  et  $\delta$  pancréatiques sont couplées par des jonctions communicantes : l'activité électrique des cellules  $\beta$  pancréatiques pourrait donc se propager aux cellules  $\delta$  pancréatiques et ainsi les stimuler, celles-ci sécrétant alors de la somatostatine afin d'inhiber la sécrétion de glucagon<sup>74</sup>.

#### 5.1.1.4.2 Régulation paracrine par les cellules $\alpha$ pancréatiques

Les cellules  $\alpha$  pancréatiques sécrètent du glucagon, une hormone hyperglycémiant sécrétée en réponse à une hypoglycémie<sup>32</sup>. Cette hormone permet notamment la glycogénolyse et la néoglucogenèse dans le foie afin d'augmenter la concentration plasmatique en glucose<sup>75</sup>. Le glucagon, malgré son effet antagoniste à l'insuline, est capable d'en stimuler la sécrétion<sup>32</sup>. L'action paracrine du glucagon sur les cellules  $\beta$  pancréatiques, par la liaison de celui-ci à son récepteur GCGR, active la protéine G stimulatrice ( $G_s$ ) couplée à l'adénylate cyclase. Son action entraîne alors une augmentation de la concentration intracellulaire en  $AMP_c$ , stimulant l'insulinosécrétion<sup>32</sup>. D'autre part, le glucagon est également capable de se lier et d'activer le récepteur au glucagon-like peptide 1 (GLP-1)  $GLP1R$  des cellules  $\beta$  pancréatiques<sup>76</sup>, un récepteur couplé aux protéines G (GPCR) de type  $G_s$ , avec une plus faible affinité que pour GCGR<sup>77</sup>. Des études menées sur des pancréas de souris ont démontré que le knock-out des récepteurs GCGR et  $GLP1R$  altérait la sécrétion d'insuline médiée par le glucagon<sup>78</sup>.

En outre, le glucagon peut également réguler la sécrétion de somatostatine par les cellules  $\delta$  pancréatiques de manière paracrine. Ces cellules expriment les récepteurs GCGR et

GLP1R, et la liaison du glucagon sur ceux-ci augmente la sécrétion de somatostatine. La sécrétion de somatostatine qui en résulte est inhibitrice de la sécrétion de glucagon, et constitue ainsi une boucle feedback négative<sup>79</sup>.

Le glucagon peut également agir directement sur les cellules  $\alpha$  pancréatiques, et ainsi réguler sa propre sécrétion de manière autocrine. L'action du glucagon sur les récepteurs GCGR et GLP1R à la surface de la cellule  $\alpha$  pancréatique active une adénylate cyclase, induisant l'élévation du contenu cytosolique en AMP<sub>c</sub>, et ainsi une diminution de sa propre sécrétion<sup>75</sup>.

#### 5.1.1.4.3 Régulation paracrine par les cellules $\delta$ pancréatiques

La somatostatine est une hormone sécrétée par les cellules  $\delta$  pancréatiques, mais également par les cellules de l'hypothalamus et du tractus gastrointestinal. Cette hormone joue un rôle inhibiteur sur les sécrétions endocrines et exocrines du pancréas<sup>32</sup>. Elle est sécrétée en réponse à une élévation de la concentration extracellulaire en glucose<sup>80</sup>, mais aussi en réponse aux acides aminés leucine et arginine<sup>81,82</sup>. L'isoforme de la somatostatine SST-14 est l'isoforme principale sécrétée par les cellules  $\delta$  pancréatiques, et a un effet inhibiteur sur la sécrétion de glucagon et d'insuline<sup>83</sup>. Dans la cellule  $\beta$  pancréatique, l'activation des récepteurs à la somatostatine SSTR couplés à une protéine G inhibitrice (Gi), principalement SSTR2, entraîne une hyperpolarisation des cellules via une activation des canaux G-protein inward rectifying K<sup>+</sup> channel (GIRK), une inhibition de l'influx de calcium par inhibition des canaux Ca<sup>2+</sup> de type P/Q, ainsi qu'une inhibition directe de l'exocytose<sup>84</sup>. Dans la cellule  $\alpha$  pancréatique, l'activation du récepteur SSTR2 entraîne une réduction de l'influx d'ions Ca<sup>2+</sup>, et une réduction de l'exocytose des granules de sécrétion du glucagon<sup>84</sup>. D'autre part, la régulation de la sécrétion de glucagon par la somatostatine peut être dépendante de l'activité des cellules  $\beta$  pancréatiques, comme décrit précédemment : les cellules  $\beta$  pancréatiques peuvent stimuler les cellules  $\delta$  pancréatiques soit par sécrétion d'insuline, soit par couplage électrique via les jonctions communicantes, afin de stimuler la sécrétion de somatostatine, et ainsi inhiber indirectement la sécrétion de glucagon<sup>74</sup>.

#### 5.1.1.4.4 Régulation paracrine par les cellules PP pancréatiques

Les cellules PP-pancréatiques sécrètent le polypeptide pancréatique (PP), qui est principalement connu pour son rôle dans la régulation de l'homéostasie énergétique, notamment en réduisant la prise alimentaire et en induisant une sensation de satiété<sup>85</sup>. Cependant, son rôle dans la régulation de la glycémie est mal compris. Certaines études ont

montré que sa sécrétion est stimulée par une concentration extracellulaire élevée en glucose, permettant ainsi l'activation de son récepteur PPYR1 situé sur les cellules  $\alpha$  pancréatiques, suivie de l'inhibition de la sécrétion de glucagon<sup>86</sup>. D'autres études ont supposé que ce peptide inhibe l'insulinosécrétion stimulée par le glucose par les cellules  $\beta$  pancréatiques murines et humaines<sup>87</sup>.

#### 5.1.1.4.5 Régulation paracrine par les cellules $\epsilon$ pancréatiques

La ghreline, hormone sécrétée par les cellules  $\epsilon$  pancréatiques, est considérée comme l'hormone de la faim. Elle joue un rôle inhibiteur sur la sécrétion d'insuline et de glucagon, de manière somatostatine-dépendante. En effet, sa liaison à son récepteur à la surface des cellules  $\delta$  pancréatiques entraîne une élévation de la concentration intracellulaire en ions  $\text{Ca}^{2+}$ , stimulant alors la sécrétion de somatostatine inhibitrice de la sécrétion d'insuline et de glucagon<sup>88</sup>.

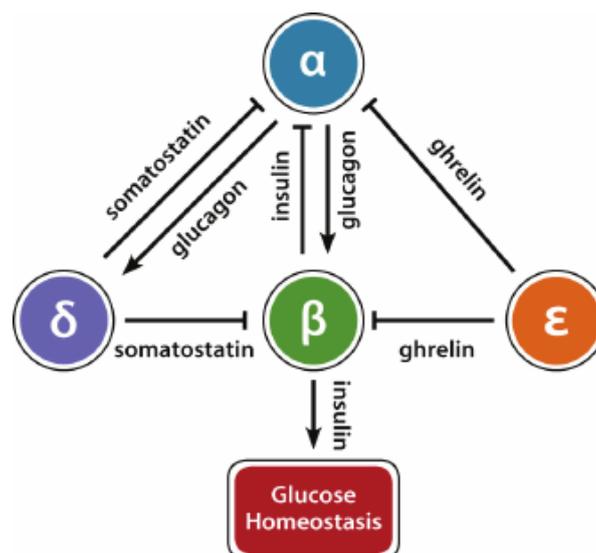


Figure 12. Interaction entre les différents types cellulaires des îlots pancréatiques. Source : Jang and Morahan, « Pancreatic Stem Cells : Unresolved Business », *Stem Cells in Clinic and Research*, 2011.

#### 5.1.1.5 Régulation de l'insulinosécrétion par les hormones incrétines

Les incrétines sont des hormones sécrétées par les cellules du tractus gastrointestinal suite au passage de nutriments dans l'intestin. Ce groupe d'hormones comprend le GLP-1 ainsi que le glucose-dependent insulintropic polypeptide (GIP)<sup>32</sup>. Une fois sécrétées, ces hormones sont transportées dans la circulation sanguine pour rejoindre les îlots pancréatiques, où elles amplifient la sécrétion d'insuline stimulée par le glucose. Ceci permet d'expliquer l'effet plus important d'une administration orale de glucose par rapport à une administration intraveineuse<sup>32</sup>. Leur effet sur la sécrétion d'insuline est repris à la figure 13.

#### 5.1.1.5.1 GLP-1

La principale hormone incrétine est le GLP-1. Il est sécrété par les cellules entéroendocrines de type L, situées dans la partie distale de l'intestin grêle et du côlon, en réponse au passage des nutriments<sup>89</sup>. Tout comme le glucagon, cette hormone est issue du clivage du proglucagon<sup>32</sup>. Le GLP-1 est impliqué dans différents processus biologiques, à savoir la stimulation de l'insulinosécrétion, l'inhibition de la sécrétion du glucagon, le ralentissement de la vidange gastrique, ou encore la diminution de la prise de nourriture<sup>32</sup>. Il est présent sous différentes isoformes dans l'organisme, mais seuls le GLP-1 (7-36) amide et le GLP-1 (7-37) sont capables de stimuler l'insulinosécrétion en réponse au glucose<sup>32</sup>. La liaison du GLP-1 à son récepteur GLP1R à la surface des cellules  $\beta$  pancréatiques permet l'activation de l'adénylate cyclase. Cette enzyme, en augmentant la concentration en AMP<sub>c</sub> cytosolique, active les protéine PKA et EPAC2 qui induisent l'activation des canaux Ca<sup>2+</sup> dépendants du voltage et des récepteurs à la ryanodine situés sur le réticulum endoplasmique, mais également la fermeture des canaux K<sup>+</sup><sub>ATP</sub>, favorisant ainsi l'influx d'ions Ca<sup>2+</sup> dans le cytoplasme<sup>90</sup>. D'autre part, cette hormone inhibe également la sécrétion de glucagon : sa liaison au récepteur GLP1R à la surface des cellules  $\alpha$  pancréatiques induit l'activation de l'adénylate cyclase et l'élévation de la concentration en AMP<sub>c</sub> en aval, entraînant de ce fait une inhibition des canaux Ca<sup>2+</sup> de type P/Q par la PKA, empêchant ainsi l'exocytose du glucagon<sup>91,92</sup>. Mais l'effet inhibiteur du GLP-1 sur la sécrétion de glucagon pourrait également passer par une activation du récepteur GCGR par celui-ci<sup>92</sup>.

#### 5.1.1.5.2 GIP

Le GIP est une incrétine sécrétée par les cellules entéroendocrines de type K, situées dans le duodénum et la partie proximale de l'intestin grêle en réponse au passage des nutriments<sup>32</sup>. Cette incrétine a divers effets dans l'organisme, à savoir un effet inhibiteur sur la sécrétion d'acide dans l'estomac, un effet stimulateur sur la sécrétion d'insuline en réponse au glucose, mais également un effet stimulateur sur la sécrétion de glucagon<sup>32</sup>. Tout comme le GLP1R, le récepteur au GIP GIPR est un GPCR couplé à une protéine G<sub>s</sub>, dont l'activation stimule l'insulinosécrétion<sup>93,94</sup>. Ceci a lieu via une élévation de la concentration cytosolique en AMP<sub>c</sub>, ainsi que l'inhibition des canaux K<sup>+</sup><sub>ATP</sub> facilitant ainsi l'exocytose des granules d'insuline<sup>95</sup>, mais également via l'activation de la PLA<sub>2</sub>, qui génère de l'acide arachidonique stimulateur de l'insulinosécrétion<sup>96</sup>. D'autre part, le GIP stimule aussi la sécrétion de glucagon,

également via la voie de l'AMPc/PKA<sup>97</sup>. L'effet du GIP sur la sécrétion de glucagon et d'insuline semble dépendre de l'état métabolique des cellules  $\alpha$  et  $\beta$  pancréatiques : ainsi, en présence d'une concentration élevée en glucose, les cellules  $\beta$  pancréatiques sont déjà dans un état stimulé, ce qui favorise l'action potentialisatrice du GIP sur la sécrétion d'insuline. Au contraire, en présence d'une faible concentration en glucose, ce sont les cellules  $\alpha$  pancréatiques qui sont stimulées, et l'effet potentialisateur du GIP sur la sécrétion de glucagon est privilégié<sup>97</sup>. Finalement, le GIP est également capable d'inhiber la libération de GLP-1<sup>98</sup>.

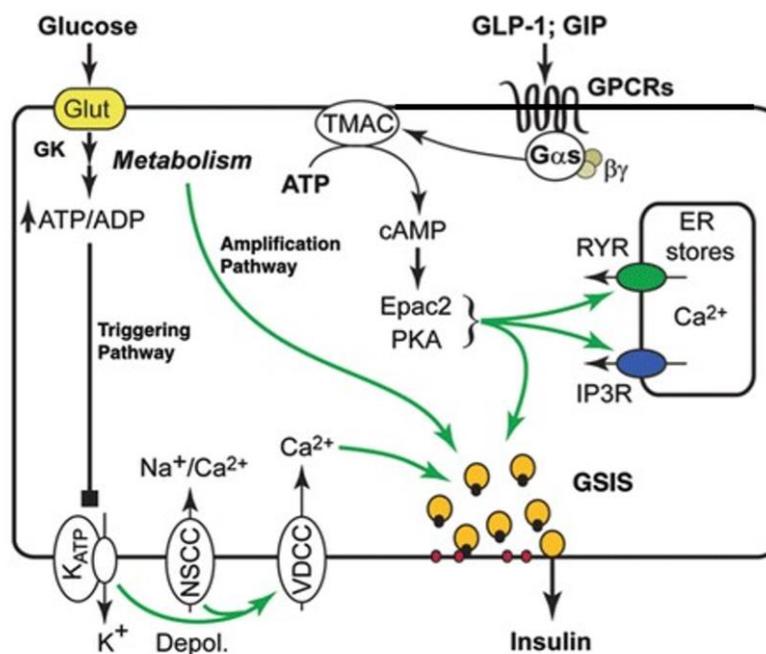


Figure 13. Effets des sécrétagogues GLP-1 et GIP sur la sécrétion d'insuline. **Glut** : transporteur de glucose, **GK** : glucokinase, **K<sub>ATP</sub>** : canal potassique ATP-dépendant, **NSCC** : canal cationique non-sélectif, **VDCC** : canal calcique voltage-dépendant, **TMAC** : adénylate cyclase transmembranaire, **RYR** : récepteur à la ryanodine, **IP3R** : récepteur à l'IP3, **ER** : réticulum endoplasmique, **GSIS** : sécrétion d'insuline stimulée par le glucose. Source : Adapté de Holz et al., « Molecular Basis of cAMP Signaling in Pancreatic Beta Cells », *Islets of Langerhans*, 2ème édition, 2016.

## 5.2 Action de l'insuline sur les tissus cibles périphériques

En temps normal, lors d'une élévation de la concentration plasmatique en glucose, celui-ci permet de stimuler la sécrétion d'insuline par les cellules  $\beta$  pancréatiques via les mécanismes décrits précédemment. Cette insuline est alors transportée dans la circulation sanguine et exerce son action sur les différents tissus cibles périphériques. Concrètement, l'action de l'insuline sur le muscle squelettique favorise la capture du glucose par celui-ci, ainsi que son stockage sous forme de glycogène ; son action sur le tissu adipeux augmente la

capture du glucose, et améliore le métabolisme lipidique ; enfin, dans le foie, l'insuline inhibe la production endogène de glucose<sup>99</sup>.

Cependant, dans le cadre du DT2, ces organes présentent généralement une insulino-résistance. Ceci, couplé à l'altération de la sécrétion d'insuline par les cellules  $\beta$  pancréatiques, entraîne une altération de la signalisation de l'insuline dans ces tissus cibles, impactant de ce fait la régulation du métabolisme énergétique. Notamment, l'insulino-résistance adipocytaire est associée à une augmentation de l'inflammation, une diminution de la sécrétion de l'adipokine adiponectine, et une accumulation de graisse dans le tissu adipeux ; l'insulino-résistance du muscle squelettique est associée à une diminution de la capture de glucose et à une réduction de la synthèse de glycogène ; enfin, l'insulino-résistance hépatique est associée à une production anormalement importante de glucose dans le foie, et à une altération des niveaux de lipides sériques<sup>28</sup>. La signalisation du récepteur à l'insuline, ainsi que les différents mécanismes moléculaires associant son action au métabolisme des tissus cibles sont repris ci-dessous.

#### 5.2.1 Activation du récepteur à l'insuline

Le récepteur à l'insuline INSR est un récepteur à activité tyrosine kinase. Il est composé de 2 sous-unités  $\alpha$  extracellulaires reliées entre elles par un pont disulfure, sur lesquelles vient se lier l'insuline, et de 2 sous-unités  $\beta$  transmembranaires portant les domaines à activité tyrosine kinase<sup>99</sup>. Chaque sous-unité  $\alpha$  est liée à une sous-unité  $\beta$  par un pont disulfure (figure 14).

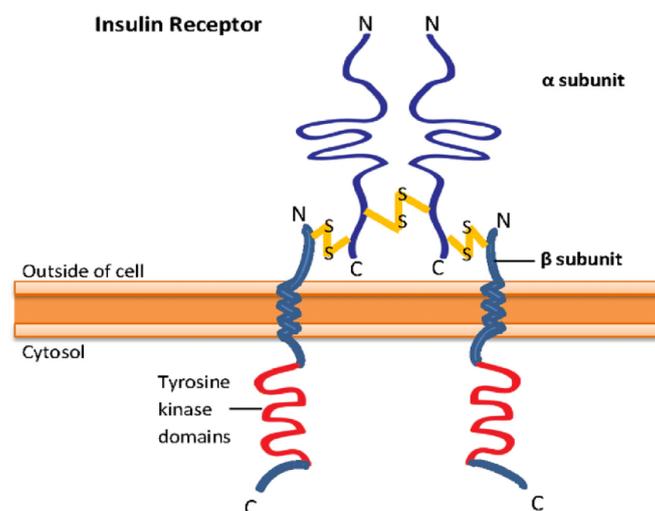


Figure 14. Structure du récepteur à l'insuline. **N** : extrémités N-terminales ; **C** : extrémités C-terminales ; **SS** : ponts disulfure. Source : Mangmool et al., «  $\beta$  Adrenergic Receptor and Insulin Resistance in the Heart », *Biomolecules and Therapeutics*, 2017.

La liaison de l'insuline sur son récepteur induit un changement de conformation de celui-ci, permettant l'autophosphorylation de plusieurs résidus tyrosine situés dans les sous-unités  $\beta$ . Ceci permet d'activer les domaines tyrosine kinase de ces sous-unités, qui pourront alors phosphoryler d'autres résidus tyrosine du récepteur, lui permettant de recruter plusieurs substrats<sup>99</sup>. Notamment, la protéine adaptatrice SRC homology 2 domain containing adaptor protein (SHC), ainsi que la protéine insuline receptor substrate (IRS) sont capables de se lier sur la Tyr972 phosphorylée, dans le domaine juxtramembranaire du récepteur, grâce à leur domaine phosphotyrosine binding (PTB). D'autres, telles que les protéines Src homology 2 b adaptor protein 1 (SH2B1) et SH2B2/APS, ou les régulateurs négatifs de la signalisation de l'insuline growth-factor receptor bound protein 10 (GRB10) et GRB14, interagissent via leur domaine Src homology 2 (SH2) avec la boucle d'activation du récepteur<sup>100</sup>. Ces protéines peuvent alors activer plusieurs voies de signalisation en aval (figure 15).

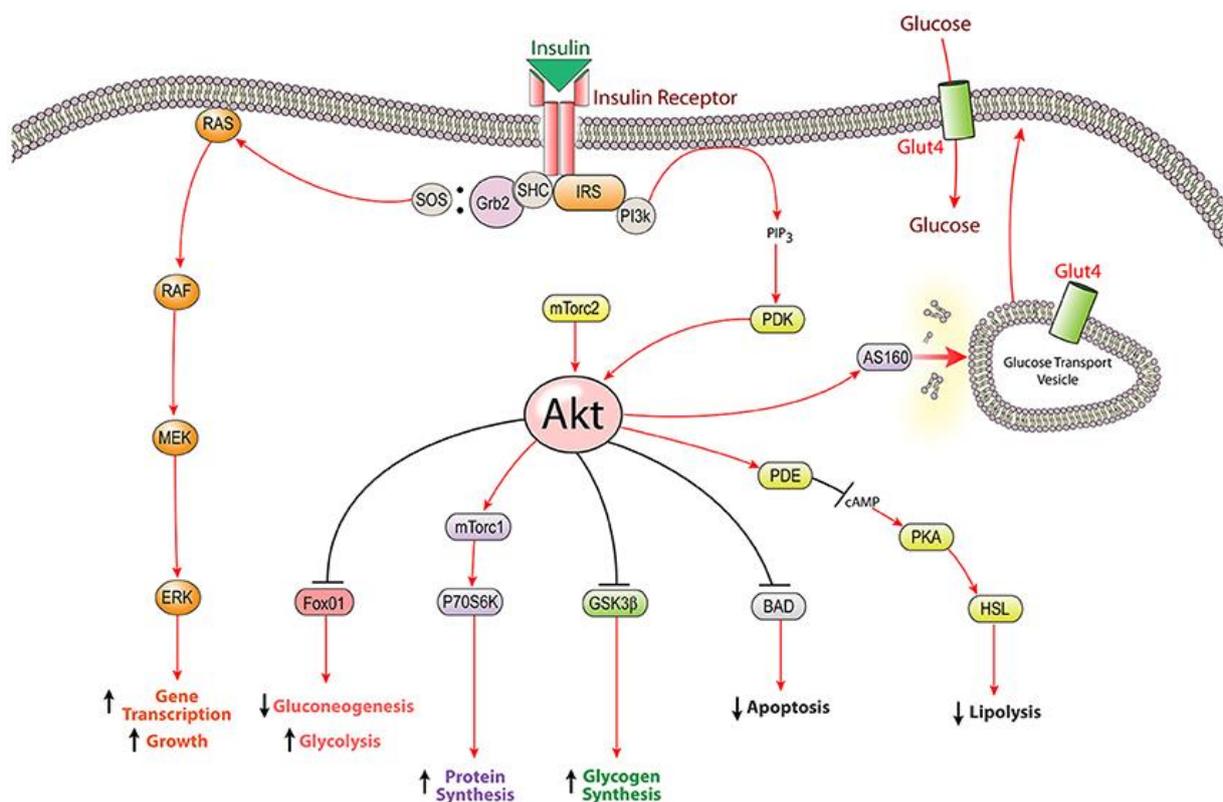


Figure 15. Voie de signalisation de l'insuline. **IRS** : insulin receptor substrate; **SHC** : SRC homology 2 domain containing adaptor protein; **Grb2** : growth-factor receptor bound protein 2; **SOS** : son of sevenless; **RAS** : rat sarcoma virus; **RAF** : rapidly accelerated fibrosarcoma; **MEK** : mitogen-activated protein kinase kinase; **ERK** : extracellular signal-regulated kinase; **PI3K** : phosphoinositide 3-kinase; **Glut4** : transporteur de glucose 4; **PIP<sub>3</sub>** : phosphatidylinositol-3,4,5-trisphosphate; **PDK** : phosphoinositide-dependent kinase; **mTorc2** : mammalian target of rapamycin complex 2; **Akt** : protéine kinase B; **AS160** : TRE-2/USP6, BUB2, cdc16 domain family member 1; **FoxO1** : forkhead box protein O1; **mTorc1** : mammalian target of rapamycin complex 1; **P70S6K** : p 70 ribosomal S6 kinase; **GSK3 $\beta$**  : glycogène synthase kinase 3  $\beta$ ; **BAD** : B-cell lymphoma 2-associated agonist of cell death; **PDE** : phosphodiesterase; **cAMP** : adenosine monophosphate cyclique; **PKA** : protéine kinase A; **HSL** : lipase hormonosensible. Source : Grote and Wright, « A role for Insulin in Diabetic Neuropathy », *Frontiers in Neuroscience*, 2016.

La signalisation en aval du récepteur de l'insuline permet notamment d'activer deux voies de signalisation importantes : la voie phosphoinositide 3-kinase (PI3K)/ protéine kinase B (PKB/AKT), impliquée dans les effets métaboliques de l'insuline, et la voie Ras sarcoma virus (RAS)/mitogen-activated protein kinase (MAPK), impliquée dans le contrôle de l'expression génique ainsi que dans la croissance et la différenciation cellulaire<sup>101</sup>. L'activation de ces voies de signalisation induite par l'insuline est reprise à la figure 15.

#### 5.2.1.1 Activation de la voie PI3K/AKT

La première protéine à être impliquée dans la transduction du signal du récepteur INSR pour activer la voie de signalisation PI3K/AKT est la protéine IRS. Il existe plusieurs isoformes de cette protéine, allant de IRS1 à 6. Cependant les protéines IRS1 et IRS2 sont considérées comme celles étant majoritairement impliquées dans les effets métaboliques de l'activation du récepteur INSR : IRS1 agit surtout dans le muscle squelettique, tandis que IRS2 est majoritairement impliquée dans la signalisation de l'insuline dans le foie et la cellule  $\beta$  pancréatique<sup>99,102</sup>. Les domaines PTB et pleckstrin homology (PH) d'IRS lui permettent de se lier au récepteur INSR, qui phosphoryle alors plusieurs résidus tyrosine dans la queue C-terminale de ce substrat. IRS peut alors servir de protéine d'amarrage pour d'autres protéines, capables de venir se lier sur les résidus tyrosine phosphorylés grâce à leurs domaines SH2.<sup>101</sup> Elle sert notamment de protéine d'amarrage pour la sous-unité régulatrice p85 de la protéine PI3K, qui inhibe alors la sous-unité catalytique p110 de PI3K<sup>103</sup>.

La protéine PI3K a pour rôle de phosphoryler le PIP2 membranaire en phosphatidylinositol-3,4,5-trisphosphate (PIP3)<sup>99,102</sup>. Ce PIP3 peut alors à son tour recruter AKT à la membrane, et peut également interagir avec l'enzyme phosphoinositide-dependent kinase-1 (PDK-1). AKT, ainsi que l'enzyme PKC $\zeta$ , peuvent alors être phosphorylées et ainsi activées par la protéine PDK<sup>102</sup>. La protéine AKT est une sérine-thréonine kinase, qui existe sous 3 isoformes différentes : AKT1, 2, et 3. Cependant, les isoformes principalement exprimées dans les tissus métaboliques tels que foie et le tissu adipeux sont les isoformes AKT1 et AKT2<sup>102</sup>. Cette protéine peut alors jouer différents rôles dans la cascade de signalisation de l'insuline, dépendant notamment des tissus étudiés<sup>104</sup> et repris dans la section 5.2.2 « Effets de l'insuline dans les tissus ».

### 5.2.1.2 Activation de la voie RAS/MAPK

La voie de signalisation RAS/MAPK implique également la protéine IRS. Cette protéine, tout comme la protéine SHC, se lie d'abord au récepteur à l'insuline. La protéine adaptatrice GRB2 peut ensuite venir s'y lier grâce à son domaine SH2. Cette protéine sert ensuite de site d'amarrage pour d'autres protéines, notamment la protéine Son of sevenless (SOS), un facteur d'échange GTP/GDP. SOS peut ainsi activer la petite GTPase membranaire RAS, qui active à son tour la protéine kinase rapidly accelerated fibrosarcoma (RAF). S'ensuit alors une cascade de phosphorylation : RAF phosphoryle et active les protéines mitogen-activated protein kinase kinase (MEK) 1 et 2, qui phosphorylent à leur tour les protéines extracellular signal-regulated kinase (ERK) 1 et 2. Ces kinases peuvent alors être transloquées dans le noyau, où elles vont réguler l'expression de différents gènes de plusieurs façons : d'une part, par phosphorylation de la p90 ribosomal S6 kinase (P90RSK), dont l'action va permettre d'activer des facteurs de transcription tels que FBJ murine osteosarcoma viral oncogene homolog (C-FOS), mais aussi par phosphorylation directe de facteurs de transcription tels que ETS like-1 (ELK-1)<sup>101,102</sup>.

### 5.2.2 Effets de l'insuline dans les tissus

#### 5.2.2.1 Effets de l'insuline dans le muscle squelettique

De par son rôle, le muscle est un tissu consommant énormément d'énergie. Pour cela, il est nécessaire que les myocytes puissent utiliser le glucose environnant. L'effet de l'insuline sur les cellules musculaires squelettiques consiste principalement à promouvoir la capture du glucose, mais également à le stocker dans le muscle sous forme de glycogène<sup>99</sup>, via les mécanismes décrits ci-dessous et repris en partie à la figure 16.

Lorsque les myocytes ne sont pas stimulés par l'insuline, le principal transporteur de glucose exprimé à la surface des cellules est GLUT1, qui permet la diffusion du glucose depuis la circulation sanguine vers le cytoplasme. GLUT4, quant à lui, est également exprimé dans les cellules, mais reste principalement contenu au sein du réseau trans-Golgien, d'endosomes, et de vésicules de stockage<sup>102,105</sup>. Lors d'une élévation de la concentration en glucose dans le corps, l'action de l'insuline sur son récepteur à la surface des myocytes induit le transport de GLUT4 à la membrane<sup>102</sup>. Cette action a lieu suite à l'activation de la voie de signalisation PI3K/AKT : en conditions basales, la petite GTPase Ras-related protein RAB, impliquée dans le transport de vésicules contenant GLUT4, est inhibée par les protéines TRE-2/BUB2/CDC16

domain family member 1 (TBC1D1) et 4 (TBC1D4/AS160). L'activation de la protéine AKT suite à la liaison de l'insuline à son récepteur phosphoryle ces protéines afin de les inactiver, levant ainsi l'inhibition de RAB et permettant le transport des vésicules contenant GLUT4 à la membrane<sup>102</sup>. En outre, la contraction musculaire permettrait également d'induire la translocation de GLUT4 à la surface des myocytes, indépendamment de l'action de l'insuline sur le muscle squelettique<sup>106</sup>. Cette action passerait par l'élévation de la concentration intracellulaire en 5'AMP dans les cellules, et l'activation de la protéine AMP-activated protein kinase (AMPK) en aval<sup>106,107</sup>.

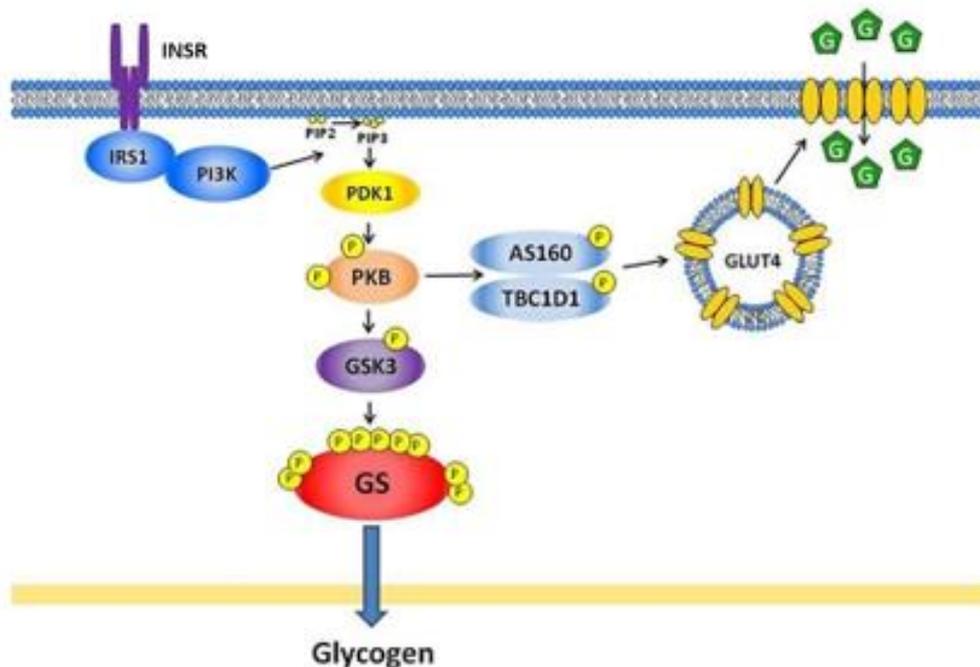


Figure 16. Actions de l'insuline dans les cellules du muscle squelettique. **IRS1** : insulin receptor substrate 1, **PIP2** : phosphatidylinositol 4,5-bisphosphate, **PIP3** : phosphatidylinositol 3,4,5-trisphosphate; **PDK1** : phosphoinositide-dependent kinase 1, **PKB** : protéine kinase B, **GSK3** : glycogène synthase kinase 3, **GS** : glycogène synthase, **TBC1D1** : TRE-2/BUB2/CDC16 domain family member 1, **AS160** : TRE-2/BUB2/CDC16 domain family member 4, **GLUT4** : transporteur de glucose 4, **G** : glucose. Source : adapté de Jensen et al., « The Role of Skeletal Muscle Glycogen Breakdown for Regulation of Insulin Sensitivity by Exercise », *Frontiers in Physiology*, 2011.

D'autre part, l'action de l'insuline à la surface des myocytes permet également de favoriser la glycolyse ainsi que la glycogénogenèse. En effet, l'insuline régule la transcription de l'enzyme hexokinase 2, dont le rôle est de phosphoryler le glucose pénétrant les cellules musculaires en glucose-6-phosphate, afin qu'il puisse s'engager dans la voie de la glycolyse<sup>108</sup>. De plus, l'insuline entraîne la déphosphorylation de la glycogène synthase ainsi que de la glycogène phosphorylase via la voie de signalisation AKT/PKB. Concrètement, l'activation de la protéine AKT en aval de la liaison de l'insuline à son récepteur permet la phosphorylation,

et ainsi l'inhibition, de la glycogène synthase kinase 3 (GSK3). L'inhibition de la glycogène synthase par la GSK3 est donc levée, et le glucose peut ainsi être stocké sous forme de glycogène<sup>109</sup>. En parallèle, l'insuline active la protéine phosphatase 1 (PP1), impliquée dans la déphosphorylation de la glycogène synthase<sup>110</sup>. Ces actions conjointes permettent l'activation de la glycogène synthase, et ainsi la synthèse de glycogène<sup>111</sup>. En outre, l'insuline promeut l'inactivation de la phosphorylase kinase par déphosphorylation de celle-ci, empêchant ainsi l'activation de la glycogène phosphorylase impliquée dans la glycogénolyse<sup>112</sup>. De plus, l'enzyme PP1, activée par l'insuline, participe également à la déphosphorylation de la glycogène phosphorylase<sup>113</sup>. Par ces mécanismes, l'insuline permet d'inhiber la glycogénolyse et favorise ainsi la synthèse de glycogène<sup>99</sup>.

#### 5.2.2.2 Effets de l'insuline dans le tissu adipeux

Le tissu adipeux sert principalement de réserve énergétique pour l'organisme via le stockage de triglycérides. L'action de l'insuline sur son récepteur, à la surface des adipocytes, joue un rôle important dans le métabolisme du tissu adipeux : en plus de participer au transport de GLUT4 à la surface des adipocytes<sup>99</sup>, l'insuline est également impliquée dans le métabolisme lipidique<sup>102</sup>, comme repris partiellement à la figure 17.

Tout d'abord, l'insuline, par différentes actions, est capable de réguler le processus de lipogénèse. Tout comme dans le muscle, l'activation de la protéine AKT, en aval du récepteur à l'insuline, induit dans les adipocytes l'inhibition de la protéine GSK3 par phosphorylation de celle-ci. En plus de favoriser le stockage de glucose sous forme de glycogène comme décrit précédemment pour le muscle, ceci empêche également l'enzyme GSK3 d'inhiber l'enzyme ATP citrate lyase, et favorise ainsi la conversion du citrate en acétyl-CoA dans le cytoplasme, permettant ainsi la synthèse d'acides gras à partir de celui-ci. De plus, l'insuline favorise la transcription de deux protéines importantes dans le processus de lipogénèse : il s'agit de la fatty acid synthase (FAS), une enzyme permettant la conversion d'acétyl-CoA et de malonyl-CoA en acides gras à longue chaîne, ainsi que de la sterol regulatory element binding protein 1 (SREBP-1), un facteur de transcription impliqué dans régulation de l'expression de gènes nécessaires à la capture et la synthèse des acides gras et des triglycérides<sup>102</sup>.

D'autre part, l'insuline exerce également une fonction anti-lipolytique. En période de jeun, la lipolyse est stimulée afin de libérer des acides gras libres, afin que le foie puisse les utiliser pour générer des corps cétoniques, source d'énergie que le système nerveux pourra

consommer à la place du glucose. Cette lipolyse est favorisée par la PKA de deux manières. D'une part, la PKA phosphoryle la lipase hormonosensible (HSL), impliquée dans la dégradation des triglycérides en glycérol et en acides gras, permettant ainsi sa translocation du cytoplasme vers la surface des gouttelettes lipidiques. D'autre part, elle phosphoryle également les périlipines. Ces protéines, situées à la surface des gouttelettes lipidiques des adipocytes, forment une barrière protectrice empêchant la dégradation des triglycérides par la HSL. Leur phosphorylation par la PKA permet à la HSL d'avoir accès au contenu des gouttelettes lipidiques, et ainsi d'y dégrader les lipides<sup>102,114</sup>. Cependant, en période d'abondance, cette lipolyse est inhibée, et le glucose capté par les adipocytes est converti en triglycérides afin de constituer une réserve énergétique<sup>102</sup>. L'inhibition de la lipolyse est rendue possible grâce à l'activation de la protéine AKT par l'insuline : AKT peut alors activer par phosphorylation l'enzyme phosphodiesterase 3B (PDE3B), permettant de diminuer la concentration intracellulaire en AMPc et ainsi diminuer l'activité de la PKA<sup>114</sup>. Par ce mécanisme, l'insuline permet donc d'inhiber la dégradation des triglycérides à la fois en inhibant l'action lipolytique de la HSL, mais aussi en maintenant la barrière protectrice, composée de périlipines, autour des gouttelettes lipidiques<sup>102</sup>.

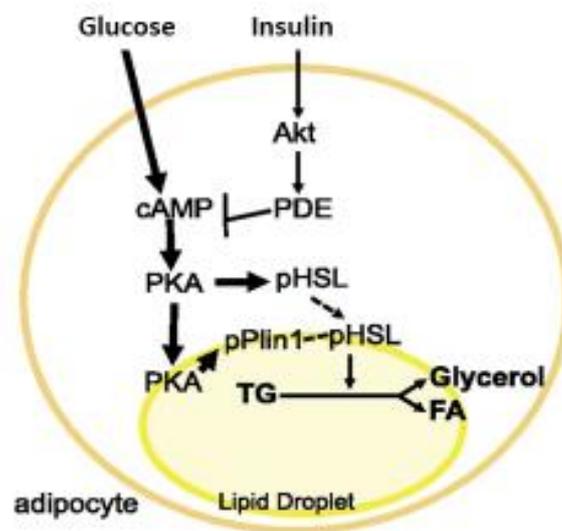


Figure 17. Actions de l'insuline dans la cellule adipeuse. **cAMP** : AMP cyclique, **PKA** : protéine kinase A, **PDE** : phosphodiesterase, **pHSL** : lipase hormonosensible phosphorylée, **pPLIN1** : périlipine 1 phosphorylée, **TG** : triglycérides, **FA** : acides gras. Source : adapté de Kuo et al., « *Pik3r1 Is Required for Glucocorticoid-Induced Perilipin 1 Phosphorylation in Lipid Droplet for Adipocyte Lipolysis* », *Diabetes*, 2017.

### 5.2.2.3 Effets de l'insuline dans le foie

Le foie est un organe important pour le stockage du glucose, mais également pour la synthèse de celui-ci. L'action de l'insuline sur les cellules hépatiques joue ainsi principalement

sur les processus de glycogénolyse et de néoglucogenèse<sup>102</sup>. Tout comme pour le muscle squelettique et le tissu adipeux, l'action de l'insuline sur les adipocytes permet l'activation de la protéine AKT, suivie de l'inhibition de la protéine GSK3, favorisant l'activité de la glycogène synthase et ainsi le stockage du glucose sous forme de glycogène<sup>102</sup>. En outre, l'insuline permet également d'inhiber la néoglucogenèse et la glycogénolyse hépatiques (figure 18).

En cas de jeun, la néoglucogenèse et la glycogénolyse sont deux processus importants, car ils permettent respectivement de produire du glucose *de novo* et de libérer du glucose à partir du stock de glycogène, en réponse à des molécules telles que le glucagon ou la noradrénaline<sup>102</sup>. En période d'abondance, l'insuline permet d'empêcher la libération de glucose à partir du foie en inhibant ces processus<sup>102</sup>.

L'insuline permet d'inhiber l'expression de deux gènes importants pour la néoglucogenèse et la glycogénolyse : il s'agit des gènes codant la phosphoénolpyruvate carboxykinase (PEPCK), une enzyme limitante de la voie de la néoglucogenèse, et la glucose-6-phosphatase (G-6-Pase), qui permet de déphosphoryler le glucose-6-phosphate retenu dans la cellule en glucose, permettant ainsi son transport hors de la cellule via le transporteur GLUT2<sup>102</sup>. Au sein de la région promotrice de ces gènes, on retrouve des éléments de réponse à l'insuline (IRE), qui permettent de réguler leur expression. En conditions basales, le facteur de transcription forkhead box O (FOXO) et la protéine peroxisome-proliférateur-activé récepteur-gamma coactivateur-1 (PGC-1) se lient aux éléments IRE dans les régions promotrices des gènes cibles, activant ainsi leur transcription. En période d'abondance, l'action de l'insuline sur les hépatocytes active la protéine AKT, qui phosphoryle alors FOXO. La phosphorylation de FOXO entraîne son exclusion du noyau, réduisant ainsi l'activation de la transcription de ces gènes et altérant de ce fait la production hépatique de glucose<sup>102,115</sup>. D'autre part, l'insuline, via l'activation de la protéine AKT, est également impliquée dans la transcription du facteur de transcription stéroïde régulateur élément-bindant 1 (SREBP-1), mais aussi dans sa translocation vers le noyau<sup>102</sup>. Une fois présent dans le noyau, SREBP-1 peut réguler l'expression de nombreux gènes contenant un élément de réponse au sérum (SRE) dans leur promoteur : on peut notamment citer les gènes codant la protéine PEPCK, mais aussi ceux codant la GSK3 et la pyruvate kinase, participant ainsi à l'activation de la glycolyse et l'inhibition de la néoglucogenèse<sup>102,116</sup>.

En outre, via son action sur SREBP-1, l'insuline peut également réguler la lipogénèse hépatique. Notamment, elle permet l'activation de l'acétyl-CoA carboxylase (ACC) par déphosphorylation de celle-ci, probablement via l'inhibition de l'AMPK, permettant ainsi la génération de malonyl-CoA servant d'intermédiaire à la synthèse des acides gras<sup>117,118</sup>. D'autre part, l'insuline active également via les protéines AKT et GSK3 l'ATP citrate lyase, permettant ainsi de convertir le citrate issu du cycle de Krebs en acétyl-CoA nécessaire à la lipogénèse<sup>119,120</sup>. De cette façon, la SREBP-1, et donc indirectement l'insuline, promeuvent la glycolyse et inhibent la néoglucogénèse, mais favorisent aussi la lipogénèse *de novo* dans le foie<sup>102</sup>.

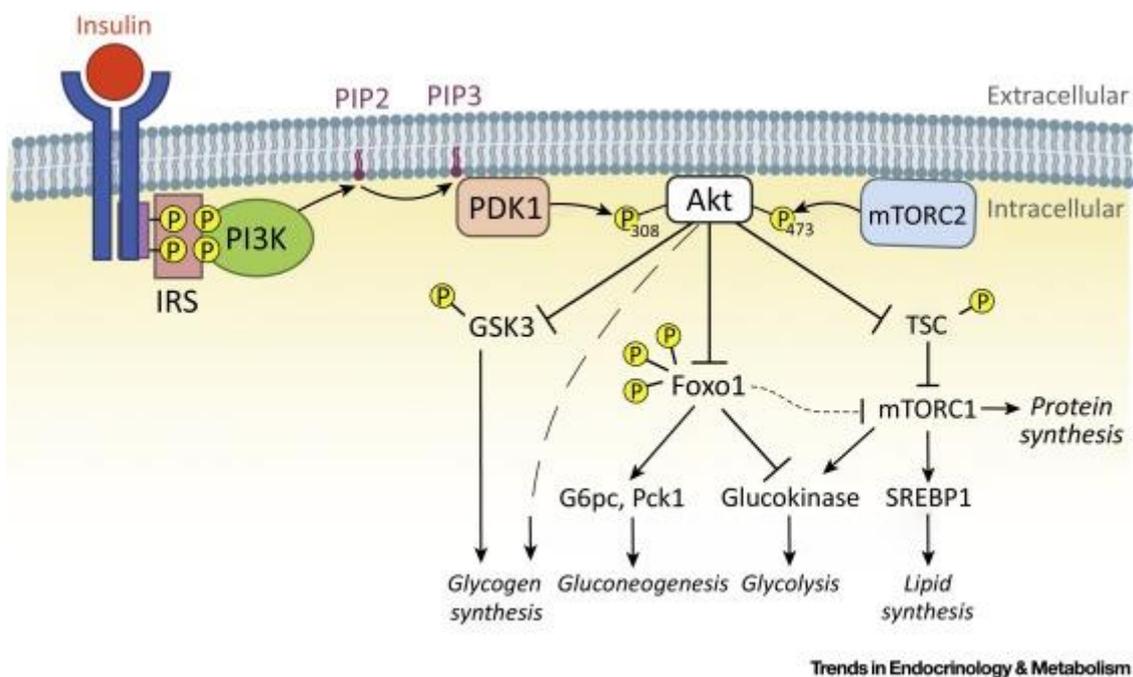


Figure 18. Actions de l'insuline dans le foie. **IRS** : insulin receptor substrate, **PI3K** : phosphoinositide 3-kinase, **PIP2** : phosphatidylinositol-4,5-bisphosphate, **PIP3** : phosphatidylinositol-3,4,5-trisphosphate, **PDK1** : phosphoinositide dependent kinase 1, **mTORC1/2** : mammalian target of rapamycin complex 1/2, **GSK3** : glycogène synthase kinase 3, **Foxo1** : Forkhead Box O1, **G6pc** : glucose-6-phosphatase, **Pck1** : phosphoenolpyruvate carboxykinase 1, **TSC** : tuberous sclerosis protein, **SREBP1** : sterol regulatory element-binding protein 1. Source : Titchenell et al., « Unraveling the Regulation of Hepatic Metabolism by Insulin », *Trends in Endocrinology and Metabolism*, 2017.

## 6 Prise en charge du diabète

Afin de maintenir une bonne homéostasie glucidique et éviter les complications associées au diabète, il est nécessaire de traiter le DT2 dès son diagnostic<sup>121,122</sup>. Ces traitements peuvent être administrés en mono-, bi- ou trithérapie, selon différentes combinaisons, en fonction de la gravité du diabète et des contre-indications possibles<sup>123</sup>. Ci-dessous sont repris les principaux traitements utilisés dans le cadre du DT2. Les associations possibles entre les différents traitements sont reprises à la figure 19.

## 6.1 Modifications du style de vie

L'alimentation et l'exercice physique, impliqués dans la balance énergétique, correspondent à la première ligne d'action contre le DT2 avant l'utilisation de traitements pharmacologiques<sup>121</sup>. Une étude prospective réalisée en 2013 a démontré qu'une modification drastique du style de vie chez des patients obèses atteints de DT2 permettait de réduire leur poids corporel ainsi que leur taux de HbA1c, reflétant une amélioration du contrôle de la glycémie<sup>124</sup>.

### 6.1.1 L'alimentation

Une adaptation du régime alimentaire des patients diabétiques permet d'améliorer la glycémie, mais également la pression sanguine, le métabolisme lipidique, et le poids corporel<sup>124</sup>. Plusieurs études cliniques ont démontré qu'un régime alimentaire adapté permettait de diminuer le taux d'hémoglobine glyquée HbA1c<sup>125</sup>.

De nombreux patients atteints de DT2 sont également atteints de surpoids ou d'obésité, qui contribuent à l'insulinorésistance et à une altération de l'insulinosécrétion, aggravant ainsi le diabète chez ces patients<sup>126</sup>. L'application d'un régime hypocalorique permet de diminuer leur poids corporel, et ainsi d'améliorer leur sensibilité à l'insuline<sup>127</sup>. En outre, les patients qui ne sont pas en surpoids, quant à eux, peuvent également voir leur régime alimentaire adapté, en conservant le même nombre de kilocalories ingérées par jour<sup>121</sup>. Ces adaptations comprennent notamment un contrôle de la quantité et de la provenance de carbohydrates ingérés, étant donné que ceux-ci contribuent directement au contrôle glycémique<sup>121</sup>.

### 6.1.2 L'activité physique

La pratique de l'exercice physique chez les patients diabétiques entraîne toute une série de conséquences bénéfiques : augmentation de la sensibilité à l'insuline, amélioration du contrôle glycémique, amélioration de la pression artérielle et du profil lipidique, perte de poids, protection cardiovasculaire, ou encore amélioration de la qualité de vie<sup>121,128,129</sup>. L'exercice permet notamment la diminution du taux de HbA1c<sup>121</sup>.

Il est conseillé aux patients diabétiques de pratiquer une activité physique régulière, au moins trois fois par semaine, en cumulant au moins 150 minutes d'activité physique par semaine<sup>5,130</sup>. De plus, il leur est également conseillé d'augmenter leur dépense énergétique dans la vie de tous les jours, par exemple en marchant plus ou en prenant les escaliers<sup>131</sup>.

## 6.2 Traitements pharmacologiques

Lorsque les modifications du style de vie ne sont pas suffisantes pour améliorer le contrôle glycémique, il est nécessaire d'administrer aux patients un traitement pharmacologique. La plupart des traitements pharmacologiques indiqués pour le traitement du diabète sont repris ci-dessous.

### 6.2.1 Traitements oraux

#### 6.2.1.1 Metformine

La metformine est un traitement oral, issu de la classe des biguanides, et est généralement considérée comme le traitement de première ligne pour le DT2<sup>123</sup>. Elle agit dans le foie, où elle inhibe la néoglucogenèse par des mécanismes AMPK-dépendants et indépendants<sup>132</sup>. Ce traitement permet une diminution de la glycémie à jeun d'environ 20%, et du taux de HbA1c de 1,5%<sup>121</sup>. Elle a également une action hypolipémiante : elle entraîne une diminution de la concentration en acides gras libres, en triglycérides et en cholestérol transporté par les lipoprotéines à faible densité (low density lipoproteins [LDL]), et une augmentation des niveaux de HDL-cholestérol<sup>121</sup>.

La metformine est le traitement optimal à administrer en monothérapie. Elle présente l'avantage d'être peu chère, sûre, sans effet sur le poids corporel, et ne provoque pas d'hypoglycémie, en plus d'avoir un effet protecteur sur le système cardiovasculaire<sup>123</sup>. Cependant, si la metformine n'est pas suffisante, elle peut être combinée en bithérapie ou trithérapie (en fonction de la gravité de l'hyperglycémie) avec d'autres traitements tels que les sulfonylurées, les thiazolidinediones (TZD), les inhibiteurs de dipeptidyl peptidase-4 (DPP-4) et du cotransporteur sodium-glucose de type 2 (SGLT2), les agonistes du GLP1R, ou encore l'insuline<sup>123,121</sup>.

La metformine peut provoquer certains effets secondaires, dont les plus fréquents sont des troubles gastrointestinaux modérés, tels que des nausées, de la diarrhée ou encore des crampes intestinales<sup>121,123</sup>. De plus, il existe plusieurs contre-indications concernant son utilisation<sup>123</sup>.

#### 6.2.1.2 Sulfonylurées et glinides

Les sulfonylurées et les glinides sont des sécrétagogues de l'insuline. Ils appartiennent à deux classes différentes d'agents hypoglycémiant, mais agissent sur la sécrétion d'insuline

via un mécanisme similaire<sup>121</sup>. Ces deux molécules peuvent se lier sur les canaux  $K^+_{ATP}$  à des endroits différents, induisant ainsi la fermeture de ces canaux. La dépolarisation membranaire qui en résulte permet ainsi un influx d'ions  $Ca^{2+}$ , qui favorise la sécrétion d'insuline<sup>121</sup>. Les sulfonylurées sont généralement utilisées en seconde ligne comme traitement du diabète<sup>133</sup>. Parmi ceux-ci, on peut citer le glyburide (ou glibenclamide), le glipizide, le gliclazide ou encore le glimepiride<sup>134</sup>. Leur effet principal est de réduire la concentration plasmatique de glucose à jeun<sup>135</sup>. Les glinides, quant à eux, comprennent le repaglinide et le nateglinide<sup>121</sup>. Ces molécules sont plus rapidement absorbées que les sulfonylurées, et ont une action plus rapide sur la stimulation de l'insulinosécrétion<sup>121,136</sup>. Ils ont pour effet de réduire la glycémie postprandiale<sup>135</sup>. Les sulfonylurées et les glinides peuvent tous les deux être employés seuls pour traiter le diabète, mais ils peuvent également être combinés à de l'insuline ou à d'autres agents hypoglycémisants<sup>121</sup>.

L'utilisation de sulfonylurées présente de nombreux avantages : ils sont peu chers, ont une bonne capacité à réduire la glycémie, et ont une bonne efficacité à long terme<sup>136</sup>. Les glinides ont quant à eux un effet un peu moins fort sur la réduction de la glycémie<sup>137</sup>.

L'utilisation de ces agents hypoglycémisants peut provoquer quelques effets secondaires. Tout d'abord, l'utilisation de tels agents sécrétagogues peut présenter une perte d'efficacité au cours du temps, liée à un épuisement des cellules  $\beta$  pancréatiques, et à l'apoptose de celles-ci<sup>138</sup>. D'autre part, leur utilisation peut induire une prise de poids corporel, mais également provoquer des hypoglycémies<sup>139</sup>.

#### 6.2.1.3 Inhibiteurs d' $\alpha$ -glucosidase

La catégorie des inhibiteurs d' $\alpha$ -glucosidase comprend 3 agents : l'acarbose, le miglitol et le voglibose<sup>140</sup>. Ils permettent de réduire les hyperglycémies postprandiales sans agir directement sur la sécrétion de d'insuline<sup>140,141</sup>. L' $\alpha$ -glucosidase est une enzyme située dans la bordure en brosse de l'intestin grêle, et permet d'hydrolyser les oligosaccharides en monosaccharides. Les inhibiteurs d' $\alpha$ -glucosidase ont une structure similaire aux oligosaccharides mais présentent une plus forte affinité que ceux-ci pour l'enzyme. En s'y liant, ils permettent de l'inhiber<sup>140,141</sup>. Ceci permet de retarder l'absorption et la digestion des carbohydrates, réduisant ainsi l'hyperglycémie postprandiale<sup>141</sup>. De plus, les carbohydrates restés dans l'intestin permettent d'augmenter la sécrétion de GLP-1, favorisant la sécrétion d'insuline<sup>121</sup>.

Les inhibiteurs d' $\alpha$ -glucosidase ont l'avantage de ne pas provoquer d'hypoglycémie<sup>123</sup>. Cependant, ils ont une efficacité moyenne, et génèrent également quelques effets secondaires gastrointestinaux modérés, parmi lesquels des flatulences, de la diarrhée ou des crampes abdominales<sup>121</sup>.

#### 6.2.1.4 Thiazolidinediones

Les TZD, bien que n'étant plus utilisées en Europe, sont encore utilisées aux Etats-Unis sous deux formes différentes : la rosiglitazone, qui est un ligand exclusif du récepteur peroxisome proliferator-activated receptor  $\gamma$  (PPAR $\gamma$ ), et la pioglitazone, qui est capable de se lier à la fois à PPAR $\gamma$  et à PPAR $\alpha$ <sup>121</sup>. Ces deux molécules, en agissant sur les récepteurs PPAR $\gamma$  et PPAR $\alpha$  permettent, d'une part, d'augmenter la sensibilité à l'insuline en agissant sur le muscle, le tissu adipeux et le foie pour augmenter l'utilisation du glucose, et d'autre part de diminuer la production de glucose<sup>121</sup>. De plus, ces agents améliorent également la glycémie en préservant la fonction des cellules  $\beta$  pancréatiques, et n'entraînent pas d'hypoglycémie<sup>121,123</sup>. Cependant, ces deux traitements ont des effets différents sur le profil lipidique. Bien que les deux molécules augmentent la concentration en HDL-cholestérol, la pioglitazone semble plus bénéfique : elle diminue la concentration en LDL-cholestérol et en triglycérides, alors que la rosiglitazone augmente la concentration en LDL-cholestérol et a un effet moins fort sur la réduction des niveaux de triglycérides sériques<sup>121</sup>.

Les TZD ont comme effet secondaire d'entraîner une augmentation du poids corporel liée à une rétention de fluide, à l'activation du récepteur PPAR $\gamma$  dans le système nerveux central augmentant l'appétit, ou encore à la surexpression de gènes participant au stockage des graisses dans les lipides<sup>121,123</sup>. D'autre part, ces molécules augmentent le risque d'insuffisance cardiaque, de fractures osseuses, et d'infarctus du myocarde<sup>123</sup>.

#### 6.2.1.5 Inhibiteurs de dipeptidyl peptidase-4

La catégorie des inhibiteurs de l'enzyme DPP-4 comprend plusieurs molécules telles que la sitagliptine, la saxagliptine, la linagliptine ou l'alogliptine. Ces inhibiteurs permettent d'empêcher la dégradation rapide du GLP-1 par l'enzyme DPP-4, permettant ainsi de maintenir la présence du GLP-1, dont la liaison au récepteur GLP1R permet d'augmenter l'insulinosécrétion et d'inhiber la sécrétion de glucagon<sup>121</sup>. Ainsi, les inhibiteurs de DPP-4 permettent d'améliorer le contrôle glycémique<sup>121,123</sup>.

Ces agents antidiabétiques sont généralement bien tolérés chez les patients, présentent peu de risque d'hypoglycémie, et n'ont pas d'effet sur le poids corporel<sup>121,123</sup>. Leurs effets secondaires comprennent le mal de tête, des infections des voies nasopharyngées et urinaires, et des troubles dermatologiques provoqués par l'immunité (angio-œdème, urticaire)<sup>123,142</sup>.

#### 6.2.1.6 Inhibiteurs du cotransporteur sodium-glucose SGLT2

Les inhibiteurs de SGLT2 comprennent la dapagliflozine, la canagliflozine et l'empagliflozine, et permettent de réduire la réabsorption rénale du glucose<sup>121</sup>. En conditions normales, les glomérules du rein permettent la filtration du glucose, et une partie de celui-ci est réabsorbée dans le tubule proximal par des transporteurs passifs ou actifs tels que le cotransporteur SGLT2. Les cotransporteurs SGLT2 du tubule proximal participent à environ 90% de la réabsorption du glucose. L'inhibition de SGLT2 permet donc d'empêcher en partie la réabsorption du glucose, qui est alors excrété dans l'urine. Ceci permet ainsi de diminuer la glycémie, indépendamment de l'action de l'insuline<sup>121</sup>.

L'action insulino-indépendante des inhibiteurs de SGLT2 permet de réduire le risque d'hypoglycémie, ainsi que le risque d'épuisement des cellules  $\beta$  pancréatiques<sup>143</sup>. Ce traitement présente également l'avantage de réduire le poids corporel ainsi que la pression sanguine systolique et diastolique<sup>143</sup>, probablement via la diurèse osmotique qu'il induit<sup>144</sup>. Finalement, ces inhibiteurs auraient également un effet protecteur sur les reins, en les protégeant de l'hyperfiltration glomérulaire associée au cotransporteur SGLT2<sup>145</sup>.

Cependant, la glycosurie provoquée par les inhibiteurs de SGLT2 peut induire des infections des voies urogénitales<sup>146</sup>. De plus, la diurèse osmotique induite par ceux-ci peut également provoquer des hypotensions orthostatiques, et une hypovolémie<sup>121,123</sup>.

### 6.2.2 Traitements injectables

#### 6.2.2.1 Agonistes du récepteur au GLP-1

Comme expliqué ci-dessus, le GLP-1 est rapidement dégradé par l'enzyme DDP-4. Pour lutter contre cela, des agonistes du GLP1R, ayant le même effet physiologique que le GLP-1 mais avec une durée d'action étendue (à court terme ou à long terme) ont été développés. On peut notamment citer l'exénatide, le liraglutide, l'albiglutide ou le dulaglutide<sup>121</sup>. Tout comme le GLP-1, ces agonistes stimulent la sécrétion d'insuline et inhibent la sécrétion de

glucagon, ralentissent la vidange gastrique, et augmentent la sensation de satiété<sup>123</sup>. Ils ont l'avantage de ne pas provoquer d'hypoglycémie, et permettent également de réduire le poids corporel et la pression sanguine. Ils diminueraient également certains facteurs de risque cardiovasculaire<sup>123</sup>.

Les effets secondaires principaux des agonistes du GLP1R sont des troubles gastrointestinaux, tels que des nausées, des vomissements ou de la diarrhée, ainsi qu'une augmentation du rythme cardiaque. D'autre part, l'administration de ces traitements peut provoquer des réactions au points d'injection, telles que des abcès, une cellulite infectieuse ou une nécrose, de manière plus fréquente que pour les injections d'insuline<sup>121</sup>.

#### 6.2.2.2 Insuline

L'insuline est utilisée comme traitement chez les patients diabétiques intolérants aux autres traitements antidiabétiques, atteints d'insuffisance rénale ou hépatique, ou atteints d'un défaut primaire des cellules  $\beta$  pancréatiques, mais est également recommandée dans le cas où les patients conservent un haut taux de HbA1c malgré une modification du style de vie et des doses maximales de traitements antidiabétiques<sup>147</sup>.

Les patients peuvent être traités avec de l'insuline humaine, ou par des analogues de l'insuline, permettant de mimer son effet. On peut notamment citer les analogues à action rapide, à action intermédiaire, ou à action prolongée<sup>148</sup>. Les analogues à action intermédiaire (insuline Neutral Protamin Hagedorn), à action longue (glargine et detemir) ou ultra-longue (degludec) permettent de reproduire l'effet basal de l'insuline<sup>148</sup>. En effet, ils fournissent une faible quantité d'insuline mais maintenue dans le temps, et permettent ainsi de maintenir un contrôle de la glycémie pendant environ 24 heures<sup>148,149</sup>. Ils permettent ainsi de supprimer la production hépatique de glucose entre les repas et pendant la nuit<sup>148</sup>. Les analogues à action rapide (aspart, lispor, glulisine), quant à eux, sont utilisés pour mimer la réponse de l'insuline endogène à un repas. Ils sont injectés avant la prise alimentaire afin de corriger l'hyperglycémie post-prandiale<sup>149</sup>. Cependant, la plupart des patients sont traités avec un mix d'insuline à action courte et prolongée, permettant de mieux contrôler la glycémie, et réduisant les risques d'hyperglycémie nocturne<sup>148</sup>.

Les injections d'insuline peuvent provoquer plusieurs effets secondaires : une hypoglycémie, un gain de poids corporel, ainsi que des réactions cutanées au point d'injection<sup>123</sup>.

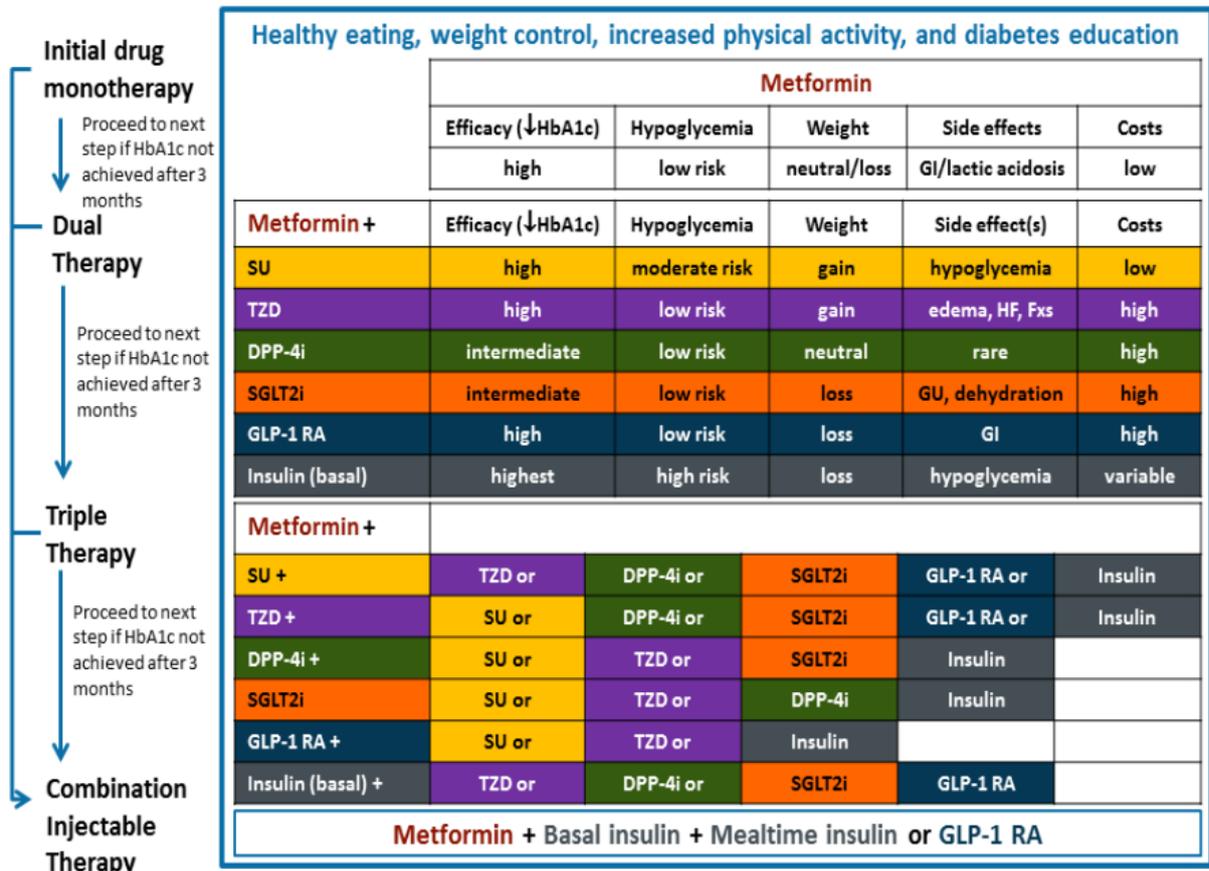


Figure 19. Recommandations générales pour le traitement du DT2. SU : sulfonyles, TZD : thiazolidinediones, DPP-4i : inhibiteurs de dipeptidyl peptidase-4, SGLT2i : inhibiteurs de SGLT2, GLP1 RA : agonistes du récepteur au GLP-1, HbA1c : hémoglobine glyquée, GI : troubles gastrointestinaux, HF : insuffisance cardiaque, Fxs : fractures, GU : troubles génito-urinaires. Source : Dagnaw et al., « National Training on Diabetes Mellitus for HCWs Participant's Manual3 », 2019.

## 7 Génétique du diabète de type 2

Le DT2 peut être causé par des facteurs environnementaux (âge, alimentation, sédentarité...) mais également par des facteurs génétiques. En fonction de son étiologie génétique, on peut le classer en 2 types : les formes monogéniques, provoquées par des mutations d'un seul gène jouant un rôle majeur dans la cellule β pancréatique, et les formes polygéniques, apparaissant quant à elles suite à l'interaction de multiples mutations génétiques au sein de plusieurs gènes avec des facteurs environnementaux<sup>150</sup>. Ces formes seront décrites en détails plus loin.

## 7.1 Les outils génétiques

Afin de pouvoir identifier les régions du génome, ou plus précisément les gènes impliqués dans l'étiologie de ces différentes formes de diabète, plusieurs stratégies ont été mises en place. Celles-ci utilisent des outils génétiques particuliers, qui se basent principalement sur des techniques de génotypage et de séquençage.

### 7.1.1 Analyses de liaison familiale

Les analyses de liaison familiale ont été largement réalisées durant les années 1990 afin d'identifier les régions chromosomiques contenant les gènes impliqués dans le développement du DT2. Celles-ci sont réalisées au sein de familles où plusieurs individus sont atteints de la maladie. Ces études supposent qu'il existe des gènes jouant un rôle majeur, et dont les mutations sont transmises sur plusieurs générations, entraînant ainsi le développement de la maladie chez ces individus<sup>151</sup>. Le but est donc d'identifier s'il existe une coségrégation entre certains marqueurs génétiques et la maladie, afin de pouvoir déterminer la région génétique qui y est associée.

Les études de liaison se basent sur la liaison génétique entre plusieurs loci. La distance génétique entre 2 loci, correspondant à la probabilité selon laquelle ils puissent être séparés par un crossing-over au moment de la méiose, est calculée en centiMorgan (cM), où 1 cM correspond à une fréquence de recombinaison de 1% entre les 2 loci. Au plus cette distance génétique est faible, au plus la probabilité que ces loci soient séparés par un crossing-over diminue : on suppose ainsi que ces loci seront transmis ensemble. Cette distance génétique est calculée sur base de la fréquence de recombinaison génétique dans la descendance : ainsi, si dans la population étudiée, la fréquence de recombinaison génétique est égale à la fréquence de non-recombinaison génétique, on considère que ces loci sont indépendants. Par contre, si la fréquence de recombinaison génétique observée est inférieure à la fréquence de non-recombinaison génétique, on considère que ces loci sont génétiquement liés. La probabilité de transmission simultanée de ces allèles à la descendance permet de calculer un LOD Score Z (logarithm of odds ou logarithme des probabilités), permettant de déterminer si la région génétique définie par ces loci est liée à la maladie. Si ce score est supérieur à 3, on suppose que la région chromosomique est liée à la maladie ; si ce score est inférieur à -2, on considère qu'il n'y a pas de liaison entre les deux.

Les études de liaison familiale sont basées sur le génotypage de plusieurs marqueurs génétiques (voir les différents marqueurs ci-dessous), espacés d'une certaine distance génétique, chez plusieurs apparentés (sains ou malades) d'une même famille. Le calcul du LOD score entre ces différents marqueurs permet ainsi d'identifier quelle est la région présentant un pic de liaison, dans laquelle se retrouve le gène associé à la maladie.

Ces études ont principalement contribué à la découverte de gènes associés au diabète monogénique, qui présente une forte pénétrance (la présence de l'allèle muté augmente très fortement la probabilité de développer la maladie). Cependant, ce type d'étude a été beaucoup moins efficace dans le cadre de l'investigation du diabète polygénique complexe, dans lequel plusieurs gènes peuvent être impliqués, et qui présente une pénétrance plus faible<sup>152</sup>.

#### 7.1.1.1 Génotypage de restriction fragment length polymorphisms

Les polymorphismes de type restriction fragment length polymorphisms (RFLP) sont définis par la présence de différents sites de restriction de long de l'acide désoxyribonucléique (ADN) génomique des individus. Ce type de polymorphisme peut être utilisé comme marqueur génétique afin d'identifier des variations génomiques entre les individus d'une même famille, grâce à la technique du Southern blot. Brièvement, l'ADN génomique des individus à étudier est fragmenté au niveau des différents sites de restriction par des enzymes. En fonction de la présence ou non de polymorphismes ayant modifié certains sites de restriction, la longueur des fragments peut différer entre les individus. Ces fragments sont ensuite séparés selon leur taille par électrophorèse, puis transférés sur une membrane de nylon, et dénaturés. Les fragments d'ADN peuvent ensuite être révélés par hybridation de ceux-ci à une sonde d'ADN marquée. Les fragments comportant un polymorphisme ayant altéré un site de restriction, n'ayant donc pas été clivé par l'enzyme, seront donc plus grands que les fragments dans lesquels le site de restriction n'a pas été modifié<sup>153,154</sup> (figure 20). La taille des fragments observés permet de ce fait de déterminer si un polymorphisme est présent dans une certaine région chromosomique chez les individus malades ou non, permettant ainsi le calcul du LOD score pour cette région. Cependant, à cause de sa complexité, le génotypage basé sur les RFLP a rapidement été abandonné au profit du génotypage par microsatellites lorsque c'était possible.

## Les marqueurs RFLP

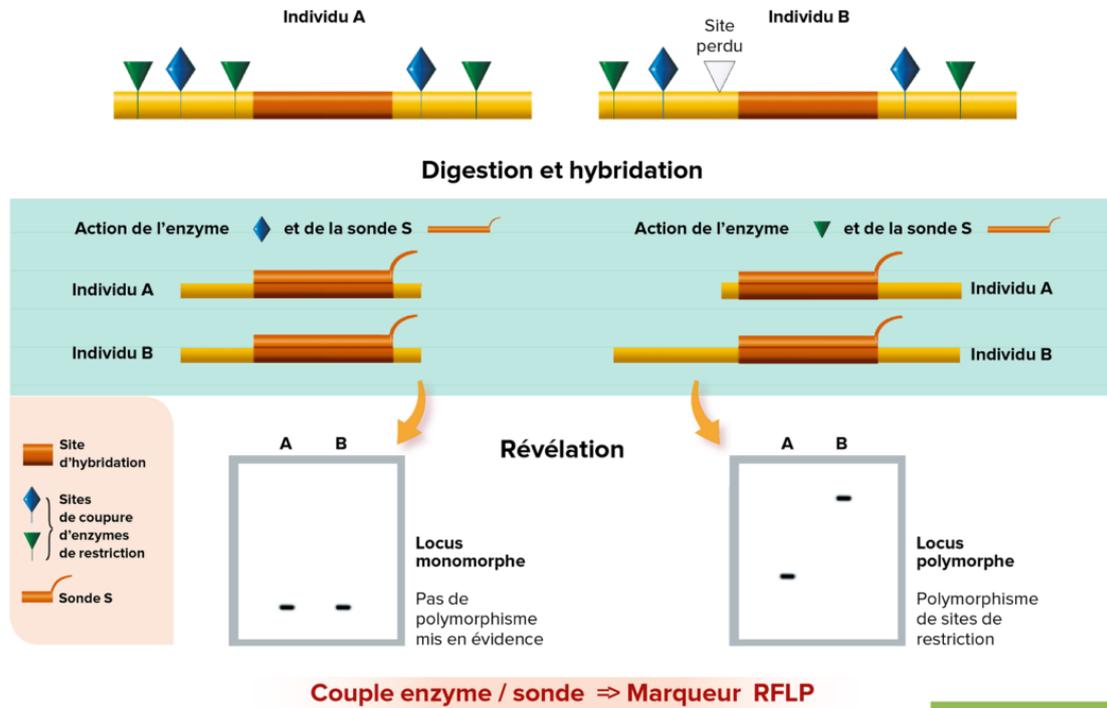
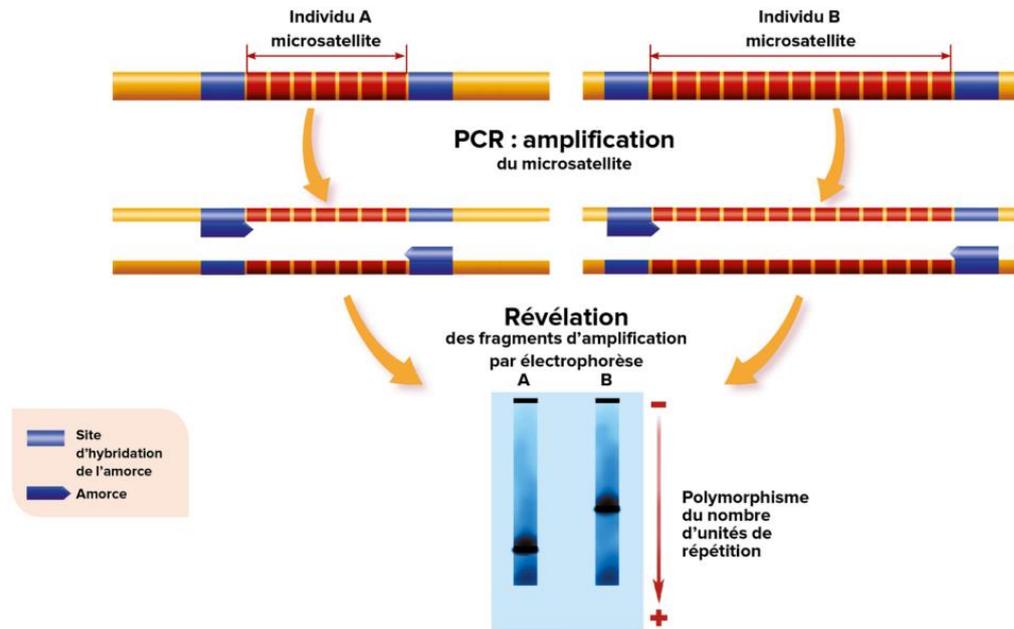


Figure 20. Technique de génotypage de RFLP. Source : issu du site web <https://www.gnis-pedagogie.org/sujet/marqueurs-moleculaires/>.

### 7.1.1.2 Génotypage des séquences microsatellites

Les séquences microsatellites, également appelées short sequence repeats (SSR), sont de courtes séquences d'ADN de 1 à 4 nucléotides répétées en tandem, hautement polymorphes, et dont le nombre de répétitions varie entre les individus d'une même famille suite aux réarrangements chromosomiques entre les chromosomes parentaux. Ainsi, une étude du polymorphisme de taille des microsatellites peut être réalisée afin d'identifier les polymorphismes de taille communs entre les individus malades dans une même famille. Le génotypage de séquences microsatellites est réalisé grâce à la technique de polymérase chain reaction (PCR). Concrètement, à partir de l'ADN génomique des individus à étudier, le locus microsatellite est amplifié à l'aide d'amorces ciblant les zones d'ADN en amont et en aval de ce locus, permettant ainsi l'amplification spécifique de cette région. Les fragments d'ADN amplifiés sont ensuite séparés selon leur taille par électrophorèse<sup>153,154</sup> (figure 21). Il est alors possible de déterminer si des polymorphismes de taille sont présents dans certaines régions génomiques chez les individus malades par rapport aux individus sains d'une même famille. Le calcul du LOD score permet ainsi de définir si ces régions sont liées au DT2.

## Les marqueurs microsatellites



Paire d'amorces spécifiques bordant le microsatellite  $\Rightarrow$  Marqueur microsatellite

© GNIS-PEDAGOGIE.ORG

Figure 21. Technique de génotypage des marqueurs microsatellites. Source : issu du site web <https://www.gnis-pedagogie.org/sujet/marqueurs-moleculaires/>.

### 7.1.1.3 Séquençage Sanger

Le séquençage Sanger est une technique de séquençage permettant de vérifier de manière précise la séquence de courtes régions d'ADN (allant jusqu'à 1000 paires de bases) d'un individu, au niveau des pics de liaison. Ainsi, ce type de séquençage permet de séquencer l'ADN génomique des individus étudiés dans les analyses de liaison familiale, uniquement au niveau du pic de liaison déterminé par génotypage, de façon à identifier quelle est la mutation exacte associée au développement de la maladie. Le principe de ce séquençage consiste en une PCR au cours de laquelle des désoxynucléotides (dNTP) ainsi qu'une plus faible concentration en didésoxynucléotides (ddNTPs) fluorescents sont utilisés pour l'extension du brin d'ADN nouvellement formé. Lorsqu'un ddNTP fluorescent est incorporé dans la séquence d'ADN au lieu d'un dNTP classique, aucun autre dNTP ne peut venir s'y lier et l'extension du brin est interrompue. De cette façon, des amplicons de différentes longueurs, dont le dernier nucléotide est fluorescent, sont obtenus. Ces amplicons sont ensuite séparés en fonction de leur taille par électrophorèse capillaire. La lecture de la fluorescence émise par chaque ddNTP, situé à l'extrémité de ces amplicons, par le séquenceur permet ainsi de définir la séquence du

brin d'ADN complémentaire (ADNc)<sup>155</sup> (figure 22). Par alignement sur une séquence de référence, il est alors possible d'identifier les mutations présentes dans la région d'ADN étudiée.

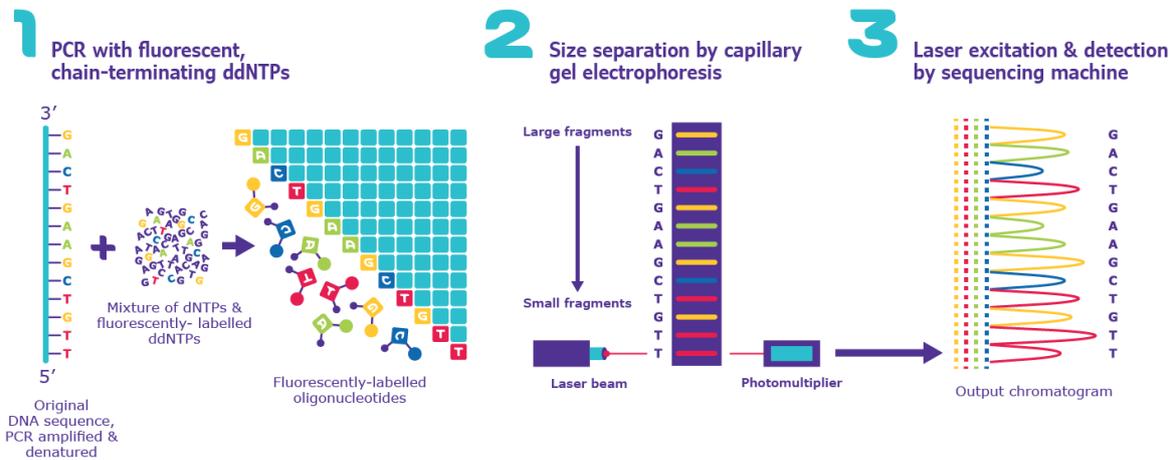


Figure 22. Technique du séquençage Sanger. Source : issu du site web <https://www.sigmaaldrich.com/BE/en/technical-documents/protocol/genomics/sequencing/sanger-sequencing>.

### 7.1.2 Approche gène-candidat

Comme leur nom l'indique, les études par approche gène-candidat se focalisent sur des gènes déjà suspectés pour être impliqués dans la pathogenèse du DT2. Ces gènes sont sélectionnés selon différents critères : rôle connu ou présumé dans la régulation de l'homéostasie glucidique (métabolisme du glucose, sécrétion de l'insuline ou transduction de son signal), association connue avec un autre type de diabète tel que le MODY, association avec un diabète ou des traits associés dans des modèles animaux, ou expression différente en condition normale ou diabétique<sup>156</sup>. Les régions codantes de ces gènes candidats sont séquencées afin d'identifier des mutations potentiellement impliquées dans le développement de la maladie.

Cette approche gène-candidat a ainsi permis d'identifier plusieurs mutations délétères dans des gènes clés impliqués dans le diabète monogénique, et également certains polymorphismes d'un seul nucléotide ou single-nucleotide polymorphisms (SNP) associés au diabète polygénique. Étonnamment, de nombreux gènes étudiés étant donné leur implication dans la sécrétion et l'action de l'insuline n'ont pas démontré d'association avec le DT2<sup>152</sup>.

### 7.1.3 Etudes d'association pangénomiques ou GWAS

Les études d'association pangénomiques ou genome-wide association studies (GWAS), publiées pour la première fois en 2007, ont permis une grande avancée dans la compréhension des événements génétiques sous-jacents au DT2. Ces études se basent sur l'identification de plusieurs milliers de SNPs fréquents, généralement non-codants (situés dans des régions intergéniques ou introniques), à l'aide de micropuces à ADN spécifiques appelées SNPs microarrays dans des cohortes de plusieurs milliers d'individus, permettant ainsi d'étudier l'association de ces SNPs avec le DT2 (lors d'études cas-contrôles, entre individus malades et individus sains) ou avec certains traits métaboliques quantitatifs associés au DT2 (lors d'études réalisées chez des individus non-diabétiques)<sup>151</sup>.

Les SNPs microarrays consistent en des puces, sur lesquelles sont fixées plusieurs milliers de courtes sondes d'ADN spécifiquement choisies pour pouvoir détecter des SNPs particuliers. Deux techniques différentes de détection des SNPs sont possibles :

- Concernant la première technique (correspondant aux puces Affymetrix), pour chaque SNP à investiguer, une sonde capable de s'hybrider aux 2 allèles est liée à la puce. Afin d'identifier les individus porteurs des allèles mutés, leur ADN génomique est fragmenté, amplifié par PCR, puis ces fragments sont ajoutés à la puce. L'hybridation des fragments d'ADN sur les sondes correspondantes permet d'induire une fluorescence variable de celles-ci, dépendant de l'allèle présent sur le fragment : s'il s'agit de l'allèle sauvage, son hybridation sera parfaite, et la fluorescence émise sera importante ; par contre, s'il s'agit de l'allèle muté, celui-ci ne s'hybridera pas correctement à la position du SNP, et la fluorescence émise sera plus faible<sup>157</sup> (figure 23). La différence d'intensité de fluorescence permet ainsi d'identifier quels sont les individus porteurs d'allèles mutés.
- Pour la deuxième technique (correspondant aux puces Illumina), les sondes utilisées ont une séquence complémentaire aux régions d'intérêt de l'ADN contenant les SNPs, mais s'arrêtent juste avant la position du SNP d'intérêt. Afin d'identifier les individus porteurs de ces mutations, l'ADN génomique des patients est fragmenté, puis ces fragments sont amplifiés et ajoutés à la puce afin de s'hybrider aux sondes d'ADN complémentaires. Des ddNTPs fluorescents ainsi que de l'ADN polymérase sont ensuite ajoutés à la puce afin de prolonger la séquence d'ADN de chaque sonde d'un

seul nucléotide, correspondant à la position du SNP d'intérêt, par complémentarité au fragment d'ADN qui s'y est hybridé. Ainsi, en fonction de la couleur du signal fluorescent émis, il est possible de déterminer quel nucléotide s'est incorporé à la position du SNP<sup>153,157</sup> (figure 23). Par complémentarité, on peut donc définir si l'individu est porteur du SNP ou non.

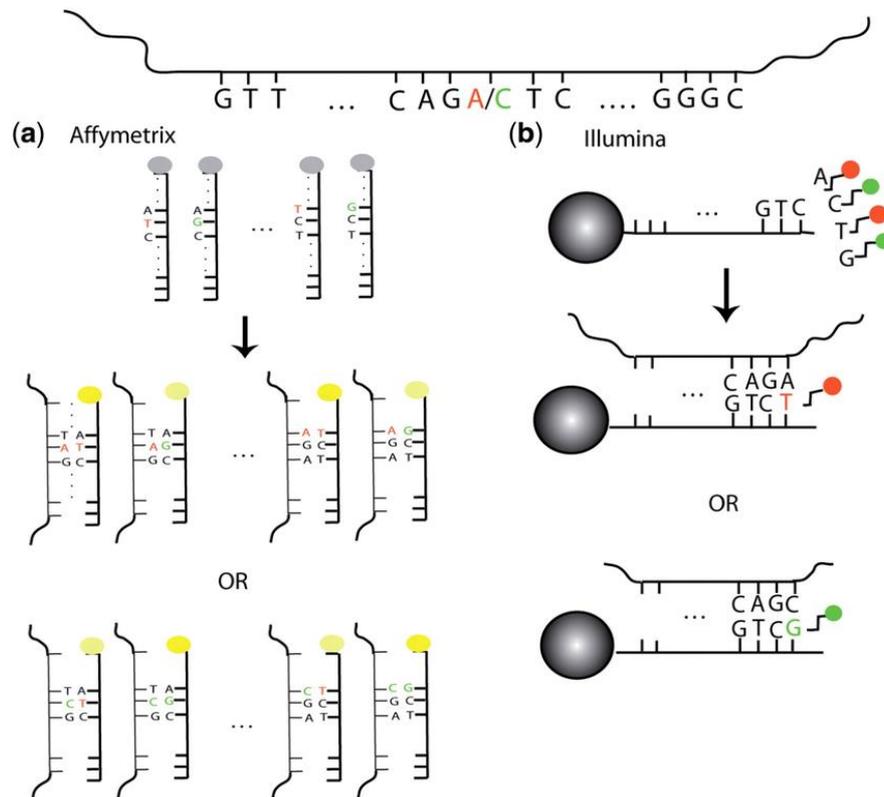


Figure 23. Techniques de génotypage de SNPs sur micropuce. (a) Le génotypage de SNPs selon la technologie Affymetrix est basé sur l'utilisation de sondes d'ADN, capables de se lier aux 2 allèles (WT ou muté), mais avec une affinité plus forte pour l'allèle sauvage, générant une forte fluorescence, par rapport à l'allèle muté (fluorescence plus faible). L'intensité de la fluorescence permet ainsi de définir quel allèle est présent dans l'ADN de l'individu étudié. (b) Le génotypage de SNPs selon la technologie Illumina est basé sur l'utilisation de sondes pouvant s'hybrider aux fragments d'ADN portant le SNP d'intérêt, mais s'arrêtant juste avant le nucléotide complémentaire au SNP. Des nucléotides fluorescents peuvent alors prolonger la séquence de la sonde par complémentarité à l'allèle porté par l'individu étudié. La couleur de la fluorescence émise permet ainsi de définir quel allèle est porté par l'individu. Source : LaFramboise, « Single nucleotide polymorphism arrays: A decade of biological, computational and technological advances », *Nucleic Acids Research*, 2009.

Des analyses statistiques sont ensuite réalisées afin d'attribuer une valeur  $P$  à chaque SNP, indiquant la probabilité que celui-ci soit associé au DT2 ou au trait métabolique quantitatif étudié. Si cette valeur  $P$  est inférieure à un certain seuil (généralement établi à  $5 \times 10^{-8}$  en appliquant la règle des tests multiples), cela signifie que le variant est bien associé au trait ou à la pathologie étudiée. De cette façon, il est possible de déterminer quels variants sont associés au trait ou à la maladie, et s'ils ont un effet protecteur ou délétère<sup>158</sup>.

Il est à noter que pour plus de facilité, les SNPs identifiés comme étant associés à la maladie ou au trait étudié sont généralement nommés selon le gène le plus proche. Cependant, cela ne signifie pas que ce gène est directement impliqué dans l'étiologie génétique de la maladie<sup>150</sup>.

Bien que les études GWAS aient grandement contribué à l'identification de multiples loci associés au DT2, elles présentent tout de même quelques faiblesses. D'abord, les micropuces sont capables de génotyper, au maximum, quelques millions de SNPs, alors que plus de 30 millions de SNPs ont été identifiés dans le génome humain. Pour pallier à cela, les données obtenues par les études GWAS sont imputées à l'aide de données issues de séquençage de génome entier, souvent issues du projet « 1000 Genomes ». Ceci permet d'estimer l'effet d'autres variants, associés à ces SNPs, sur le risque de développer un DT2<sup>150</sup>. D'autre part, la plupart des variants génétiques identifiés comme étant associés avec le DT2 par les études GWAS sont des variants fréquents, avec une fréquence d'allèle mineur (minor allele frequency [MAF]) supérieure à 5%. Cependant, ces variants fréquents n'expliquent qu'une petite partie de l'héritabilité du DT2 : chaque variant considéré seul augmente de moins de 15% le risque de développer la maladie, et ensemble, ils n'expliquent que 15% de l'héritabilité du DT2<sup>150,159</sup>. De plus, ces SNPs peuvent ne pas être directement associés à la maladie. En effet, ces polymorphismes sont présents dans des régions en déséquilibre de liaison, et servent de « porte-drapeaux » à un ensemble de SNPs transmis ensemble. Ainsi, un SNP mis en évidence par la puce à ADN ne signifie pas nécessairement qu'il est causal, mais que d'autres SNPs présents dans le même locus génétique peuvent contribuer à l'étiologie génétique de la maladie ou du trait métabolique étudié<sup>150,158</sup>. En outre, il est également possible que ces SNPs identifiés par GWAS et associés à la maladie n'aient pas d'effet direct sur le gène qui y est associé, mais qu'ils impactent d'autres gènes situés plus loin sur le chromosome, comme cela a déjà été démontré pour des variants associés au gène *FTO* qui impactent en réalité le gène *IRX3*<sup>160</sup>, ou même des gènes situés ailleurs dans le génome<sup>150</sup>.

#### 7.1.4 Séquençage de nouvelle génération ou NGS

Le séquençage de nouvelle génération ou Next Generation Sequencing (NGS) est une technologie développée il y a seulement quelques années. Contrairement au séquençage Sanger, qui ne permet de lire qu'une petite séquence d'ADN de manière ponctuelle, le séquençage de type NGS permet de séquencer l'ensemble du génome de milliers d'individus

en une fois, très rapidement, et à faible coût. Cette technologie a donc permis une grande avancée dans la génétique : alors que le premier séquençage du génome humain dans le cadre du « Human Genome Project » a duré 13 ans et a coûté 3 milliards de dollars<sup>161</sup>, il est désormais possible de séquencer le génome entier d'un individu en quelques jours, pour environ 1000 dollars. Le séquençage NGS peut être utilisé pour séquencer le génome entier (whole genome sequencing [WGS]), uniquement l'exome (whole exome sequencing [WES]), ou encore un panel de gènes particuliers sélectionnés pour leur implication potentielle dans une maladie par exemple<sup>162</sup>. Le séquençage NGS permet notamment l'identification de variants très rares (MAF inférieure à 1%) et de faible fréquence (MAF comprise entre 1 et 5%), qui sont difficiles à identifier avec les autres outils génétiques étant donné leur faible fréquence dans la population.

Il existe différents types de séquençage NGS, mais le séquençage par synthèse est la technique majoritairement utilisée. Celui-ci est composé de 4 étapes : la préparation de la librairie d'ADN, l'amplification de l'ADN, le séquençage à proprement parler, puis l'alignement des fragments obtenus sur le génome de référence (figure 24). La préparation de la librairie est réalisée grâce à la fragmentation de l'ADN génomique en fragments aléatoires par sonication de l'échantillon ou par l'utilisation d'enzymes de restriction. Si une région particulière du génome est ciblée (WES ou panel de gènes), les fragments d'ADN correspondant aux régions d'intérêt sont sélectionnés par hybridation à des sondes d'ADN spécifiques. Ensuite, les fragments sont liés à des adaptateurs, permettant ainsi la liaison de la librairie de fragments sur une cellule en verre, appelée flow cell, et permettant également de marquer chaque fragment par un index afin de retrouver l'individu d'origine pour chaque fragment. Les fragments d'ADN sont ensuite amplifiés par PCR sur la flow cell, puis le séquençage par synthèse à proprement parler est réalisé. Brièvement, des amorces de séquençage s'apparient sur les adaptateurs liés aux fragments d'ADN, puis de l'ADN polymérase ainsi que des nucléotides fluorescents, portant une molécule empêchant la liaison d'autres nucléotides à leur suite, sont ajoutés à la flow cell. Ainsi, le premier nucléotide complémentaire à celui du fragment d'ADN étudié vient rallonger la séquence d'ADN de la sonde. Le signal fluorescent émis par ce nucléotide est ensuite lu par le séquenceur. La molécule empêchant l'extension de la synthèse du brin d'ADN ainsi que le fluorophore sont ensuite clivés du nucléotide, puis un nouveau cycle est à nouveau lancé. De cette façon, par

ajout et lecture successifs des différents nucléotides complémentaires au fragment d'ADN, il est possible de lire la séquence de chacun d'entre eux. La séquence de chacun des reads obtenus est alors alignée sur la séquence d'un génome de référence<sup>163</sup>. Il est ainsi possible de reconstituer la séquence entière du génome d'un individu dans le cas du WGS, ou les régions d'intérêt ciblées dans le cas du WES ou de l'étude d'un panel de gènes. De cette façon, on peut donc retrouver quels sont les variants rares de certains gènes, présents chez les individus étudiés, par rapport au génome de référence.

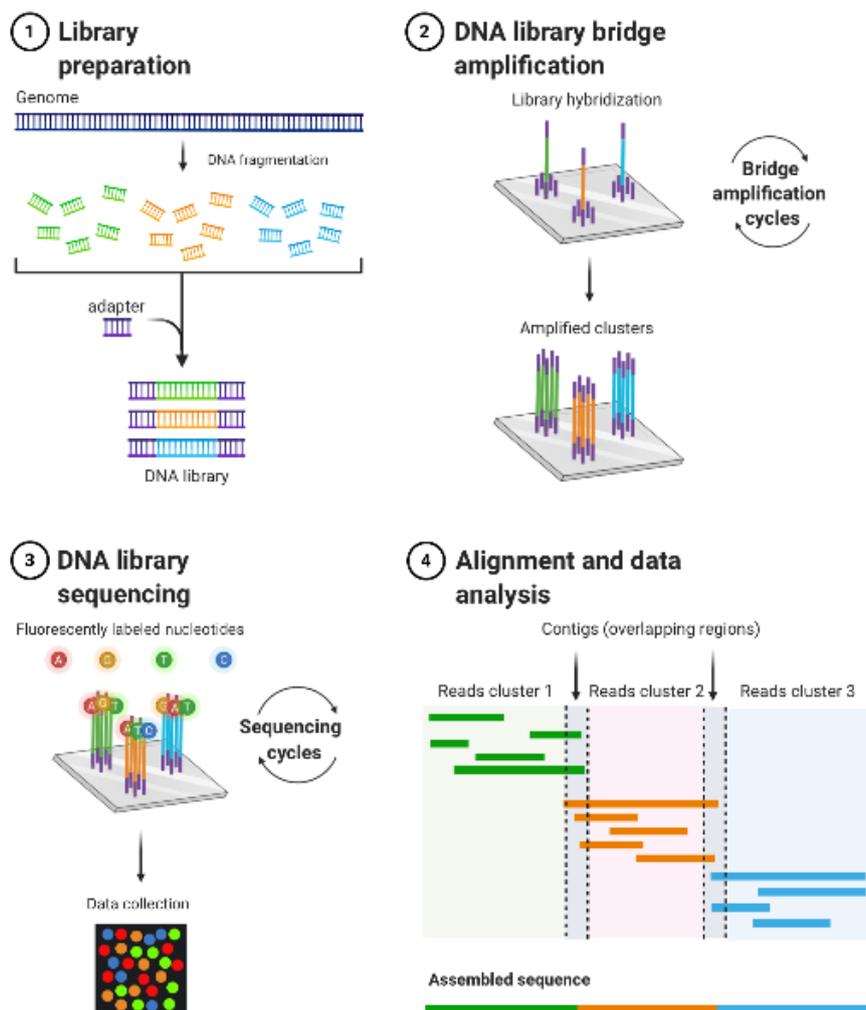


Figure 24. Technique du séquençage NGS. Source : issu du site web <https://app.biorender.com/biorender-templates/figures/5c95180bc2753f33003dacd3>.

Afin de déterminer si les variants identifiés par séquençage NGS sont associés au risque de développer la maladie, des études d'association sont réalisées entre ces variants rares et la maladie. Cependant, leur faible fréquence réduit fortement la puissance statistique des

études, étant donné que très peu d'individus atteints de ces variants sont disponibles pour étudier le phénotype : il est donc difficile d'établir une association entre ces variants et la maladie. Ces études sont donc généralement réalisées non pas entre un variant et la maladie, mais entre un cluster (ou une agrégation) de variants et la maladie. En augmentant le nombre de variants étudiés simultanément, et donc le nombre de patients chez qui le phénotype peut être investigué, on peut augmenter la puissance statistique et ainsi déterminer s'il existe une association entre ces variants et le DT2<sup>164</sup>. De cette façon, il est possible de confirmer l'association d'un gène avec la maladie, ou de découvrir de nouveaux gènes qui y sont associés.

## 7.2 Les formes monogéniques du diabète de type 2

Les formes monogéniques du DT2 correspondent à 1-2% des cas de diabète<sup>165</sup>. Elles sont provoquées par une mutation dans un seul gène impliqué dans le développement ou la fonction de la cellule  $\beta$  pancréatique, qui est suffisante pour causer à elle seule l'apparition de la maladie<sup>166</sup>. Ces mutations sont rares, avec une MAF inférieure à 1%.

Parmi les formes monogéniques du diabète, on peut notamment citer le diabète de type MODY (maturity-onset diabetes of the young), ainsi que le diabète néonatal<sup>167</sup>, qui seront présentés plus en en détails ci-après.

### 7.2.1 Le MODY

Le diabète de type MODY est défini comme étant une forme familiale de diabète autosomique dominante, à apparition précoce (avant l'âge de 25 ans), et qui est généralement associée à un défaut primaire de sécrétion d'insuline<sup>168</sup>. Ce type de diabète est provoqué par des mutations dans des gènes jouant un rôle important dans la fonction de la cellule  $\beta$  pancréatique, parmi lesquels on peut citer *GCK* (identifié par notre laboratoire), *HNF1A* et *HNF4A*, entre autres. Au total, ce sont 14 gènes qui ont été identifiés comme étant impliqués dans le MODY. Les 6 types de MODY principaux sont associés à des mutations dans les gènes *GCK*, *HNF1A*, *HNF4A*, *HNF1B*, *PDX1*, et *NEUROD1*, et seront expliqués plus en détails ci-dessous. Les autres types de MODY connus sont quant à eux provoqués par des mutations des gènes *CEL*, *PAX4*, *INS*, *ABCC8*, *KCNJ11* et *APPL1*, et potentiellement par des mutations des gènes *KLF11* et *BLK*, bien que leur implication soit controversée<sup>167</sup>. Les différents types de

MODY, ainsi que leur fréquence, leur physiopathologie, le phénotype associé et leur traitement sont repris dans le tableau 1.

Plusieurs critères, basés notamment sur des caractéristiques cliniques, ont été établis afin de diagnostiquer un diabète MODY chez les patients<sup>169</sup>. Cependant, ce diabète reste souvent mal diagnostiqué. En 2010, il a été estimé qu'environ 80% des patients souffrant de MODY sont diagnostiqués comme étant atteints d'un DT1 ou DT2<sup>170</sup>. Le diagnostic génétique, sur lequel nous reviendrons plus tard, est donc important<sup>171</sup>.

#### 7.2.1.1 *HNF4A*/MODY1

Le MODY1 est provoqué par une mutation du gène *HNF4A* codant le facteur de transcription hépatocyte nuclear factor-4 $\alpha$  (HNF-4 $\alpha$ ), et correspond à 5 à 10% des cas de MODY<sup>172</sup>. Ce facteur de transcription est exprimé dans le foie, l'intestin, les reins et les îlots pancréatiques<sup>173-175</sup>. HNF-4 $\alpha$  est impliqué dans la régulation du métabolisme et du transport du glucose, en agissant sur l'expression de gènes tels que ceux codant le transporteur GLUT2, la glycéraldéhyde 3-phosphate déshydrogénase (GAPDH), la pyruvate kinase, ou encore la sous-unité Kir6.2 des canaux K<sup>+</sup><sub>ATP</sub><sup>176,177</sup>, et la préproinsuline<sup>178</sup>. De plus, ce gène, tout comme *HNF1A* et *HNF1B*, est capable de réguler l'expression génique au cours du développement embryonnaire<sup>176</sup>. En outre, il a été démontré que ces facteurs de transcription font partie du même réseau de régulation transcriptionnelle, et sont capables de se réguler mutuellement<sup>179,180</sup>.

L'implication du gène *HNF4A* dans le développement du MODY1 a été démontrée grâce à la combinaison de plusieurs études. Une première étude de liaison familiale, basée sur le génotypage de 75 polymorphismes génétiques a été réalisée chez différents membres d'une même famille (composée de plus de 275 individus sur 5 générations), dans laquelle le diabète suit un mode de transmission autosomique dominant. Cette étude a démontré une coségrégation entre le bras long du chromosome 20 et le développement du diabète MODY, et a suspecté le gène *ADA*, codant l'adénosine déaminase comme le gène causal<sup>181,182</sup>. Cependant, une seconde étude, basée sur l'étude de 71 marqueurs génétiques du bras long du chromosome 20, a permis de restreindre la région associée au MODY1 à une région de 7 Mb uniquement, dans laquelle le gène *HNF4A* est présent<sup>183</sup>. Ainsi, en 1996, un séquençage des 11 exons de ce gène a été réalisé chez certains individus d'une même famille, et a permis l'identification de la mutation non-sens p.Q268\* qui coségrège avec le développement de la

maladie<sup>184</sup>. Par la suite, d'autres mutations de ce gène ont été identifiées comme étant associées au développement d'un MODY<sup>185</sup>, confirmant l'implication de ce gène dans le développement du MODY1.

Les patients souffrant de MODY1 présentent une réduction de la fonction  $\beta$  pancréatique, marquée par une réduction progressive de l'insulinosécrétion, mais conservent une bonne sensibilité à l'insuline<sup>185</sup>. Ils présentent également une réduction des niveaux de protéines apolipoprotéines A1 et A2 (APOA1 et APOA2) sériques et de HDL-cholestérol<sup>185</sup>, une altération de la sécrétion de glucagon, et des complications microvasculaires<sup>186</sup>. Étonnamment, les mutations du gène *HNF4A* sont associées à une macrosomie fœtale et une hypoglycémie néonatale liée à une hyperinsulinémie<sup>187</sup>.

L'identification du type de MODY dont souffrent les patients permet de leur proposer un traitement adapté. Ainsi, concernant les individus souffrant de MODY1, ceux-ci peuvent être traités avec des sulfonylurées<sup>185,188</sup>. Si cela n'est pas suffisant, on peut combiner ce traitement avec des inhibiteurs de DPP-4, des agonistes du GLP1R, ou encore de l'insuline<sup>171,189</sup>.

#### 7.2.1.2 *GCK/MODY2*

Le MODY2, provoqué par des mutations du gène *GCK*, correspond à 30 à 50% des cas de MODY<sup>172</sup>. Comme décrit précédemment, la *GCK* est une enzyme qui permet de catalyser la première réaction de la glycolyse, consistant en la phosphorylation du glucose en glucose-6-phosphate. Les mutations du gène *GCK* sont associées à une altération de son activité enzymatique, entraînant ainsi une diminution du flux glycolytique dans les cellules  $\beta$  pancréatiques<sup>190</sup>.

Le rôle du gène *GCK* dans le développement du diabète MODY2 a été démontré pour la première fois en 1992 par notre laboratoire, par une étude de liaison réalisée chez 16 familles d'origine française, permettant une investigation génétique de 304 membres dont 105 sont atteints de MODY. Par le génotypage de marqueurs microsatellites, il a été démontré que le locus du gène *GCK*, situé sur le chromosome 7, est lié à une forme de MODY nommée MODY2<sup>191</sup>. Cette étude fut la première à prouver que le DT2 peut être causé par des facteurs génétiques, et que l'insulinorésistance n'est pas le seul paramètre qui entraîne le développement du DT2<sup>150</sup>. Des mutations ponctuelles de *GCK* associées au MODY2 ont par la

suite été décrites<sup>192</sup>. Il est à noter que le MODY2 est hérité de manière autosomique dominante, et est associé à des mutations hétérozygotes du gène *GCK*. La présence de mutations homozygotes est quant à elle associée à un diabète néonatal permanent<sup>172</sup>.

Les patients atteints de MODY2 présentent une hyperglycémie modérée liée à une mauvaise sensibilité au glucose, nécessitant ainsi une glycémie plus élevée pour stimuler la sécrétion d'insuline<sup>193</sup>. Cette hyperglycémie est également favorisée par une altération de la glycogénogenèse, et une augmentation de la néoglucogenèse hépatique<sup>194</sup>.

Les patients atteints de MODY2 ne répondent ni aux hypoglycémifiants oraux, ni à l'insuline, et il n'est pas intéressant de les traiter pharmacologiquement<sup>195</sup>. Un bon contrôle de leur alimentation est suffisant pour contrôler leur hyperglycémie<sup>169</sup>.

#### 7.2.1.3 *HNF1A*/MODY3

Le MODY3 est provoqué par des mutations du gène *HNF1A* codant le facteur de transcription HNF-1 $\alpha$ , et correspond au type de MODY le plus fréquent en représentant 30 à 65% des cas<sup>172</sup>. HNF-1 $\alpha$  est exprimé dans le pancréas et le foie, mais aussi dans d'autres organes tels que l'intestin et les reins<sup>196,197</sup>. Tout comme HNF-4 $\alpha$ , avec lequel il peut interagir, HNF-1 $\alpha$  régule également l'expression de gènes impliqués dans le transport du glucose ainsi que dans la sécrétion d'insuline, tels que celui codant le transporteur de glucose GLUT2<sup>198</sup>.

L'implication du gène *HNF1A* dans le développement du MODY3 a été démontrée par notre laboratoire en 1995. Cette étude, basée sur le génotypage de 97 marqueurs microsatellites (répétitions de la séquence CA) dans le génome d'individus issus de 12 familles d'origine française atteintes de MODY, a démontré l'existence d'une coségrégation entre une région de 10 cM située dans le bras long du chromosome 12, et le développement du diabète de type MODY, suggérant l'implication d'un gène situé dans cette région chromosomique dans le développement de la maladie<sup>199</sup>. Une seconde étude a permis de réduire cet intervalle à 5 cM<sup>200</sup>. Finalement, une étude réalisée en 1996 s'est focalisée sur l'étude de cette région chromosomique chez des individus issus de 7 familles différentes dont certains membres sont atteints de MODY, afin d'investiguer les gènes présents dans ce locus, parmi lesquels le gène *HNF1A*. Grâce au séquençage de ce gène, il a été démontré que plusieurs mutations de celui-ci coségrègent avec le développement d'un diabète MODY<sup>201</sup>. Des recherches effectuées

ensuite par d'autres équipes ont permis d'identifier d'autres mutations de ce gène responsables du MODY3<sup>202</sup>.

Les patients atteints de MODY3 présentent généralement une hyperglycémie modérée liée à une mauvaise sécrétion d'insuline en réponse au glucose durant l'enfance, qui se transforme par la suite en diabète durant l'adolescence ou au début de l'âge adulte<sup>203</sup>. Il s'agit d'une forme de diabète assez sévère, et les patients développent généralement une dépendance à l'insuline, ainsi que des complications microvasculaires<sup>204</sup>. Etant donné que HNF-4 $\alpha$  est exprimé dans les cellules tubulaires rénales, les patients peuvent également souffrir de glycosurie liée à une altération du transport du glucose dans les tubules rénaux<sup>205</sup>.

Tout comme pour les patients atteints de MODY1, ceux atteints de MODY3 peuvent recevoir un traitement à base de sulfonylurées, qui peut être complété avec un agoniste du GLP1R, des inhibiteurs de DPP-4, ou de l'insuline<sup>171,189</sup>.

#### 7.2.1.4 *PDX1*/MODY4

Le gène *PDX1* code le facteur de transcription pancreatic and duodenal homeobox 1, et correspond à moins de 1% des cas de MODY<sup>189</sup>. Ce gène est impliqué dans le développement du pancréas, mais également dans sa fonction. Notamment, *PDX1* est exprimé dans la cellule  $\beta$  pancréatique, où il régule l'expression des gènes codant la préproinsuline, le transporteur au glucose GLUT2, et la GCK<sup>169</sup>. Des mutations de ce gène sont donc impliquées dans une agénésie pancréatique, des problèmes de développement des cellules  $\beta$  pancréatiques, et une altération de la sécrétion d'insuline<sup>206,207</sup>.

C'est en 1997 qu'il a été démontré par étude d'association familiale que des mutations de *PDX1* étaient associées au développement du MODY4. Cette étude a été réalisée dans la famille d'un individu présentant une agénésie du pancréas liée à la mutation homozygote p.P63Rfs\*60 située dans l'exon 1 du gène *PDX1*, et dont plusieurs membres présentaient un diabète transmis de manière autosomique dominante. Par génotypage de cette mutation chez plusieurs membres de la famille, une coségrégation entre cette mutation à l'état hétérozygote et le développement d'un MODY a été démontrée<sup>208</sup>. D'autres variants causant la maladie ont été identifiés par la suite<sup>209</sup>.

Les patients atteints de MODY4 peuvent présenter un phénotype variable, allant d'une simple intolérance au glucose à un diabète insulino-dépendant<sup>169</sup>. Ils sont généralement

traités avec de l'insuline, ou bien certains antidiabétiques oraux tels que la metformine ou les inhibiteurs de DPP-4<sup>186</sup>.

#### 7.2.1.5 *HNF1B*/MODY5

Le gène *HNF1B* code le facteur de transcription HNF-1 $\beta$ , et est responsable de moins de 5% des cas de MODY<sup>172</sup>. Ce gène est impliqué dans le développement de plusieurs organes, tels que le pancréas, les reins, le tractus urinaire, le foie, les poumons et les intestins<sup>189</sup>. Des mutations de ce gène peuvent donc engendrer des problèmes liés au mauvais développement de ces organes tels qu'une agénésie ou une hypoplasie pancréatique<sup>210,211</sup>, une insuffisance rénale, ou encore la formation de reins polykystiques et de méganéphrons<sup>212</sup>.

L'implication de mutations du gène *HNF1B* dans le développement du MODY5 a été mise en évidence en 1997. Etant donné que les gènes *HNF1A* et *HNF4A* sont impliqués dans le développement du type MODY, cette étude s'est intéressée au gène *HNF1B* faisant partie de la même famille. Le séquençage de ce gène chez 57 individus d'origine Japonaise atteints de MODY a permis d'identifier plusieurs substitutions nucléotidiques de ce gène, parmi lesquelles la mutation p.R177\* présente chez un individu. L'étude de la famille de cet individu a ensuite démontré la coségrégation de cette mutation avec le diabète de type MODY, confirmant ainsi le rôle de ce gène dans le développement du MODY5<sup>213</sup>. Cette étude a été suivie de l'identification d'autres variants de *HNF1B* impliqués dans le développement de la maladie<sup>214</sup>.

Les patients atteints de MODY5 présentent généralement une dyslipidémie, avec un faible taux de HDL-cholestérol mais une triglycéridémie élevée. Leur diabète se manifeste généralement à l'adolescence ou au début de l'âge adulte, et évolue souvent vers une dépendance à l'insuline. Ces patients peuvent également présenter une dysfonction rénale<sup>189</sup>.

Ces patients sont généralement traités avec de l'insuline, étant donné qu'ils répondent mal aux agents hypoglycémisants oraux tels que les sulfonylurées<sup>171</sup>.

#### 7.2.1.6 *NEUROD1*/MODY6

Le gène *NEUROD1*, codant le facteur de transcription neuronal differentiation 1 (*NEUROD1*) est associé au MODY6, et correspond à moins de 1% des cas de MODY<sup>189</sup>. *NEUROD1* est impliqué dans le développement pancréatique et neuronal. Il participe

également à l'activation de la transcription de gènes tels que ceux codant la préproinsuline, la sous-unité SUR1 des canaux  $K^+_{ATP}$ , ou encore la GCK en se liant à leur promoteur<sup>186</sup>.

En 1999, une première étude s'est intéressée à l'implication du gène *NEUROD1* dans le développement du DT2. Au cours de cette étude, l'exon 2 de ce gène ainsi que les introns flanquants ont été séquencés chez 94 individus atteints de DT2 issus de familles présentant un diabète familial transmis de manière autosomique dominante, et 96 individus contrôles. Cela a permis d'identifier 4 mutations dans l'exon 2 du gène, parmi lesquelles 2 d'entre elles (p.R111L et p.H206Pfs\*38), affectant sa capacité à se lier au promoteur de l'insuline et son domaine de transactivation, ont été identifiées à l'état hétérozygote chez 2 individus différents atteints de DT2, et absentes chez les individus contrôles. Ces 2 mutations ont ensuite été identifiées chez d'autres membres de la famille de ces 2 individus porteurs de celles-ci, mettant en avant la coségrégation de ces mutations avec le développement du DT2<sup>215</sup>. En 2001, une étude de liaison familiale islandaise, réalisée dans 3 familles présentant une histoire familiale de MODY, a réalisé un génotypage de 39 marqueurs microsatellites situés sur 9 chromosomes différents. Cette étude a permis de mettre en évidence la liaison de marqueurs, dont un situé à proximité du gène *NEUROD1*, avec le MODY dans l'une des familles étudiées. L'exon 2 et les introns flanquants de ce gène ont alors été séquencés, permettant ainsi l'identification du variant p.E110K coségrégant avec le diabète MODY6 dans cette famille<sup>216</sup>, suivi de l'identification de nouveaux variants dans une étude japonaise<sup>217</sup>. Il est à noter que si la présence de mutations hétérozygotes de ce gène est associée au développement d'un MODY, la présence de mutations homozygotes est quant à elle impliquée dans le développement d'un diabète néonatal ainsi que de troubles neurologiques et de difficultés d'apprentissage<sup>218</sup>.

Les patients atteints de MODY6 présentent une altération de la sécrétion d'insuline en réponse en glucose, avec une sévérité variable<sup>186</sup>. Ces patients peuvent également développer des complications microvasculaires<sup>219</sup>.

La prise en charge des patients atteints de MODY6 ne nécessite généralement pas de traitement à l'insuline<sup>219</sup>.

Tableau 1. Tableau reprenant les différents types de MODY connus, le gène impliqué ainsi que la protéine qu'il code, la fréquence du type de MODY par rapport à l'ensemble des MODY, la physiopathologie et les caractéristiques cliniques associées, ainsi que les traitements envisagés. Adapté de Yahaya et Ufuoma, "Genetics and Pathophysiology of Maturity-onset Diabetes of the Young (MODY): A Review of Current Trends", Oman Medical Journal, 2014; Delvecchio et al., "Treatment Options for MODY Patients: A Systematic Review of Literature", Diabetes Therapy, 2020; Urakami, "Maturity-onset diabetes of the young (MODY): current perspectives on diagnosis and treatment", Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy, 2019; Naylor et al., "Maturity-Onset Diabetes of the Young Overview", GeneReviews, 2018.

MODY et gène impliqué	Protéine impliquée	Fréquence (% des MODY)	Physiopathologie	Caractéristiques cliniques	Traitements
<b>MODY1 – HNF4A</b>	Hepatocyte nuclear factor-4 $\alpha$ Facteur de transcription	5-10%	Dysfonction de la cellule $\beta$ pancréatique	Diminution progressive de l'insulinosécrétion Faibles niveaux d'apolipoprotéines et de triglycérides Macrosomie néonatale et hypoglycémie hyperinsulinique transitoire Complications microvasculaires	Régime alimentaire, sulfonylurées, inhibiteurs de DPP-4, agonistes du GLP-1R, insuline
<b>MODY2 – GCK</b>	Glucokinase Enzyme	30-50%	Altération de la sensibilité au glucose	Hyperglycémie modérée Complications microvasculaires rares	Pas de traitement médicamenteux, régime alimentaire
<b>MODY3 – HNF1A</b>	Hepatocyte nuclear factor-1 $\alpha$ Facteur de transcription	30-65%	Dysfonction de la cellule $\beta$ pancréatique	Diminution progressive de l'insulinosécrétion Hypoglycémie hyperinsulinique transitoire (parfois) Glycosurie Complications microvasculaires	Sulfonylurées, meglitinides, inhibiteurs de DPP-4, agonistes du GLP-1R, inhibiteurs de SGLT-2, insuline
<b>MODY4 – PDX1</b>	Pancreatic duodenal homeobox 1 Facteur de transcription	<1%	Dysfonction de la cellule $\beta$ pancréatique	Diabète modéré Altération de l'insulinosécrétion Altération du développement des cellules $\beta$ pancréatiques)	Régime alimentaire, agents hypoglycémisants oraux, insuline
<b>MODY5 – HNF1B</b>	Hepatocyte nuclear factor-1 $\beta$ Facteur de transcription	<5%	Dysfonction de la cellule $\beta$ pancréatique	Diminution progressive de l'insulinosécrétion, avec dégradation du contrôle glycémique Hypoplasie pancréatique Malformations rénales et génito-urinaires Complications microvasculaires	Sulfonylurées, repaglinide, insuline
<b>MODY6 – NEUROD1</b>	Neurogenic differentiation 1 Facteur de transcription	<1%	Dysfonction de la cellule $\beta$ pancréatique	Altérations de la morphogenèse pancréatique et de la différenciation des cellules $\beta$ pancréatiques Hyperglycémie variable	Régime alimentaire, agents hypoglycémisants oraux, insuline

<b>MODY7 – KLF11</b>	Krüppel-like factor 11 Facteur de transcription	<1%	Dysfonction de la cellule $\beta$ pancréatique	Diminution de la sensibilité à l'insuline Hyperglycémie modérée	Agents hypoglycémiants oraux, insuline
<b>MODY8 – CEL</b>	Carboxyl ester lipase Enzyme lipolytique	<1%	Dysfonction du pancréas endocrine et exocrine	Altération de la fonction pancréatique endocrine et exocrine Fibrose et lipomatose	Agents hypoglycémiants oraux, insuline
<b>MODY9 – PAX4</b>	Paired box 4 Facteur de transcription	<1%	Dysfonction de la cellule $\beta$ pancréatique	Altération du développement et de la différenciation des cellules $\beta$ pancréatiques Hyperglycémie progressive, parfois cétoacidose	Régime alimentaire, agents hypoglycémiants oraux, insuline
<b>MODY10 – INS</b>	Insuline Hormone hypoglycémiante	<1%	Mutation du gène de l'insuline	Altération de la synthèse d'insuline, hyperglycémie et diabète	Régime alimentaire, agents hypoglycémiants oraux, insuline
<b>MODY11 – BLK</b>	$\beta$ lymphocyte kinase Enzyme tyrosine kinase	<1%	Défaut de l'insulinosécrétion	Altération relative de la sécrétion d'insuline, hyperglycémie et diabète	Régime alimentaire, agents hypoglycémiants oraux, insuline
<b>MODY12 – ABCC8</b>	ATP binding cassette subfamily C member 8 Sous-unité des canaux $K^+_{ATP}$	<1%	Dysfonction des canaux $K^+_{ATP}$	Altération de la sécrétion d'insuline, diabète	Agents hypoglycémiants oraux (sulfonylurées)
<b>MODY13 – KCNJ11</b>	Inward-rectifier potassium channel, subfamily J, member 11 Sous-unité des canaux $K^+_{ATP}$	<1%	Dysfonction des canaux $K^+_{ATP}$	Altération de la sécrétion d'insuline, diabète	Régime alimentaire, sulfonylurées
<b>MODY14 – APPL1</b>	Adaptor protein, Phosphotyrosine interacting with PH domain and Leucine Zipper 1	<1%	Défaut de l'insulinosécrétion	Altération de la sécrétion d'insuline, hyperglycémie et diabète	Régime alimentaire, agents hypoglycémiants oraux, insuline

### 7.2.2 Le diabète néonatal

Le diabète néonatal se développe généralement au cours des 6 premiers mois de vie, mais il peut apparaître jusqu'à 12 mois après la naissance<sup>220,221</sup>. Ce type de diabète est très rare, et son incidence a été estimée entre 1:90 000 et 1:400 000 naissances<sup>222-224</sup>. Il est provoqué par des mutations dans un peu plus d'une vingtaine de gènes<sup>221</sup>, qui sont généralement impliqués dans le développement du pancréas, l'apoptose des cellules  $\beta$  pancréatiques, ou encore la régulation et la maturation de l'insuline<sup>225</sup>. Ces mutations apparaissent la plupart du temps *de novo*, bien que des cas familiaux de diabète néonatal aient déjà été démontrés<sup>226</sup>.

La majorité des cas de diabète néonatal sont associés à un défaut de l'insulinosécrétion, lié à une altération du développement ou de la fonction des cellules  $\beta$  pancréatiques, ainsi qu'à leur destruction progressive<sup>221</sup>. Sur base du phénotype associé, le diabète néonatal peut être classé en deux catégories : le diabète néonatal permanent et le diabète néonatal transitoire<sup>222</sup>. Certaines de ces mutations impliquées dans le développement de ces maladies sont associées au développement de syndromes sévères, tels que les syndromes de Wolcott-Rallison ou de Wolfram, comprenant d'autres troubles en plus du diabète<sup>227,228</sup>. Les gènes principaux impliqués dans le développement de ces pathologies sont expliqués plus en détails ci-dessous. La liste exhaustive des gènes qui ont été identifiés comme étant associés au diabète néonatal permanent ou transitoire, ainsi que les caractéristiques cliniques associées sont reprises dans le tableau 2.

#### 7.2.2.1 Le diabète néonatal transitoire

Le diabète néonatal transitoire est défini comme une forme de diabète qui apparaît durant les 6 premières semaines de vie, et qui se résorbe aux alentours de 18 mois<sup>229</sup>. Cependant, environ 50% de ces patients redévelopperont un diabète plus tard dans leur vie<sup>225</sup>. Il correspond à un peu plus de la moitié des cas de diabète néonatal<sup>224,230</sup>.

La majorité des cas de diabète néonatal transitoire sont provoqués par des anomalies génétiques dans le chromosome 6q24, mais il peut également être causé, dans une moindre mesure, par des mutations dans les gènes potassium inwardly rectifying channel subfamily J member 11 (*KCNJ11*) et ATP binding cassette subfamily C member 8 (*ABCC8*), codant les sous-unités du canal  $K^+_{ATP}$ <sup>230</sup>.

### 7.2.2.2 Le diabète néonatal permanent

Le diabète néonatal permanent correspond à un peu moins de la moitié des cas de diabète néonatal<sup>224</sup>. Il a tendance à apparaître un peu plus tardivement que la forme transitoire, mais contrairement à celle-ci, il ne se résorbe pas avec le temps<sup>231</sup>. Le diabète néonatal permanent a des caractéristiques cliniques très similaires à la forme transitoire, et il est difficile de les différencier chez les nouveaux-nés<sup>231</sup>. La cause la plus courante de diabète néonatal permanent est la présence de mutations dans les gènes *KCNJ11* et *ABCC8*<sup>231</sup>, mais il peut également être provoqué par des mutations présentes dans le gène *INS*<sup>232</sup>.

### 7.2.2.3 Gènes et régions chromosomiques associés au diabète néonatal

#### 7.2.2.3.1 Locus 6p24

La majorité des cas de diabète néonatal transitoire sont liés à des anomalies des gènes soumis à empreinte parentale présents sur le locus 6q24 : ils correspondent à 68% des cas<sup>222,230</sup>. Cette région contient notamment le gène *ZAC*, codant la protéine à doigts de zinc zinc activated ion channel (*ZAC*) impliquée dans l'apoptose et l'arrêt du cycle cellulaire, et le gène *hydatiform mole-associated and imprinted (HYMAI)*<sup>229,233</sup>. Plusieurs types d'altérations de cette région chromosomique ont été reportés : perte de l'empreinte parentale suite à une disomie uniparentale paternelle du chromosome 6, menant à la présence des deux allèles paternels uniquement<sup>234</sup>, duplication paternelle, menant à la présence de 3 copies du locus 6q24, dont deux proviennent du père<sup>235</sup>, ou bien perte de la méthylation de l'allèle maternel, menant à son activation<sup>236</sup>.

Chez ces patients, le diabète apparaît généralement durant les premiers jours ou les premières semaines de vie, mais disparaît souvent durant la première année<sup>222</sup>. Cependant, il y a de fortes chances qu'une hyperglycémie réapparaisse durant l'adolescence, et celle-ci peut persister jusqu'à l'âge adulte<sup>237</sup>. Le diabète peut également être accompagné d'une macroglossie et d'une hernie ombilicale<sup>222</sup>.

Les enfants atteints de cette pathologie sont généralement traités avec de l'insuline<sup>222</sup>.

#### 7.2.2.3.2 Gènes *KCNJ11* et *ABCC8*

Les mutations activatrices des gènes *KCNJ11* et *ABCC8* sont la première cause de diabète néonatal permanent et la deuxième cause de diabète néonatal transitoire<sup>222</sup>, avec une fréquence plus importante des mutations du gène *KCNJ11* chez les cas de diabète néonatal

permanent, et du gène *ABCC8* chez les cas de diabète néonatal transitoire<sup>221</sup>. La grande majorité de ces mutations sont hétérozygotes, et exercent un effet dominant<sup>220</sup>. Ces gènes codent les deux sous-unités des canaux  $K^+_{ATP}$  : le gène *KCNJ11* code la sous-unité Kir6.2, formant le pore des canaux  $K^+_{ATP}$ , tandis que le gène *ABCC8* code la sous-unité régulatrice SUR1, sur laquelle peut venir se lier l'ATP afin d'inactiver le canal<sup>238,239</sup>. En temps normal, en présence de glucose, la cellule  $\beta$  pancréatique est capable de générer de l'ATP permettant d'inactiver le canal  $K^+_{ATP}$  en s'y liant. Ceci permet d'induire une dépolarisation membranaire qui permet l'ouverture des canaux  $Ca^{2+}$  et ainsi l'exocytose de l'insuline. Si des mutations activatrices de ce canal sont présentes dans l'un de ces gènes, celui-ci reste en conformation « ouverte », ce qui maintient une hyperpolarisation membranaire et empêche donc la sécrétion d'insuline en réponse à une hyperglycémie<sup>222</sup>.

Les patients atteints de ces mutations sont généralement plus petits à la naissance, et développent leur diabète durant leurs six premiers mois de vie<sup>222</sup>. De plus, étant donné que les canaux  $K^+_{ATP}$  sont également exprimés dans le cerveau, des mutations du gène *KCNJ11* peuvent provoquer des troubles neuroaux<sup>240</sup>. Certains individus atteints d'un diabète néonatal causé par des mutations de ce gène présentent également des troubles tels qu'un trouble du déficit de l'attention avec hyperactivité (TDAH), un retard du développement, ou encore des crises d'épilepsie avec une sévérité variable, pouvant aller jusqu'au syndrome de DEND regroupant tous ces troubles<sup>222</sup>.

Les patients porteurs de mutations dans les gènes *KCNJ11* et *ABCC8* initialement traités avec de l'insuline peuvent généralement basculer vers un traitement aux sulfonylurées à haute dose, qui permet de restaurer la fermeture du canal  $K^+_{ATP}$  et ainsi la sécrétion d'insuline<sup>222</sup>. Les sulfonylurées pourraient également améliorer les troubles neurologiques<sup>241</sup>.

#### 7.2.2.3.3 Gène *INS*

Les mutations du gène *INS*, codant la préproinsuline, sont la deuxième cause la plus courante de diabète néonatal permanent<sup>242</sup>. Les mutations de ce gène impliquées dans le diabète néonatal sont hétérozygotes dominantes ou homozygotes récessives<sup>221</sup>. Les mutations dominantes sont situées dans des régions critiques de la préproinsuline, et semblent avoir un impact sur son folding<sup>243</sup>. La protéine présente alors une conformation

Tableau 2. Tableau reprenant les différents gènes associés au diabète néonatal transitoire et/ou permanent, selon l'anomalie qui y est associée. Le rôle de ces gènes, l'hérabilité de leurs mutations, ainsi que les caractéristiques cliniques associées y sont également repris. Adapté de Rubio-Cabezas et al., « Diabetes Mellitus in Neonates and Infants: Genetic Heterogeneity, Clinical Approach to Diagnosis, and Therapeutic Options », *Hormone Research in Paediatrics*, 2013.

Gène	Rôle	Hérabilité	Caractéristiques cliniques
<b>Gènes associés à une anomalie du développement pancréatique</b>			
<b>ZAC/HYMAI</b>	ZAC : facteur de transcription HYMAI : long ARN non-codant	Variable (soumis à empreinte parentale)	Diabète néonatal transitoire, macroglossie et hernie ombilicale possibles
<b>ZFP57</b>	Facteur de transcription	Récessif	Diabète néonatal transitoire ± macroglossie ± retard du développement ± défauts ombilicaux ± cardiopathie congénitale
<b>PDX1</b>	Facteur de transcription	Récessif	Diabète néonatal permanent + agénésie pancréatique et stéatorrhée
<b>PTF1A</b>	Sous-unité α du facteur de transcription 1 du pancréas (PTF1A)	Récessif	Diabète néonatal permanent + agénésie pancréatique et stéatorrhée + hypoplasie ou aplasie cérébelleuse + dysfonction du centre respiratoire
<b>HNF1B</b>	Facteur de transcription	Dominant	Diabète néonatal transitoire + hypoplasie pancréatique et kystes rénaux
<b>RFX6</b>	Facteur de transcription	Récessif	Diabète néonatal permanent + atrésie intestinale + agénésie de la vésicule biliaire
<b>GATA6</b>	Facteur de transcription	Dominant	Diabète néonatal permanent + cardiopathie congénitale + troubles biliaires
<b>GLIS3</b>	Facteur de transcription	Récessif	Diabète néonatal permanent + hypothyroïdisme congénital + glaucome + fibrose hépatique + kystes rénaux
<b>NEUROG3</b>	Facteur de transcription	Récessif	Diabète néonatal permanent + anendocrinose entérique
<b>NEUROD1</b>	Facteur de transcription	Récessif	Diabète néonatal permanent + hypoplasie cérébelleuse + troubles visuels + surdité
<b>PAX6</b>	Facteur de transcription	Récessif	Diabète néonatal permanent + microphthalmie + malformations cérébrales
<b>Gènes associés à une fonction cellulaire β pancréatique anormale</b>			
<b>KCNJ11</b>	Sous-unité Kir6.2 du canal K <sup>+</sup> <sub>ATP</sub>	Apparition <i>de novo</i> ou dominant	Diabète néonatal transitoire ou permanent ± syndrome de DEND
<b>ABCC8</b>	Sous-unité SUR1 du canal K <sup>+</sup> <sub>ATP</sub>	Apparition <i>de novo</i> ou dominant ou récessif	Diabète néonatal transitoire ou permanent ± syndrome de DEND
<b>INS</b>	Préproinsuline	Récessif	Diabète néonatal transitoire ou permanent
<b>GCK</b>	glucokinase	Récessif	Diabète néonatal permanent

<b>SLC2A2</b>	Transporteur de glucose GLUT2	Récessif	Syndrome de Fanconi-Bickel, composé d'un diabète néonatal permanent, une hypergalactosémie et une dysfonction hépatique
<b>SLC19A2</b>	Transporteur de thiamine 1	Récessif	Syndrome de Roger, composé d'un diabète néonatal permanent, d'anémie mégaloblastique thiamine-dépendante, et de surdité
<b>Gènes associés à une destruction des cellules <math>\beta</math> pancréatiques</b>			
<b>INS</b>	Préproinsuline	Apparition <i>de novo</i> ou dominant	Diabète néonatal permanent
<b>EIF2AK3</b>	Protéine kinase PERK induite par le stress du réticulum endoplasmique	Récessif	Syndrome de Wolcott-Rallison, composé d'un diabète néonatal permanent, une dysplasie squelettique, ainsi qu'une dysfonction hépatique récurrente
<b>IER3IP1</b>	Impliqué dans la fonction du réticulum endoplasmique	Récessif	Diabète néonatal permanent + microcéphalie, lissencéphalie, encéphalopathie épileptique
<b>FOXP3</b>	Facteur de transcription	Lié au chromosome X	Syndrome IPEX, composé d'une entéropathie autoimmune, d'eczéma, d'un hypothyroïdisme autoimmun et de niveaux élevés d'IgE
<b>WFS1</b>	Canal ionique	Récessif	Syndrome de Wolfram, composé d'un diabète néonatal permanent + atrophie optique $\pm$ diabète insipide $\pm$ surdité

anormale et est de ce fait retenue dans le réticulum endoplasmique, générant un stress de cet organe. En réponse à ce stress, la cellule génère une réponse cellulaire appelée « unfolded protein response », qui, si elle n'est pas efficace, mène à l'apoptose des cellules  $\beta$  pancréatiques<sup>242,243</sup>. Les mutations récessives peuvent quant à elles avoir plusieurs effets : génération d'une protéine tronquée, mutations du promoteur altérant son activité, absence de site de début de transcription, ou encore altération de la stabilité de l'ARNm, menant toutes à une diminution de la synthèse de l'insuline<sup>244</sup>.

Les patients porteurs de mutations hétérozygotes du gène *INS* sont atteints de diabète néonatal permanent<sup>221</sup>. Ils ont, de manière similaire aux patients porteurs de mutations des gènes *KCNJ11* et *ABCC8*, un retard de croissance à la naissance<sup>221</sup>. Leur diabète a tendance à se développer légèrement plus tard (11 semaines) que les patients souffrant de mutations dans le gène *KCNJ11* (8 semaines)<sup>242</sup>. Concernant les porteurs de mutations homozygotes du gène *INS*, ceux-ci peuvent développer un diabète néonatal transitoire ou permanent<sup>221</sup>. ceux-ci ont un poids plus faible à la naissance que les porteurs de mutations hétérozygotes, et leur diabète se développe plus tôt (1 semaine)<sup>244</sup>.

Ces patients sont généralement traités avec de l'insuline<sup>242</sup>.

### 7.2.3 Importance du diagnostic génétique du diabète monogénique

Le diagnostic génétique du diabète monogénique est important pour la prise en charge des patients. L'arrivée du séquençage NGS a été un élément majeur dans le diagnostic génétique de cette maladie. Tout d'abord, en plus d'avoir confirmé l'association déjà existante entre certains gènes et le développement du diabète monogénique, cette technologie a également permis l'identification de nouveaux gènes associés à la maladie<sup>245</sup>, permettant ainsi de mettre un nom sur le gène déficient chez des patients chez lesquels la cause du MODY était encore inconnue. D'autre part, la technologie NGS permet également une bonne rentabilité du diagnostic génétique : tandis que le séquençage Sanger est cher (entre 500 et 2000 dollars par gène étudié) et peu efficace dans le diagnostic des maladies monogéniques, la technologie NGS permet de tester en parallèle un large panel de gènes identifiés comme étant impliqués dans les formes monogéniques de la maladie pour moins de 1000 dollars, réduisant fortement les coûts du diagnostic<sup>171</sup>. Le diagnostic génétique permet également de réduire les frais de santé au long terme grâce à la médecine de précision, notamment en adaptant le traitement

Tableau 3. Critères 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
Très forte preuve de pathogénicité (very strong)	PVS1	Variant nul (variant non-sens, frameshift, 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
Forte preuve de pathogénicité (strong)	PS1	Variant induisant un changement d'acide aminé identique à un autre, déjà établi pour un autre variant pathogénique, peu importe le changement de nucléotide
	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
Preuve modérée de pathogénicité (moderate)	PM1	Variant situé dans un point chaud (hot spot) de mutations, ou dans un domaine fonctionnel critique ne contenant pas de variation bénigne
	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, ou ExAC
	PM3	Pour les maladies récessives, variant détecté en trans 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 à 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
	PM6	Variant supposé apparu <i>de novo</i> , sans confirmation paternelle et maternelle
Preuve secondaire de pathogénicité (supporting)	PP1	Coségrégation avec la maladie chez plusieurs membres de la famille atteints, 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
	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

des patients selon le gène atteint, évitant ainsi l'administration de traitements inefficaces, ainsi qu'en prévenant l'apparition de complications associées à la maladie et de comorbidités<sup>171</sup>.

Lorsque les variants identifiés par séquençage de l'ADN du patient sont situés dans des gènes connus pour être impliqués dans des maladies génétiques rares, comme le diabète monogénique, leur pathogénicité peut être évaluée à l'aide des critères établis par l'American College of Medical Genetics and Genomics (ACMG)<sup>246</sup>. Il s'agit d'une liste de critères soutenant un effet pathogène très fort, fort, modéré, ou secondaire des variants sur base de données épidémiologiques, informatiques, ou encore fonctionnelles, reprises dans le tableau 3. La combinaison de ces critères, suivant certaines règles reprises dans le tableau 4, permet de définir la pathogénicité ou non d'un variant. Ainsi, il est possible de déterminer si un variant identifié nouvellement a un impact sur la fonction du gène. La caractérisation des variants identifiés chez les patients par séquençage permet ainsi d'affiner le diagnostic génétique.

Tableau 4. 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	Combinaisons de critères possibles
Pathogène	1 critère de pathogénicité très fort (PVS1) et : <ul style="list-style-type: none"> <li>- ≥ 1 critère fort (PS1-PS4) ou</li> <li>- ≥ 2 critères modérés (PM1-PM6) ou</li> <li>- 1 critère modéré (PM1-PM6) et 1 critère secondaire (PP1-PP5) ou</li> <li>- ≥ 2 critères secondaires (PP1-PP5)</li> </ul>
	≥ 2 critères de pathogénicité forts (PS1-PSS4)
	1 critère de pathogénicité fort (PS1-PS4) et : <ul style="list-style-type: none"> <li>- ≥ 3 critères modérés (PM1-PM6) ou</li> <li>- 2 critères modérés (PM1-PM6) et ≥ 2 critères secondaires (PP1-PP5) ou</li> <li>- 1 critère modéré (PM1-PM6) et ≥ 4 critères secondaires (PP1-PP5)</li> </ul>
Probablement pathogène	1 critère très fort (PVS1) et 1 critère modéré (PM1-PM6)
	1 critère fort (PS1-PS4) et 1- 2 critères modérés (PM1-PM6)
	1 critère fort (PS1-PS4) et ≥ 2 critères secondaires (PP1-PP5)
	≥ 3 critères modérés (PM1-PM6)
	2 critères modérés (PM1-PM6) et ≥ 2 critères secondaires (PP16-PP5)
1 critère modéré (PM1-PM6) et ≥ 4 critères secondaires (PP1-PP5)	

Le diagnostic génétique est déjà bien appliqué dans le cadre du diabète néonatal. En effet, en cas d'apparition d'un diabète chez des enfants de moins de 6 mois, un test génétique peut être réalisé afin de déterminer de quel type de diabète néonatal l'enfant est atteint. Le diagnostic génétique du diabète néonatal présente de nombreux avantages : il permet d'adapter le traitement de l'enfant, d'expliquer certaines caractéristiques cliniques pouvant être associées au diabète, d'anticiper les futures complications cliniques qui pourront apparaître, et également d'agir rapidement sur les comorbidités possibles<sup>171</sup> (figure 25).

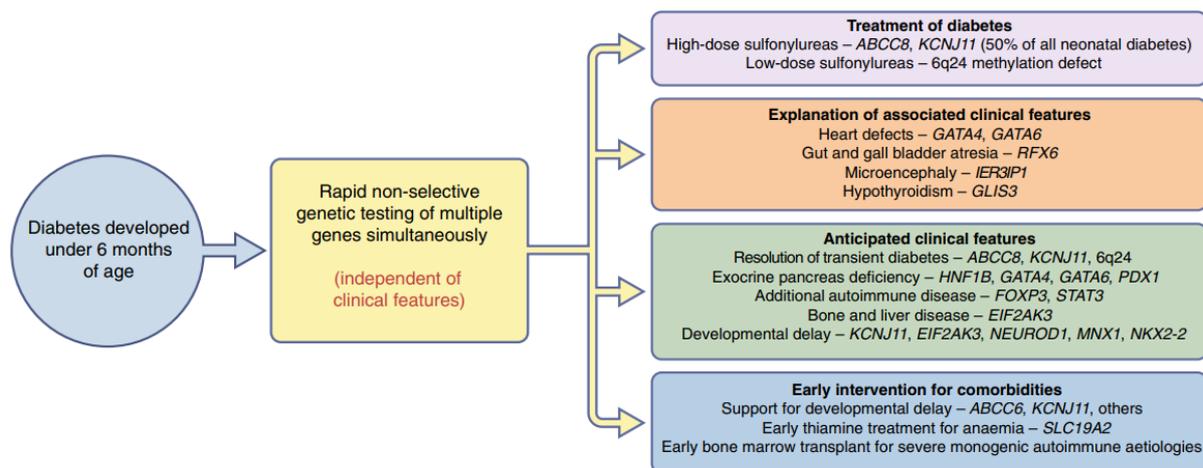


Figure 25. Bénéfices du diagnostic génétique chez les enfants présentant un diabète moins de 6 mois après la naissance. Source : Hattersley and Patel, « Precision diabetes : learning from monogenic diabetes », Diabetologia, 2017.

Cependant, le diagnostic génétique dans le cadre du diabète MODY est moins courant. En effet, la médecine de précision est un concept relativement nouveau qui n'est pas encore ancré en routine. Ce diagnostic génétique est préconisé notamment suivant le calcul de la probabilité d'être atteint de MODY, basé sur les caractéristiques cliniques présentés par le patient, le traitement suivi, ainsi que la présence d'autres cas de diabète dans la famille (figure 26). Cependant, ces critères ne sont pas forcément retrouvés chez tous les patients qui en sont atteints, et d'autre part, plusieurs caractéristiques cliniques du diabète MODY, du DT1 et du DT2 commun se chevauchent (IMC, histoire familiale de diabète, âge d'apparition de la maladie ou encore taux de HbA1c) entraînant de ce fait une confusion. Ainsi, seule la moitié des individus atteints d'un MODY sont correctement diagnostiqués, ce qui réduit le nombre d'individus enclins à être testés pour définir la cause génétique de leur MODY<sup>171</sup>. Néanmoins, le diagnostic génétique du diabète monogénique présente un grand intérêt, étant donné que comme expliqué précédemment, le diagnostic génétique de la maladie permet d'adapter le traitement du patient (figure 26).

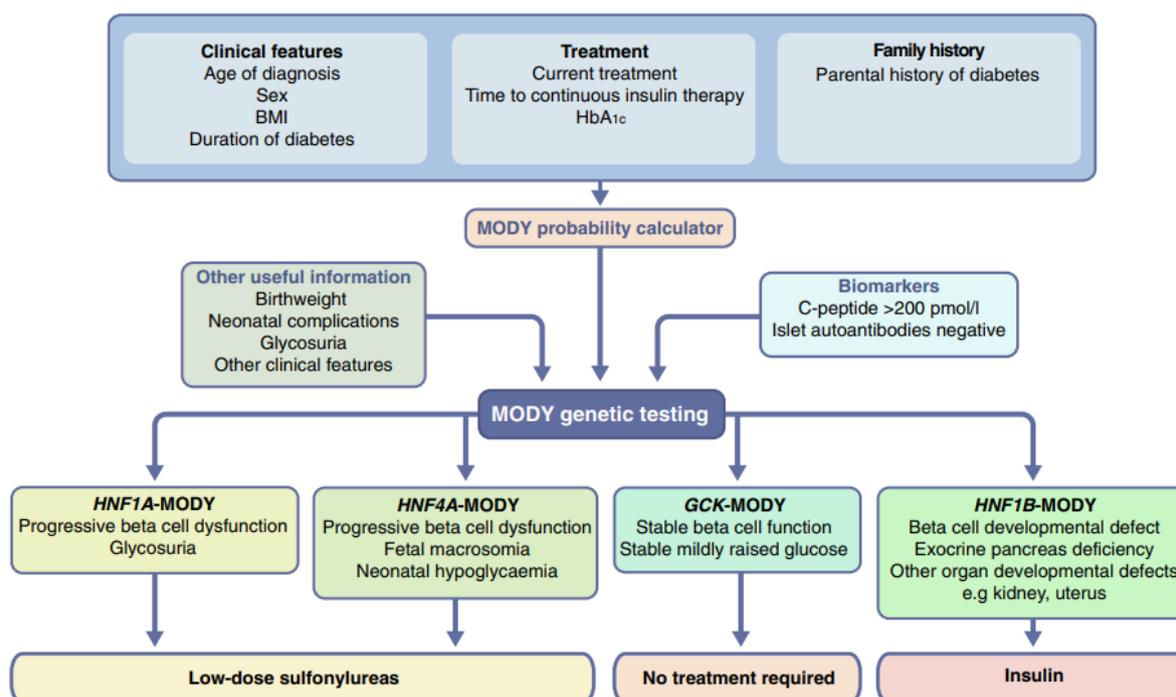


Figure 26. Diagnostic génétique du diabète de type MODY et adaptation du traitement du diabète en conséquence. Source : Hattersley and Patel, « Precision diabetes : learning from monogenic diabetes », *Diabetologia*, 2017.

## 7.3 Les formes polygéniques du diabète de type 2

### 7.3.1 Continuum entre les formes monogéniques et polygéniques

Les formes polygéniques du DT2 correspondent à la forme commune de cette pathologie, représentant environ 90% des cas. Il s'agit d'une maladie polygénique multifactorielle complexe, qui est dans un continuum avec le diabète monogénique. Pour rappel, dans le cas du diabète monogénique, des mutations rares mais présentant un effet fort sur le risque de développer la maladie, et situées dans un seul gène jouant un rôle majeur dans la fonction ou le développement des cellules  $\beta$  pancréatiques sont suffisantes pour provoquer le développement de la maladie (figure 27). Cependant, contrairement aux variants rares impliqués dans le DT2 monogénique, les variants identifiés comme étant associés aux formes communes du DT2 sont principalement des variants fréquents, situés dans des régions non-codantes, et ayant un faible effet sur le risque de développer la maladie (figure 28). Ces variants fréquents ne peuvent donc pas provoquer la maladie à eux seuls : ainsi, c'est la présence de ce type de mutations dans plusieurs gènes, couplée à la présence de facteurs environnementaux (tels que l'alimentation ou la sédentarité), qui vont mener à l'apparition de la maladie<sup>150</sup>.

L'héritabilité du DT2 commun a été estimée à 25-70%, en fonction de l'âge auquel la maladie est apparue et du seuil de glycémie pris en compte<sup>247-249</sup>. En 2015, le consortium DISCOTWIN, regroupant presque 35 000 paires de jumeaux discordants pour le DT2 et permettant ainsi d'étudier l'implication de la génétique dans l'étiologie du DT2, a estimé que l'héritabilité de la maladie pourrait même atteindre 75%<sup>249</sup>. Cependant, les différents variants, majoritairement fréquents, identifiés jusqu'ici n'expliquent qu'une faible proportion de cette héritabilité. Il a donc été supposé, en lien avec le continuum entre les formes monogéniques et polygéniques de la pathologie, que des variants rares (MAF inférieure à 1%) et de faible fréquence (MAF inférieure à 5%) situés dans les régions codantes des gènes, impactant ainsi directement leur fonction, pourraient être associés au développement des formes polygéniques du DT2 avec un effet plus fort que les variants fréquents (figure 27), et ainsi expliquer en partie la part d'héritabilité encore inconnue<sup>150,250</sup>.

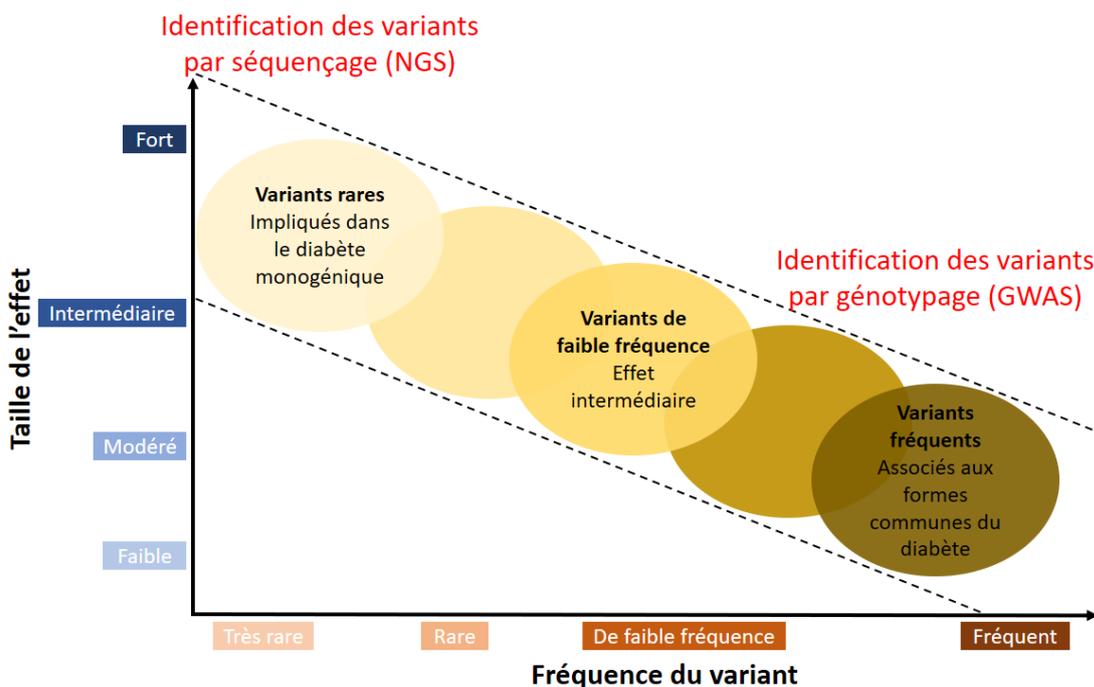


Figure 27. Fréquence et effets des variants associés au DT2. Source : adapté à partir de Manolio et al., « Finding the missing heritability of complex diseases », Nature, 2009.

### 7.3.2 Identification de variants fréquents associés au DT2 commun

#### 7.3.2.1 Identification de gènes associés au DT2 commun par études de liaison familiale

Afin d'identifier les gènes et les variants associés aux formes polygéniques du DT2, plusieurs études basées sur les outils génétiques décrits précédemment ont été réalisées. Notamment, plus de 20 études d'association familiales basées principalement sur le

génomotypage ont été réalisées à cette fin. Celles-ci ont permis d'identifier certains gènes comme étant impliqués dans le développement du DT2 commun, repris ci-dessous.

- En 1996, une étude d'association familiale, réalisée chez 408 individus atteints ou non de DT2 issus de 170 familles mexicano-américaines, a réalisé un génotypage de 490 marqueurs génétiques. Le génotypage de 18 de ces marqueurs a ensuite été répliqué dans un second groupe de 110 individus de la même origine. Ainsi, un locus de susceptibilité majeur a été identifié dans le bras long du chromosome 2<sup>251</sup>, où se situe le gène *CAPN10*, codant la cystéine protéase calpaïne 10<sup>375</sup>.
- En 2002, une étude menée par notre laboratoire a réalisé un séquençage du gène *ADIPOQ* et de sa région promotrice proximale, étant donné que la protéine adiponectine, codée par ce gène, est connue pour moduler la sensibilité à l'insuline et l'homéostasie glucidique. 12 SNPs, et 4 mutations rares non-synonymes situés dans l'exon 3 de ce gène ont été identifiés. Les 10 SNPs les plus fréquents, ainsi que les 4 mutations non-synonymes ont ensuite été génotypés chez 1373 et 1246 individus causasiens atteints de DT2 ou d'obésité, respectivement, ainsi que chez 743 membres issus de 148 familles atteints de DT2, permettant d'identifier plusieurs de ces mutations comme étant associées au DT2<sup>252</sup>.
- En 2004, une étude finlandaise s'est intéressée à la région chromosomique 20q13, dont l'association avec le DT2 avait été déjà préalablement démontrée par plusieurs études<sup>253</sup>. Par génotypage de 291 marqueurs génétiques dans une étude cas-contrôle composée d'un peu moins de 800 individus danois malades issus de familles présentant une histoire familiale de diabète, ainsi que d'environ 400 individus contrôles, cette étude a identifié un SNP situé dans la région promotrice du gène *HNF4A* associé au DT2 et aux traits métaboliques associés, ainsi que 9 autres SNPs présents dans la région promotrice et les exons ce gène, suggérant que ce gène est associé au DT2<sup>254</sup>.
- En 2005, une étude menée par notre unité de recherche et basée sur le génotypage de plusieurs SNPs du gène *ENPP1*, codant la protéine ectonucleotide pyrophosphatase/phosphodiesterase 1, a permis d'identifier plusieurs SNPs de ce gène chez des individus obèses. Parmi ceux-ci, le SNP K121Q, identifié dans une cohorte de 2430 individus (comprenant 1225 individus obèses et 1205 individus sains

normoglycémiques) et répliqué dans 184 familles atteintes d'obésité, ainsi que les SNPs IVS20delT-11 et A/G+1044TGA, identifiés dans une cohorte composée de 719 individus (comprenant 421 individus obèses et 298 individus contrôles), capables de moduler l'effet du variant K121Q ont été identifiés. L'haplotype composé de ces 3 SNPs a été démontré comme étant associé au DT2 et à l'obésité<sup>255</sup>.

- En 2006, une étude Islandaise s'est intéressée au lien entre le DT2 et le bras long du chromosome 10, étant donné qu'une étude précédente avait suggéré un lien entre ce locus et le DT2. Un génotypage de 228 séquences microsatellites a été réalisé dans cette région chromosomique, dans une cohorte islandaise composée de 1185 individus atteints de DT2 et 931 individus contrôles. Cette étude a ainsi identifié un microsatellite dans l'intron 3 du gène *TCF7L2*, codant le facteur de transcription 7-like 2, comme étant associé au DT2. Cette association a ensuite été confirmée par réplification dans des cohortes danoises et américaines<sup>256</sup>.

Bien que les analyses de liaison familiales aient permis de démontrer une association entre quelques variants fréquents de ces gènes et le risque de DT2 polygénique, elle n'ont pas permis de découverte majeure dans le cadre du diabète polygénique : les variants du gène *TCF7L2* sont ceux qui augmentent le plus fortement ce risque, avec un odds ratio (OR) égal à 1,50, signifiant que les individus porteurs de ces variants ont un risque 1,5 fois plus important de développer la maladie<sup>256</sup>. De plus, seules les associations entre le DT2 et les variants des gènes *HNF4A* et *TCF7L2* ont pu être répliquées par GWAS<sup>257,258</sup>, mettant ainsi en doute l'implication des autres gènes identifiés par études d'association familiales dans le développement de la maladie.

#### 7.3.2.2 Identification de gènes associés au DT2 commun par approche gène-candidat

En parallèle de l'approche basée sur les études d'association familiales, d'autres études se sont focalisées sur l'investigation de gènes-candidats, et ont permis d'identifier plusieurs variants fréquents associés à la forme commune de la maladie<sup>150</sup>.

Plusieurs de ces études se sont intéressées à l'étude de gènes dont l'implication était déjà connue dans le diabète monogénique. En effet, alors que des mutations rares de ces gènes, présentant un effet fort sont impliquées dans les formes monogéniques, il a été supposé que des mutations plus fréquentes, avec un effet plus faible, pourrait être associées aux formes polygéniques. Ces études basées sur l'approche gène-candidat ont ainsi permis

d'identifier, par génotypage des gènes *GCK*, *HNF1B*, et *WFS1*, déjà connus pour être impliqués dans le développement des *MODY2*, *MODY5*, et du syndrome de Wolfram respectivement, plusieurs SNPs associés au DT2 commun<sup>259-261</sup>.

D'autres études se sont intéressées quant à elles à des gènes sur base de leur implication dans la fonction de la cellule  $\beta$  pancréatique et l'insulinorésistance<sup>151</sup>.

- Le séquençage du gène *KCNJ11*, codant la sous-unité Kir6.2 des canaux  $K^+_{ATP}$  et jouant un rôle prépondérant dans la sécrétion d'insuline, chez 72 familles caucasiennes atteintes de DT2, suivi du génotypage des différents variants chez 191 individus diabétiques et 119 individus normoglycémiques a démontré que ce gène est associé au développement du DT2 commun<sup>262</sup>. Il est à noter que des mutations de ce gène ont ensuite été démontrées comme étant impliquées dans le diabète néonatal transitoire et permanent<sup>238</sup> et le *MODY13*<sup>263</sup>.
- Le gène *PPARG*, codant le récepteur PPAR $\gamma$  impliqué dans la différenciation et la fonction adipocytaires et également impliqué dans la résistance à l'insuline, et ciblé par les des TZD, a également été considéré comme un gène-candidat. Une étude a démontré, par génotypage des exons codants de ce gène, que 2 mutations à effet dominant négatif de celui-ci sont associées à une insulinorésistance, ainsi qu'à un DT2 et une hypertension<sup>264</sup>. Le génotypage d'un variant fréquent situé dans la région codante de ce gène a également démontré son association avec le DT2<sup>150,265</sup>.
- D'autre part, des études se sont intéressées au gène *IRS1* en tant que gène-candidat, étant donné son rôle majeur dans la transduction du signal de l'insuline dans les tissus cibles. Le génotypage de ce gène suivi de son séquençage chez 86 individus atteints de DT2 et 76 contrôles normoglycémiques a permis d'identifier plusieurs polymorphismes de ce gène associés au DT2 commun<sup>391</sup>.

Cependant, les études basées sur l'approche gène-candidat n'ont pas permis la découverte de nombreux gènes associés à la maladie : étonnamment, de nombreux gènes étudiés suite à leur implication dans la sécrétion et l'action de l'insuline n'ont pas démontré d'association avec le DT2<sup>152</sup>.

### 7.3.2.3 Identification de SNPs associés au DT2 commun par études GWAS

L'arrivée des études GWAS a permis une grande avancée dans la compréhension de la génétique du DT2 commun. La première vague d'étude GWAS a consisté en des études cas-contrôles du DT2, par génotypage de plusieurs milliers de SNPs chez des individus atteints de la maladie ou non. Une de ces premières études a été réalisée par notre laboratoire. Celle-ci a consisté en une étude cas-contrôles, dans laquelle 392 935 SNPs différents ont été génotypés dans une cohorte composée de 1363 individus d'origine européenne, comprenant 661 individus atteints de DT2 et 614 individus normoglycémiques. 57 SNPs identifiés comme associés au DT2 dans cette première cohorte ont ensuite été génotypés dans une seconde cohorte, composée de 2617 individus atteints de DT2 et 2894 individus contrôles. Cette étude a ainsi permis de confirmer l'association entre le gène *TCF7L2* et la maladie, et d'identifier de nouveaux gènes de susceptibilité : *SLC30A8*, *HHEX*, ainsi que le locus *EXT2-ALX4* qui n'a cependant pas été répliqué<sup>257</sup>. Cette étude a été suivie de plusieurs autres GWAS ainsi que de méta-analyses GWAS, combinant l'étude de plusieurs cohortes simultanément de manière à augmenter la puissance statistique de l'étude et ainsi améliorer l'identification de nouveaux gènes de susceptibilité<sup>150</sup>. Ensemble, ces différentes études ont ainsi permis d'identifier les gènes de susceptibilité *ADAMTS9*, *CDC123/CAMK1D*, *CDKAL1*, *CDKN2A/B*, *IGF2BP2*, *JAZF1*, *NOTCH2*, *RBMS1*, *THADA*, et *TSPAN8/LGR5*<sup>266-270</sup>. D'autre part, plusieurs études GWAS ont identifié le gène *FTO* comme étant associé au DT2. Cependant, cette association disparaît suite à un ajustement au BMI : son association est donc indirecte et médiée par l'adiposité<sup>271-273</sup>. Une seconde vague d'études GWAS, basée sur l'analyse de traits métaboliques associés au DT2 tels que la glycémie à jeun, l'insulinémie à jeun, ou encore la triglycéridémie, ont par la suite été réalisées chez des individus non-diabétiques<sup>150</sup>. Ces études ont permis d'identifier les gènes *MTNR1B*, *DGKB*, *PROX1*, *ADCY5* et *GCKR* comme étant associés à la pathologie<sup>266,274-276</sup>. Cependant, les gènes associés au DT2 identifiés dans les études cas-contrôle, et ceux associés à des traits métaboliques liés au DT2 chez des individus non-diabétiques se recouvrent peu : les gènes associés à la régulation des traits métaboliques pourraient donc être différents de ceux impliqués dans la physiopathologie de la maladie<sup>150,275</sup>. D'autres études GWAS ont par la suite été réalisées, dans plusieurs ethnies différentes, et ont permis d'identifier de nombreux loci génétiques associés au DT2 ou aux traits métaboliques associés<sup>150</sup>. Ainsi, ces études GWAS ont permis d'identifier 240 loci génétiques associés au DT2.

Cependant, contrairement aux variants rares impliqués dans le DT2 monogénique, capables de provoquer à eux seuls la maladie, et dont le diagnostic génétique permet d'identifier le gène atteint et d'adapter le traitement du patient, les différents SNPs associés au DT2 polygénique sont généralement situés dans des régions non-codantes et ont un faible effet sur le risque de développer la maladie. Ces SNPs ne permettent donc pas de prédire l'apparition ou l'évolution de la maladie chez les patients, et ne peuvent donc pas contribuer à l'élaboration d'une médecine de précision vis-à-vis du DT2 commun.

### 7.3.3 Contribution de variants rares à la forme commune du DT2

Le séquençage NGS, déjà largement utilisé pour découvrir de nouveaux variants associés au diabète monogénique et également réalisé dans le cadre du dépistage génétique de celui-ci, peut également être utilisé afin d'investiguer l'étiologie génétique du DT2 polygénique par identification de variants rares et de faible fréquence associés à celui-ci.

Plusieurs études, basées sur le séquençage NGS, ont d'ores et déjà permis d'identifier de nouveaux gènes associés au DT2. Par exemple, en 2016, une méta-analyse réalisée à partir de données de séquençage WES issues de 6504 cas et 6436 contrôles a permis de démontrer une association entre l'agrégation de tous les variants rares situés dans 12 gènes liés au MODY et le risque de DT2 (même si les analyses gène par gène n'étaient pas significatives)<sup>277</sup>. Une autre méta-analyse basée quant à elle sur des données de WES obtenues chez 20 791 cas et 24 440 contrôles a confirmé l'association entre l'agrégation de variants rares délétères situés dans tous les gènes du diabète monogénique et le risque de DT2<sup>278</sup>.

Une étude réalisée en 2020 par notre laboratoire a permis d'établir une association entre plusieurs gènes actionnables impliqués dans le MODY, qui peuvent être utilisés comme cibles thérapeutiques, et le DT2 grâce au séquençage NGS. Le séquençage de 33 gènes associés au MODY chez 6348 individus issus de l'étude RaDiO, comprenant 2178 individus atteints de DT2 commun et 4170 individus normoglycémiques, a permis d'identifier 1408 variants rares situés dans les régions codantes de ceux-ci, avec un enrichissement chez les cas (53,8% de porteurs chez les cas *versus* 45,7% de porteurs chez les contrôles)<sup>279</sup>. Par application des critères ACMG, 167 variants ont été classés comme étant pathogènes ou probablement pathogènes (P/LP), avec un enrichissement chez les cas. Plus particulièrement, les gènes *GCK* et *HNF1A*, c'est-à-dire les gènes les plus fréquemment mutés dans le MODY, sont associés à un risque important de développer un DT2 (OR = 7,08 et 7,40, respectivement). Cette étude a

ensuite été répliquée pour les gènes actionnables grâce à des données de séquençage WES fournies dans la UK Biobank (comprenant 2151 individus atteints de DT2 et 33 718 individus contrôles normoglycémiques), le Healthy Nevada Project (comprenant 1518 individus atteints de DT2 et 4171 individus contrôles normoglycémiques), ainsi que dans le Accelerating Medicines Partnership (AMP) T2D Knowledge Portal (comprenant 19 852 individus atteints de DT2 et 6871 individus contrôles). Ainsi, 227 variants P/LP additionnels ont été identifiés. La méta-analyse de ces 4 cohortes, combinant ainsi 25 699 cas et 48 930 contrôles, a mis en évidence un enrichissement en variants P/LP dans les gènes actionnables du MODY chez les cas, probablement influencé par les variants présents dans les gènes *GCK* (OR = 4,52), *HNF4A* (OR = 2,09), *KCNJ11* (OR = 2,38) et *ABCC8* (OR = 1,35)<sup>279</sup>.

Ces différentes études ont ainsi mis en évidence l'importance du séquençage NGS dans l'identification de variants rares associés aux formes polygéniques du DT2. L'identification de ce type de variant pourrait contribuer à une médecine de précision chez ces patients : étant donné que ces variants ont un effet fort sur le risque de développer la maladie, le traitement des patients porteurs de ces variants pourrait être adapté en fonction des gènes atteints, permettant ainsi une meilleure prise en charge de ceux-ci ainsi qu'une réduction des coûts des soins de santé, comme décrit précédemment.

#### 7.3.4 Combinaison du séquençage NGS et de la génétique fonctionnelle

L'investigation de l'effet fonctionnel des différents variants associés au DT2 est un élément important pour la compréhension de l'effet de ces variants sur l'étiologie génétique de la maladie. Cette approche est très peu utilisée concernant l'étude des variants fréquents : en effet, étant donné qu'ils ont un faible effet sur le risque de développer la maladie et qu'ils ne sont pas situés dans des régions codantes, leur effet fonctionnel est difficile à évaluer dans des modèles cellulaires ou animaux<sup>150</sup>. Certaines études ont tout de même réussi cet exploit, en démontrant que des variants associés au DT2 et à la glycémie à jeun situés dans des régions enhancer de certains gènes sont capables d'y altérer la liaison de facteurs de transcription, réduisant ainsi l'activité enhancer<sup>280</sup>. Cependant, cette approche est très intéressante dans le cadre de l'étude des variants rares. Comme expliqué précédemment, ces variants sont généralement situés dans des régions codantes, et il est peu probable qu'ils soient associés à d'autres variants perte de fonction (loss-of-function [LoF]) dans des régions en déséquilibre de liaison, comme c'est le cas pour les variants fréquents : ainsi, leur effet impacte

directement le gène dans lequel ils se situent, et implique donc directement ce gène dans la pathogenèse du DT2. Ces variants peuvent être facilement étudiés dans des modèles *in vitro* ou *in vivo*, permettant ainsi de mieux comprendre quel effet exerce le gène sur divers processus biologiques, et par quels mécanismes ces variants peuvent impacter son activité. Ainsi, la génétique fonctionnelle est un outil important, permettant de mieux comprendre les causes génétiques du DT2<sup>150,250</sup>. Ci-dessous sont reprises quelques études illustrant l'importance des études fonctionnelles de variants rares dans le cadre du DT2 commun.

#### 7.3.4.1 Le gène *MTNR1B*

Le gène *MTNR1B* code le récepteur 2 à la mélatonine (MT<sub>2</sub>), un GPCR. L'activation des récepteurs MT1 et MT2 par la neurohormone mélatonine, sécrétée par la glande pinéale suivant un rythme circadien, entraîne l'activation de protéines G<sub>i</sub> modulant ainsi l'activité de l'adénylate cyclase et de la voie de signalisation ERK1/2 en aval<sup>281,282</sup>. Des études précédentes avaient déjà pu démontrer par GWAS l'association de SNPs fréquents non-codants du gène *MTNR1B* avec le DT2 et la glycémie à jeun<sup>274,276</sup>. Par un reséquençage d'exons de *MTNR1B* chez 7932 individus d'origine européenne comprenant 2186 cas, 40 variants non-synonymes, parmi lesquels deux variants fréquents, deux variants de faible fréquence et 36 variants rares du gène ont pu être identifiés. Trois agrégations de variants ont été formées selon leur fréquence, et les études d'association réalisées pour chaque agrégation ont montré une augmentation du risque de DT2 uniquement pour l'agrégation de variants très rares (OR = 3,31,  $P = 1,64 \times 10^{-4}$ )<sup>283</sup>.

La deuxième partie de cette étude a cherché à déterminer si ces variants non-synonymes du gène *MTNR1B* ont un effet sur la fonction du récepteur MT<sub>2</sub>. L'effet de ces 40 variants a été évalué par des expériences de génétique fonctionnelle dans un modèle cellulaire, notamment leur expression à la surface de la cellule, leur capacité à lier leur ligand, ainsi que leur capacité à moduler l'activité de la protéine G<sub>i</sub> et de la voie de signalisation ERK1/2. Ces expériences ont démontré qu'aucun de ces variants n'altère l'expression membranaire du récepteur, mais que certains induisent une diminution, voire une absence de liaison du ligand, et que d'autres, altérant ou non la liaison du ligand, induisent également une altération de l'activation de la protéine G<sub>i</sub> et de la voie de signalisation ERK1/2 en aval. Au total, 4 variants rares abolissent complètement la fonction du récepteur MT<sub>2</sub>, et 10 variants entraînent une altération partielle de celle-ci. Ces 4 variants ont également été génotypés

chez 11 854 autres individus d'origine française. Par étude d'association, il a été démontré que l'agrégation de ces 4 variants augmente significativement le risque de développer un DT2 (OR = 3,88,  $P=5,37 \times 10^{-3}$ ). L'agrégation des 10 variants ayant un effet partiel sur la fonction du récepteur montre une tendance à augmenter le risque de développer la maladie (OR = 2,66,  $P = 4,09 \times 10^{-4}$ ). Les variants neutres, quant à eux, ne présentent pas d'association avec la maladie<sup>283</sup>.

Cette étude permet également d'envisager que l'utilisation d'agonistes du récepteur MT<sub>2</sub> pourrait contribuer à l'amélioration du phénotype des patients diabétiques porteurs de ces mutations. Ainsi, on peut envisager l'utilisation future d'agonistes de ce récepteur comme traitement dans le cadre de la médecine de précision.

#### 7.3.4.2 Le gène *PPARG*

Le gène *PPARG* code le récepteur PPAR $\gamma$ , qui est impliqué dans la différenciation adipocytaire. Comme décrit précédemment, ce récepteur est la cible des TZD, permettant d'améliorer la sensibilité à l'insuline<sup>284</sup>. Des premières études avaient déjà pu démontrer l'association entre le SNP fréquent p.P12A de ce gène et une diminution du risque de développer un DT2<sup>265,285</sup>. Dans cette nouvelle étude, le séquençage des exons du gène chez 19 752 individus d'origine multiethnique, composé pour moitié de cas et pour moitié de contrôles, a permis d'identifier 53 variants non-synonymes de ce gène. Parmi ces variants, le variant fréquent p.P12A a également été retrouvé, et son association avec une diminution du risque de DT2 a été confirmée (OR = 0,85,  $P = 0,0006$ ). Les études d'association réalisées pour l'agrégation composée des 52 autres variants rares n'a par contre montré aucune association avec le risque de DT2<sup>284</sup>.

Les variants ont alors été classés selon qu'ils soient prédits comme étant bénins ou délétères selon des outils bioinformatiques. L'étude d'association réalisée sur l'agrégation de variants prédits comme étant délétères *in silico* n'a cependant pas montré d'association significative avec le DT2, démontrant encore une fois la faiblesse des prédictions informatiques<sup>284</sup>. L'effet des différents variants a donc été évalué par des expériences de génétique fonctionnelle, basées sur la capacité de ceux-ci à induire la différenciation adipocytaire dans un modèle cellulaire de préadipocytes humains suite à une stimulation avec l'agoniste de PPAR $\gamma$  rosiglitazone. Sur les 53 variants initialement identifiés, 41 d'entre eux induisent une différenciation adipocytaire à des niveaux similaires au récepteur wild-type

(WT), et ont donc été considérés comme neutres. Les 12 variants restants sont associés à une diminution de la différenciation adipocytaire, et ont donc été considérés comme variants LoF. Cependant, 4 d'entre eux retrouvent une activité similaire au récepteur WT en présence de concentration plus importante d'agoniste<sup>284</sup>. Une nouvelle étude d'association a ensuite été réalisée, basée cette fois-ci sur la classification des variants en tant que neutres ou LoF selon les expériences fonctionnelles. Alors que l'agrégation de variants neutres n'est pas associée au DT2 (OR = 1,17,  $P = 0,6$ ), une augmentation significative du risque de développer cette maladie a pu être démontrée chez les porteurs de mutations LoF (OR = 7,22,  $P = 0,005$ )<sup>284</sup>.

En plus de démontrer l'association entre le DT2 et le gène *PPARG*, cette étude suggère également que des individus atteints de mutations LoF partielles du gène *PPARG* pourraient être traités avec l'agoniste rosiglitazone, mais pas ceux porteurs de mutations LoF totales, qui ne sont plus capables de répondre à l'agoniste<sup>284</sup>.

#### 7.3.4.3 Le gène *HNF1A*

Comme décrit précédemment, le gène *HNF1A* code le facteur de transcription HNF-1 $\alpha$ . Par réalisation de WES dans une cohorte comprenant 1794 individus atteints de DT2 et 1962 contrôles d'origine latino, 1,2 millions de variants ont été identifiés. Parmi les variants rares identifiés dans cette étude, le variant faux-sens p.E508K, présent dans l'exon 8 du gène *HNF1A* et déjà connu comme étant associé au *MODY3*, a été identifié. Ce variant est enrichi chez les cas (OR = 5,48,  $P = 4,4 \times 10^{-7}$ ). La réplication de cette étude dans 2 autres cohortes a montré une association nominale de ce variant avec le DT2 (OR = 4,16,  $P = 0,0013$ ). Combinées ensemble, ces 3 études ont démontré que ce variant est significativement associé à une augmentation du risque de développer la maladie (OR = 4,96,  $P = 2,39 \times 10^{-9}$ )<sup>286</sup>.

Des expériences de génétique fonctionnelle ont ensuite été réalisées dans différents modèles cellulaires afin de déterminer si le variant rare p.E308K, situé dans le domaine de transactivation du facteur de transcription, affecte sa fonction. En comparaison à la protéine HNF-1 $\alpha$  WT, le variant p.E308K entraîne une diminution de l'activité transcriptionnelle. Des expériences s'intéressant à la localisation cellulaire de HNF-1 $\alpha$  ont montré qu'alors que le récepteur WT est surtout exprimé dans le noyau, le variant p.E308K est exprimé également dans le cytosol, suggérant ainsi une altération de sa translocation vers le noyau<sup>286</sup>.

Comme décrit précédemment, les patients atteints de MODY3 sont sensibles aux sulfonylurées, et leur traitement à l'insuline peut donc être remplacé par celles-ci. On peut donc supposer que les patients atteints de DT2 commun, et porteurs de mutations du gène *HNF1A*, pourraient voir leur traitement adapté.

# **MON TRAVAIL DE THESE**

## Mon travail de thèse

Le DT2 est une maladie métabolique chronique, provoquée à la fois par des facteurs environnementaux, mais également génétiques. Depuis les années 1990, plusieurs études, basées sur les approches d'association familiales et les approches gènes-candidats, ont permis d'identifier quelques gènes de susceptibilité associés aux formes communes du DT2. L'arrivée des GWAS en 2007 a permis l'identification d'un très grand nombre de SNPs généralement fréquents associés à la maladie. Cependant, ces études présentent quelques faiblesses. Tout d'abord, les SNPs identifiés par GWAS ne sont pas forcément directement impliqués dans l'étiologie de la maladie. Ceux-ci sont en effet généralement situés dans des régions non-codantes du génome, et sont nommés selon le gène le plus proche, sans pour autant avoir la certitude que ces variants aient un impact sur la fonction du gène qui y est associé. De plus, ces variants fréquents ont un faible effet sur le risque de la maladie, et d'autre part, ils n'expliquent qu'une faible proportion de l'héritabilité du DT2 : une grande proportion d'héritabilité manquante reste encore à investiguer. Il a donc été supposé que des variants rares pourraient avoir un plus fort effet sur le risque de développer la maladie, et ainsi expliquer, au moins en partie, cette héritabilité manquante.

L'arrivée du séquençage NGS a été une révolution concernant la compréhension de l'architecture génétique du DT2. La possibilité du séquençage d'exomes dans de larges cohortes permet l'identification de variants rares dans des régions codantes du génome, pouvant directement impacter la fonction du gène. Notamment, l'investigation de ces variants rares par génétique fonctionnelle dans des modèles *in vitro* ou *in vivo* permet de définir quels sont les variants ayant un effet LoF ou gain de fonction (gain-of-function [GoF]) sur le gène, et ainsi affiner les études d'association pour identifier de nouveaux gènes associés à la maladie, permettant d'approfondir nos connaissances concernant son héritabilité. De plus, l'étude de ces variants permet de comprendre quels sont les mécanismes associant le gène à la pathologie. Finalement, l'identification de ces variants rares chez les patients atteints de DT2 permet de se diriger vers une médecine de précision : le traitement du patient pourrait être adapté en fonction de l'anomalie génétique sous-jacente.

Les différents gènes qui ont été étudiés au cours de ce travail sont issus de l'étude RaDiO<sup>279</sup> (***Rare variants involved in Diabetes and Obesity***), qui a pour but l'identification de nouveaux

gènes associés au DT2 et aux traits métaboliques associés via notamment l'investigation de variants rares. Pour cela, les régions codantes de plus de 1200 gènes d'intérêt (incluant des gènes codant des GPCRs ou des facteurs de transcription, des gènes associés à d'autres types de diabète, des gènes identifiés par études GWAS...) ont été séquencées par NGS (Illumina) chez plus de 10 000 individus, incluant des participants atteints de DT2 et/ou d'obésité et des contrôles, afin d'y identifier des variants.

## 1 Le projet OPRD1

Parmi les différents gènes étudiés, nous avons décidé de nous intéresser aux variants du gène *OPRD1*, qui code le récepteur  $\delta$  aux opioïdes (DOP), identifiés par l'étude RaDiO.

Le récepteur DOP, ainsi que les récepteurs  $\mu$  (MOP), codé par le gène *OPRM1*, et  $\kappa$  (KOP), codé par le gène *OPRK1*, font partie de la famille des GPCRs couplés aux protéines  $G_i$ <sup>287</sup>. Ces récepteurs sont exprimés dans plusieurs régions du système nerveux, notamment dans les structures cérébrales impliquées dans l'émotion et la récompense, ainsi que dans les circuits neuronaux périphériques associés à la nociception<sup>288</sup>. Ces récepteurs sont principalement ciblés pour le traitement de la douleur<sup>289,290</sup>.

Les récepteurs aux opioïdes sont activés par liaison d'agonistes sur ceux-ci, qui peuvent être naturels tels que l'opium, semi-synthétiques comme l'héroïne, totalement synthétiques comme le fentanyl, ou encore endogènes, comme par exemple les endorphines<sup>291</sup>. Concernant le récepteur DOP, celui-ci a pour ligands endogènes principaux la métrénképhaline et la  $\beta$  endorphine<sup>292</sup>.

La liaison d'opioïdes agonistes sur le domaine extracellulaire N-terminal de ces récepteurs entraîne un changement de conformation de ceux-ci, permettant la liaison du GTP sur la protéine  $G_i$ , située à leur extrémité intracellulaire C-terminal. Ceci induit le scindement de la protéine G en une sous-unité  $G_\alpha$  et une sous-unité  $G\beta\gamma$ , qui peuvent activer différentes voies cellulaires. Notamment, cela permet l'inhibition de l'adénylate cyclase, ce qui entraîne une diminution du contenu intracellulaire en AMPc<sup>287,293</sup>. D'autre part, la liaison des ligands opioïdes sur ces récepteurs permet également la phosphorylation des récepteurs par les protéines G-protein coupled receptor kinases 2/3 (GRK2/3), menant ensuite à leur internalisation par les protéines  $\beta$  arrestines 1/2<sup>289</sup>. Le complexe formé par le récepteur aux opioïdes et la  $\beta$  arrestine peut activer plusieurs cascades de signalisation en aval, parmi

lesquelles la voie de signalisation des mitogen-activated protein kinases (MAPK)<sup>289,294</sup>, dont les plus étudiées sont les protéines ERK1/2<sup>289</sup>. Cependant, concernant le récepteur DOP, l'activation des protéines ERK1/2 peut se faire indépendamment de l'internalisation du récepteur<sup>295</sup> : celle-ci se fait via la cascade de signalisation en aval de la sous-unité Gβγ et de la protéine RAS<sup>296</sup>.

Plusieurs éléments nous ont poussé à étudier le gène *OPRD1* en particulier. Tout d'abord, comme expliqué précédemment, ce gène code un récepteur de la famille des GPCRs, qui sont des cibles thérapeutiques privilégiées. En effet, sur les 1100 cibles thérapeutiques connues aujourd'hui, environ 12% d'entre elles correspondent à des GPCRs<sup>297</sup>. D'autre part, la plupart des études réalisées chez des consommateurs d'opioïdes (sous forme d'opium ou d'héroïne) ont montré que cette addiction peut entraîner des altérations de l'homéostasie métabolique, notamment en diminuant l'appétit, en diminuant le poids, en diminuant la sécrétion d'insuline ou encore en augmentant la glycémie, avec parfois des effets contradictoires en fonction des études<sup>298</sup>.

Par ailleurs, bien que les gènes *OPRD1*, *OPRM1* et *OPRK1* soient tous les 3 exprimés dans le système nerveux central, seul *OPRD1* est fortement exprimé dans le pancréas<sup>288</sup> (figure 28). Via des analyses de données de RNA-seq issues d'îlots pancréatiques humains (n = 11) et murins (n = 14), obtenus à partir de la base de données Gene Expression Omnibus<sup>299-302</sup>, nous avons démontré que le gène *OPRD1* est le seul à être exprimé dans les îlots pancréatiques humains, et qu'aucun des gènes codant les récepteurs aux opioïdes n'est exprimé dans les îlots pancréatiques de souris (figure 29A). En outre, l'analyse de données de RNA-seq obtenues à partir de cellules β pancréatiques humaines purifiées, et à partir d'îlots pancréatiques humains déplétés en cellules β pancréatiques, a démontré que l'expression d'*OPRD1* est enrichie dans la préparation de cellules β pancréatiques *versus* la préparation des îlots déplétés en cellules β pancréatiques, alors que l'expression des gènes *OPRM1* et *OPRK1* est nulle<sup>303</sup> (figure 29B). Ainsi, *OPRD1* était pour nous un excellent gène-candidat dans le cadre du DT2 et de l'obésité.

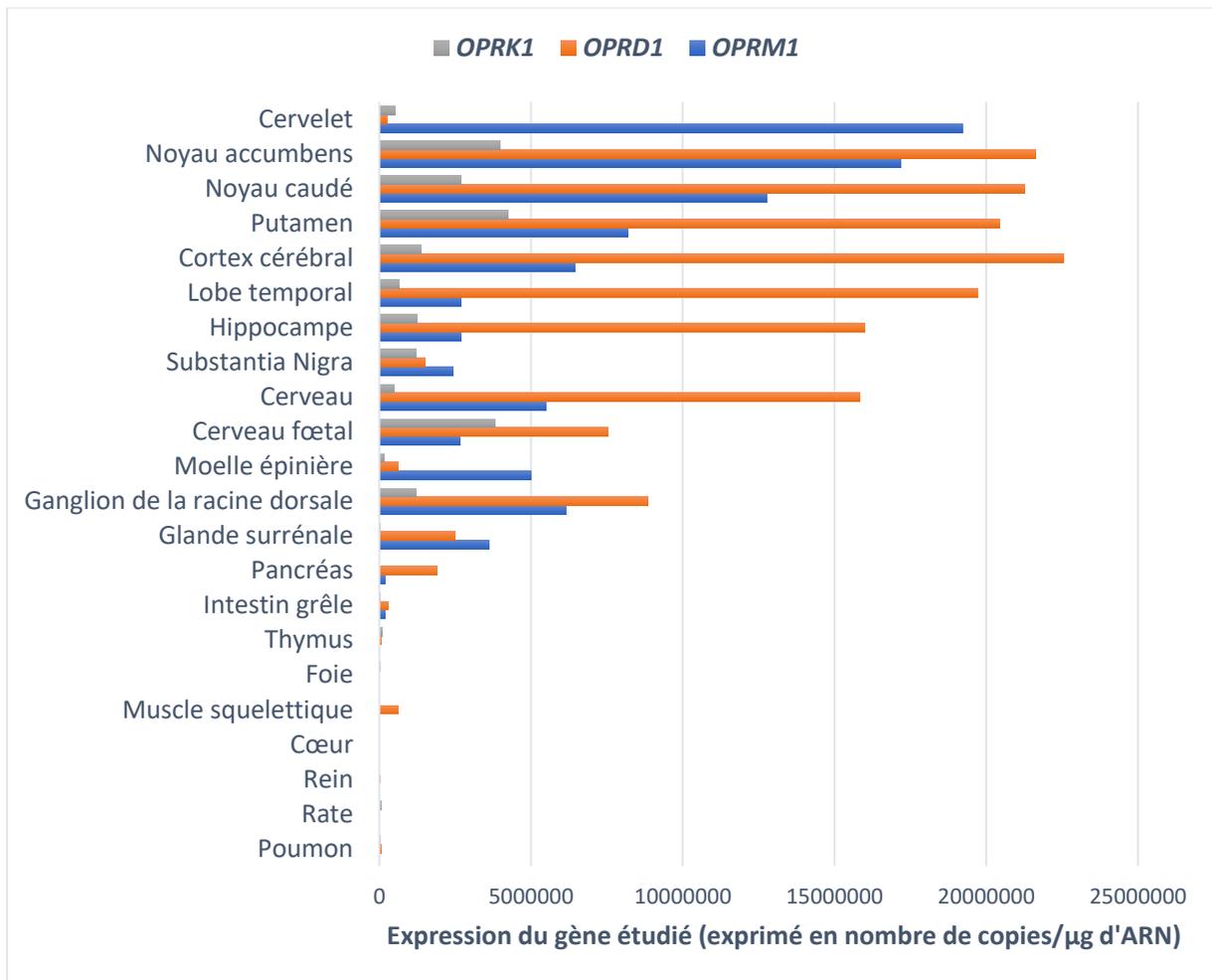


Figure 28. Expression tissulaire des gènes OPRK1, OPRD1 et OPRM1 chez l'humain. Graphique réalisé à partir des données issues de Peng et al ; « Opioid receptor expression in human brain and peripheral tissues using absolute quantitative real-time RT-PCR », *Drug and Alcohol Dependence*, 2013.

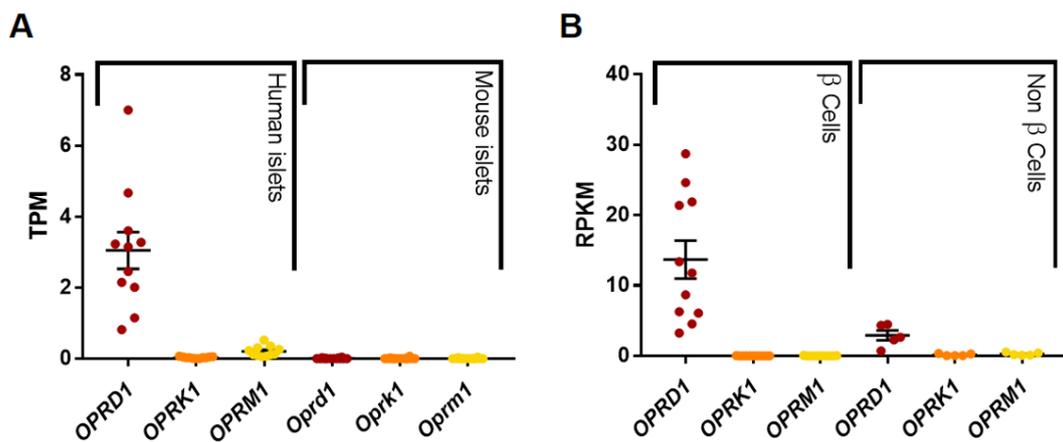


Figure 29. A) Expression des gènes OPRD1, OPRK1 et OPRM1 dans les îlots pancréatiques humains et des gènes Oprd1, Oprk1 et Oprm1 dans les îlots pancréatiques de souris. L'expression des gènes a été quantifiée par des expériences de RNA-seq. Les données concernant les îlots humains ont été obtenues à partir de 11 échantillons indépendants<sup>299,300</sup>, celles concernant les îlots de souris ont été obtenues à partir de 14 échantillons indépendants<sup>301,302</sup>. Ces résultats représentent la moyenne des TPM  $\pm$  SEM. **TPM** : transcripts per million. B) Expression des gènes OPRD1, OPRK1 et OPRM1 dans des préparations de cellules  $\beta$  pancréatiques humaines ( $\beta$  cells), et dans des préparations d'îlots pancréatiques humains déplétés en cellules  $\beta$  pancréatiques (Non  $\beta$  cells)<sup>303</sup>. Les données ont été téléchargées à partir de <ftp://jungle.unige.ch/BetaCellRNA-seq/>. Ces résultats représentent la moyenne des RPKM  $\pm$  SEM. **RPKM** : reads per kilobase million.

Dans le but de déterminer l'effet de chacun des variants du gène *OPRD1* identifiés dans l'étude RaDiO, nous avons réalisé des expériences fonctionnelles de type luciférase pour chacun d'entre eux. Ces expériences nous ont ainsi permis de classer ces variants en tant que variants GoF, LoF, ou neutres. Nous avons également investigué l'impact de chaque variant sur l'expression du récepteur DOP ainsi que sur sa localisation cellulaire par des expériences de western blot et d'immunofluorescence. Afin de déterminer si ces variants sont associés au DT2 ainsi qu'à d'autres traits métaboliques avec un effet délétère ou protecteur, des études d'association ont finalement été réalisées pour chaque agrégation de variants rares GoF ou LoF, puis répliquées chez d'autres groupes d'individus. Nous avons également investigué l'effet du gène *OPRD1* sur la régulation de l'insulinosécrétion.

## 2 Le projet *GLIS3*

Le gène *GLIS3* code le facteur de transcription GLI-Similar family zinc finger 3 (GLIS3). Les protéines GLIS1 à 3 font partie de la famille des facteurs de transcription Krüppel-like à doigts de zinc, qui ont la particularité de contenir un domaine en doigt de zinc très conservé<sup>304</sup>. Ce domaine en doigt de zinc permet la liaison des facteurs de transcription GLIS à des domaines particuliers de la région promotrice des gènes cibles, appelés « GLIS binding sites » (GLISBS)<sup>304</sup>. En s'y liant, les protéines GLIS1 à 3 peuvent ainsi jouer un rôle activateur ou répresseur sur la transcription. Cependant, GLIS1 et GLIS3 semblent être plutôt activateurs, tandis que GLIS2 est plutôt répresseur de la transcription<sup>304</sup>. Ces protéines sont exprimées durant le développement embryonnaire, mais également chez l'adulte, de façon spécifique au tissu et au type cellulaire<sup>305</sup>. Ainsi, bien que ces protéines soient toutes fortement exprimées dans le rein, elles présentent tout de même certaines disparités tissulaires : GLIS1 est exprimé dans le placenta, le tissu adipeux brun, le cerveau, le thymus, le côlon et les testicules<sup>306</sup> ; GLIS2 est exprimé dans le poumon, la prostate, le côlon, le cerveau, le foie, le cœur et les intestins<sup>307,308</sup>. Concernant le facteur de transcription GLIS3, celui-ci est principalement exprimé dans le rein, le pancréas et la thyroïde, et plus faiblement dans le cœur, le foie, le thymus, l'utérus, les ovaires, le cerveau et les poumons<sup>309-311</sup>.

Nous nous sommes spécifiquement intéressés à *GLIS3* pour plusieurs raisons. Tout d'abord, *GLIS3* est un gène déjà connu pour être impliqué dans une forme monogénique de diabète sévère, selon une hérédité autosomique récessive<sup>312</sup>. L'étude princeps a été réalisée par Senée et al., en 2006<sup>310</sup>. Celle-ci a mis en évidence la présence de plusieurs

mutations homozygotes rares du gène *GLIS3* (incluant une insertion homozygote c.2067\_2068insC, et des délétions de 426 et 149 kb dans la région 5'UTR du gène, NM\_152629.2) impliquées dans le développement d'un diabète néonatal permanent, couplé à une hypothyroïdie congénitale et au développement de reins polykystiques<sup>310</sup>. D'autres études menées par Dimitri et al. ont par la suite confirmé la pathogénicité des mutations rares homozygotes ou hétérozygotes composites du gène *GLIS3* et le développement du diabète néonatal, ainsi que d'autres anomalies, notamment des dysmorphies faciales<sup>313-315</sup>. Cependant, les phénotypes cliniques associés à ces mutations sont très variables. Cela pourrait être lié aux différents transcrits du gène qui sont touchés<sup>310</sup>, ou à l'effet spécifique des mutations sur la fonction de *GLIS3*, qui peuvent s'étendre d'une altération partielle de son activité à une perte totale de celle-ci<sup>310,311,314</sup>. D'autre part, *GLIS3* est un gène de susceptibilité au DT1<sup>312</sup>. En effet, plusieurs études GWAS ont permis d'identifier dans des populations de différentes origines, des SNPs du gène *GLIS3* associés au développement d'un DT1<sup>316-318</sup>. De plus, le variant p.A908V, identifié dans une cohorte japonaise avec une MAF de 0,7% chez les individus contrôles mais absent chez les individus atteints de DT1, est associé à une diminution du risque de DT1<sup>319</sup>. Par ailleurs, plusieurs études GWAS ont permis d'identifier des SNPs de *GLIS3* associés au DT2 ou aux traits métaboliques associés, tels que la glycémie à jeun ou une altération de la fonction  $\beta$  pancréatique<sup>312</sup>.

La protéine GLIS3 joue divers rôles importants dans le pancréas. Tout d'abord, ce facteur de transcription participe à la différenciation des îlots pancréatiques via la transactivation du gène *NGN3*, codant le facteur de transcription neurogénine 3<sup>312</sup>, qui joue un rôle crucial dans la différenciation des cellules précurseurs endocriniennes pancréatiques<sup>320</sup>. Des études ont montré une réduction importante de l'aire des îlots pancréatiques chez les souris *Glis3*<sup>-/-</sup>, qui ne correspond qu'à 15% de l'aire des îlots des animaux *Glis3*<sup>+/+</sup> : ces îlots sont plus petits, présentent une altération de leur morphologie, ainsi qu'une réduction du nombre de cellules endocriniennes qui les constituent<sup>321</sup>. Dans le pancréas embryonnaire de ce modèle, l'expression du gène *Ngn3* est réduite, mais la surexpression du gène *Glis3* permet de rétablir l'expression de *Ngn3*<sup>321</sup>.

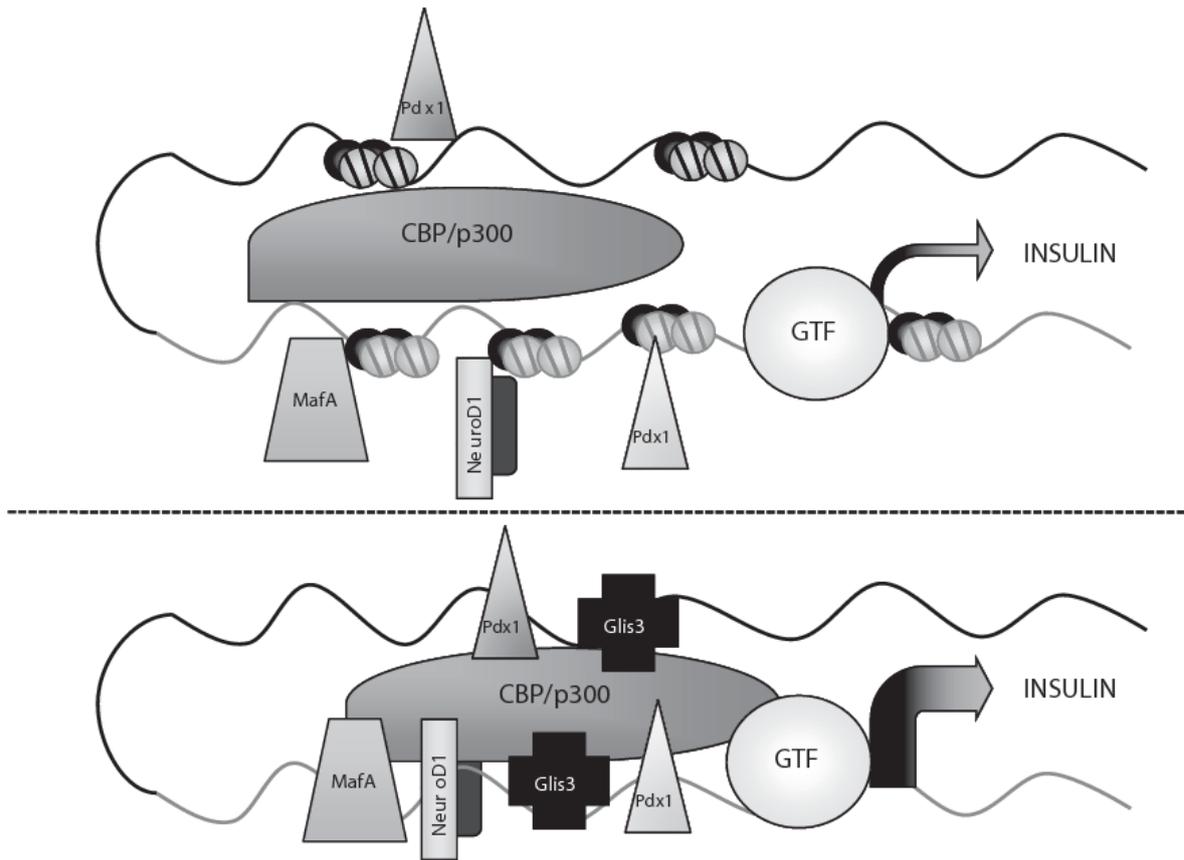


Figure 30. Mécanismes d'activation de la transcription du gène *INS* par les facteurs de transcription *GLIS3*, *MAFA*, *NEUROD1* et *PDX1*. En l'absence de *GLIS3* (en haut), les facteurs de transcription *MAFA*, *NEUROD1* et *PDX1* sont associés de manière transitoire avec la région promotrice du gène *INS*. En présence de *GLIS3* (en bas), celui-ci se lie sur des domaines de liaison *GLIS* binding sites (*GLISBS*) situés dans la région promotrice proximale du gène *INS*. *GLIS3* peut ainsi recruter le coactivateur *CBP/p300*, ainsi que les facteurs de transcription *MAFA*, *NEUROD1* et *PDX1*. Ceci permet de stabiliser l'interaction des facteurs de transcription sur le promoteur. Ce complexe d'activation peut alors activer la transcription du gène de l'insuline par son interaction avec des facteurs de transcription généraux (*GTF*). Source : Zeruth and Jetten, « The Krüppel-like protein Gli-similar 3 (*Glis3*) functions as a key regulator of insulin transcription », *Molecular Endocrinology*, 2013.

La protéine *GLIS3* est également impliquée dans la régulation de la transcription du gène *INS*<sup>312</sup>. Des études réalisées dans des cellules issues d'insulinome de rat *INS-1 832/13* ont démontré qu'une surexpression de *GLIS3* dans ces cellules stimule l'expression du gène *Ins2*, tandis que le KO de *GLIS3* réduit l'expression des gènes *Ins1* et *Ins2*<sup>311</sup>, mettant ainsi en évidence son implication dans la régulation de l'expression du gène de l'insuline. Plusieurs facteurs de transcription sont capables de réguler la transcription du gène de l'insuline, comme illustré à la figure 30. Parmi ceux-ci, les facteurs de transcription *PDX1*, *MAFA* BZIP transcription factor A (*MAFA*) et *NEUROD1* activent sa transcription en se liant respectivement aux boîtes A, aux éléments C, et aux boîtes E situés dans sa région promotrice<sup>322</sup>. Cependant, la région promotrice du gène de l'insuline contient également deux domaines *GLISBS*, sur lesquels la protéine *GLIS3* peut venir se lier<sup>311,323</sup>. *GLIS3*, après s'être lié sur ces domaines *GLISBS*, est capable d'interagir avec les facteurs de transcription *PDX1*, *MAFA* et *NEUROD1*

afin d'agir en synergie sur l'activation de la transcription du gène *INS*<sup>311,324</sup>. Des mutations dans les domaines GLISBS, ou de la protéine GLIS3 altérant son domaine en doigt de zinc ou son domaine de transactivation entraîne une diminution de la transactivation du gène de l'insuline par la combinaison des facteurs NEUROD1, MAFA et PDX1<sup>324</sup>. La nécessité de la liaison de GLIS3 sur le promoteur de l'insuline pour activer sa transcription a également été démontrée in vivo : chez des souris chez lesquelles le gène *Glis3* a été inactivé, on observe une diminution de l'expression de l'insuline, menant à l'apparition d'un diabète fulminant et à la mort des animaux<sup>325</sup>. En outre, le gène *GLIS3* pourrait également être impliqué dans la régulation de l'insulinosécrétion. Ceci est suggéré par le fait que chez des souris déficientes pour le gène *Glis3*, une diminution de l'expression des gènes *Glut2*, codant le transporteur de glucose majeur de la cellule  $\beta$  pancréatique, et *Abcc8*, codant la sous-unité SUR1 des canaux  $K^+_{ATP}$ , a été observée<sup>323,325</sup>.

La protéine GLIS3 est également impliquée dans l'expansion de la masse de cellules  $\beta$  pancréatiques associée à un état d'insulinorésistance (obésité ou grossesse), permettant ainsi d'augmenter la sécrétion d'insuline en réponse à l'insulinorésistance<sup>312</sup>. Une étude réalisée chez des souris *Glis3*<sup>+/-</sup> a mis en évidence que, tandis qu'une alimentation classique n'induit pas d'altération du métabolisme chez les animaux, une alimentation riche en graisse entraîne un diabète couplé à une hypoinsulinémie, provoqués par une altération de l'expansion de la masse de cellules  $\beta$  pancréatiques<sup>325</sup>. Cette expansion se fait principalement via la prolifération des cellules  $\beta$  pancréatiques<sup>312,326</sup>. GLIS3 est impliqué dans le contrôle de cette prolifération via la régulation de l'expression du gène *CCND2* codant la cycline D2<sup>325</sup>, impliquée dans la croissance des cellules  $\beta$  pancréatiques en post-natal et l'expansion de la masse de ces cellules en conditions d'insulinorésistance<sup>327-329</sup>.

Jusqu'à présent, aucune étude n'avait jusqu'alors pu montrer un lien fort entre des variants rares hétérozygotes du gène *GLIS3*, qui ont un effet beaucoup plus fort que les variants fréquents, et le DT2, voire un impact direct de variants rares hétérozygotes dans une forme monogénique autosomique dominante de type MODY, comme cela a été démontré pour les gènes *GCK* et *PDX1*. La seule étude réalisée sur des variants hétérozygotes rares de ce gène date de 2019 : il s'agit d'une étude danoise basée sur le séquençage des régions codantes du gène dans une cohorte d'environ 9000 individus témoins ou diabétiques, qui a permis d'identifier 88 variants non-synonymes du gène *GLIS3* parmi lesquels 79 sont très rares

(avec une MAF < 0,1%)<sup>330</sup>. L'agrégation de ces variants rares (sans sélection particulière) présente un enrichissement chez les individus atteints de DT2 (OR = 1,37 ;  $P = 0,04$ )<sup>330</sup>. De plus, chez les individus diabétiques, ces variants rares sont associés à une augmentation du taux de HbA1c (7,20% chez les porteurs de mutation et 6,80% chez les non-porteurs,  $P = 0,02$ ). Cependant, aucune analyse fonctionnelle n'a été réalisée. C'est pourquoi nous avons également choisi d'approfondir l'étude de ce gène par de la génétique fonctionnelle.

Afin d'investiguer l'effet fonctionnel des variants rares du gène *GLIS3* identifiés dans l'étude RaDiO, nous avons réalisé des expériences de luciférase. Les critères ACMG ont ensuite été appliqués à chacun des variants LoF afin d'identifier les variants P/LP de ce gène. Des études d'association ont ensuite été réalisées entre les variants rares P/LP et le DT2 afin de déterminer si ces variants sont associés au risque de développer la maladie.

## **MATERIEL & METHODES**

# Matériel et méthodes

## 1 Etudes génétiques

### 1.1 Participants aux études

#### 1.1.1 Participants à l'étude RaDiO

L'étude des gènes *OPRD1* (NM\_000911.4) et *GLIS3* (NM\_001042413.1) a débuté par un séquençage d'exomes, effectué sur de l'ADN prélevé dans le plasma sanguin de 6791 individus (pour *OPRD1*) et 5471 individus (pour *GLIS3*) issus de l'étude RaDiO<sup>279</sup>. Cette étude comprend des individus issus de plusieurs études épidémiologiques : 1) l'étude D.E.S.I.R. (Données Epidémiologiques sur le Syndrome d'InsulinoRésistance), qui est une étude prospective française s'étalant sur une période de 9 ans, comprenant des hommes et des femmes habitant l'ouest de la France<sup>331</sup> ; 2) des individus recrutés par l'UMR 1283/8199, par le service de nutrition de l'hôpital Hôtel-Dieu (Paris, France), ou par le CEPH (Centre d'Etude du Polymorphisme Humain, hôpital Saint-Louis, France)<sup>332,333</sup> ; 3) des participants recrutés par le service d'endocrinologie de l'hôpital de Corbeil-Essonnes (Corbeil-Essonnes, France)<sup>332</sup> ; 4) ainsi que des individus recrutés par l'association Fleurbaix Laventie Ville Santé<sup>334</sup>.

Pour chaque participant, plusieurs informations concernant son profil métabolique sont connues, telles que son IMC, son statut glycémique, son profil lipidique, ou encore sa pression sanguine. Ainsi, les participants sont considérés comme ayant un poids corporel normal lorsque leur IMC est inférieur à 25 kg/m<sup>2</sup>, en surpoids lorsqu'il est compris entre 25 et 30 kg/m<sup>2</sup>, et obèses s'il est supérieur à 30 kg/m<sup>2</sup>. Les participants sont considérés comme atteints d'hyperglycémie si leur glycémie à jeun est supérieure à 5,6 mmol/L et/ou s'ils prennent un traitement hypoglycémiant. Les participants sont considérés comme atteints de DT2 si leur glycémie à jeun est supérieure à 7 mmol/L et/ou s'ils prennent un traitement hypoglycémiant, sans la présence d'anticorps. Les participants contrôles utilisés dans les études cas-contrôles de l'hyperglycémie ou du DT2 ont une glycémie à jeun inférieure à 5,6 mmol/L, n'ont aucun traitement hypoglycémiant et sont âgés de plus de 40 ans. En outre, les taux de HDL-cholestérol et de triglycérides n'ont été évalués que chez les patients qui ne sont pas sous traitement hypolipémiant. Enfin, la pression artérielle systolique et diastolique a été évaluée uniquement chez les participants qui ne prennent pas de traitement contre l'hypertension.

### 1.1.2 Participants à la UK Biobank

Les données issues de la UK Biobank ont permis de répliquer notre étude concernant le gène *OPRD1* dans une seconde cohorte, composée de 187 242 individus. La UK Biobank est une base de données, reprenant des résultats de séquençage d'exomes, et permettant ainsi de connaître la séquence codante de plusieurs gènes chez des centaines de milliers d'individus. Les participants à la UK Biobank sont des individus originaires du Royaume-Uni, et ont été recrutés entre mars 2006 et décembre 2010.

De la même façon que pour la cohorte issue de l'étude RaDiO, plusieurs informations concernant le profil métabolique de chacun des individus issus de la UK Biobank sont disponibles. Ainsi, comme pour l'étude RaDiO, les individus sont considérés comme ayant un poids normal s'ils ont un IMC inférieur à 25 kg/m<sup>2</sup>, en surpoids s'il est compris entre 25 et 30 kg/m<sup>2</sup>, et obèses s'il est supérieur à 30 kg/m<sup>2</sup>. En outre, concernant l'étude cas-contrôles du DT2, les individus sont considérés comme des « cas » diabétiques s'ils répondent à l'un des critères suivants : ils ont un taux de HbA1c supérieur ou égal à 48 mmol/mol Hb, ils sont traités avec de l'insuline, ou ils ont été diagnostiqués comme étant diabétiques par un médecin. Les individus sont exclus de l'étude s'ils sont atteints de DT1, ou s'ils ont présenté un diabète gestationnel. Les individus sont considérés comme des « contrôles » s'ils ont un taux de HbA1c inférieur à 42 mmol/mol Hb, et s'ils ont au moins 40 ans. Les individus sont exclus de l'étude cas-contrôles s'ils sont atteints de DT1, de diabète non-insulinodépendant, de diabète gestationnel ou d'un autre type de diabète, s'ils ont été diagnostiqués comme souffrant de diabète par un médecin, s'ils ont une glycémie élevée ou s'ils sont traités avec de l'insuline.

## 1.2 Identification des variants du gène *OPRD1*

### 1.2.1 Séquençage des gènes *OPRD1* et *GLIS3* et identification des variants dans l'étude RaDiO

Dans le cadre de l'étude RaDiO, un séquençage ciblé des gènes *OPRD1* et *GLIS3* a été réalisé par NGS. Pour cela, un enrichissement ciblé de ces gènes a été réalisé à l'aide du kit NimbleGen SeqCap EZ Choice XL (Roche Molecular Diagnostics, Pleasanton, Etats-Unis), permettant de séquencer spécifiquement certaines régions du génome. Le processus est repris à la figure 31.

Brièvement, afin de préparer la librairie de fragments d'ADN, 1 µg d'ADN issu du plasma sanguin des individus étudiés est fragmenté par sonication à l'aide de l'ultrasonicateur

E220 (Covaris, Woburn, Etats-Unis). Les extrémités des fragments d'ADN obtenus sont réparées de façon à obtenir des extrémités franches, puis liées à des adaptateurs à l'aide du kit KAPA High Throughput Library Preparation kit (Roche) via l'appareil Microlab STARlet (Hamilton, Reno, Etats-Unis). Ces adaptateurs ont plusieurs rôles : ils permettent la liaison de la librairie à la flow cell de séquençage en formant des clusters, servent de site d'ancrage pour les amorces de séquençage, et servent d'index aux différents fragments, permettant ainsi d'identifier l'individu d'origine. Les fragments sont ensuite amplifiés par PCR.

Après sélection selon leur taille et quantification de ceux-ci à l'aide de l'appareil LabChip GX (PerkinElmer, Waltham, Etats-Unis), 24 échantillons sont combinés ensemble, de façon à former un pool d'ADN de 1 µg, puis sont hybridés à un pool de sondes marquées à la biotine SeqCap EZ (Roche) et sont incubés avec des billes de streptavidine durant 72 heures à 47°C. Les différents fragments capturés sur les billes sont alors purifiés avec le kit SeqCap Hybridization and Wash Kit (Roche), au moyen de l'appareil Bravo Automated Liquid Handling Platform (Agilent, Santa Clara, Etats-Unis), puis sont amplifiés à l'aide du kit KAPA HiFi HotStart ReadyMix (Roche). Ils sont ensuite quantifiés avec l'appareil LabChip GX (PerkinElmer), ainsi qu'avec l'appareil Qubit (Invitrogen / Thermo Fisher Scientific, Waltham, Etats-Unis) par dosage fluorométrique. Les différents fragments d'ADN obtenus sont ensuite séquencés sur l'appareil HiSeq 4000 (Illumina, San Diego, Etats-Unis), suivant un protocole de séquençage paired-end (2x150 bp), permettant de séquencer les fragments à partir des deux extrémités.

Le séquençage Illumina est basé sur la stratégie de séquençage par synthèse, dont le principe est repris à la figure 32. Brièvement, les fragments d'ADN sont liés à une plaque en verre appelée flow cell par l'intermédiaire des adaptateurs. Des amorces de séquençage s'apparient sur les adaptateurs, puis les fragments d'ADN sont directement amplifiés par PCR sur la flow cell afin de former un cluster de fragments d'ADN identiques, qui permettra d'amplifier le signal fluorescent obtenu par la suite. De l'ADN polymérase ainsi que des nucléotides fluorescents, couplés à un élément empêchant la suite de la synthèse du brin d'ADN, sont ensuite ajoutés à la flow cell. Le nucléotide complémentaire à celui du brin d'ADN à séquencer est intégré au fragment d'ADN nouvellement formé, sans qu'aucun autre nucléotide ne puisse venir s'y lier. A ce stade, la fluorescence émise est lue par le séquenceur, et permet de savoir quelle base a été intégrée au fragment d'ADN en formation. Ensuite, le fluorophore ainsi que l'élément empêchant la suite de la synthèse du brin d'ADN sont clivés, la flow cell est lavée,

puis de l'ADN polymérase et les différents nucléotides fluorescents sont à nouveau ajoutés à la flow cell. Ainsi, au cours de la formation du brin d'ADN, chaque base peut être lue par le séquenceur, et il est possible de reconstituer la séquence de départ.

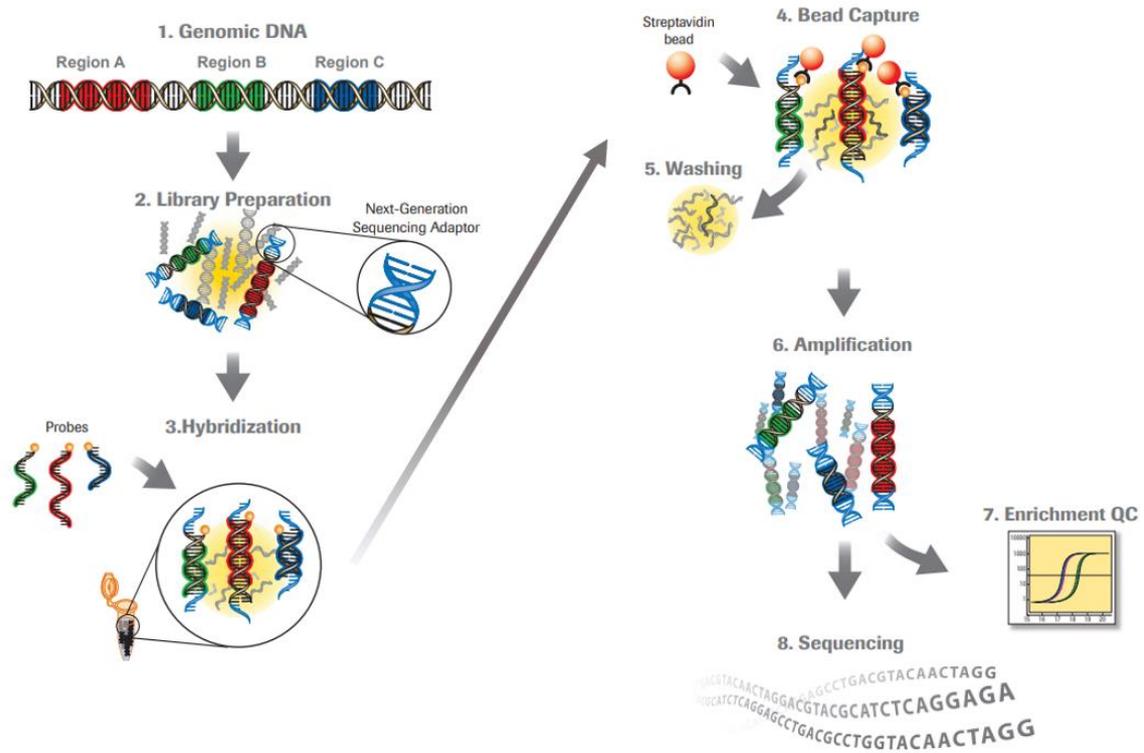


Figure 31. Principe du séquençage de nouvelle génération. Source : adapté à partir du site web [https://www.ndsu.edu/pubweb/~mcclean/ctig/ctigfall2011/05227887001\\_SeqCapBroch\\_Oct2011.pdf](https://www.ndsu.edu/pubweb/~mcclean/ctig/ctigfall2011/05227887001_SeqCapBroch_Oct2011.pdf)

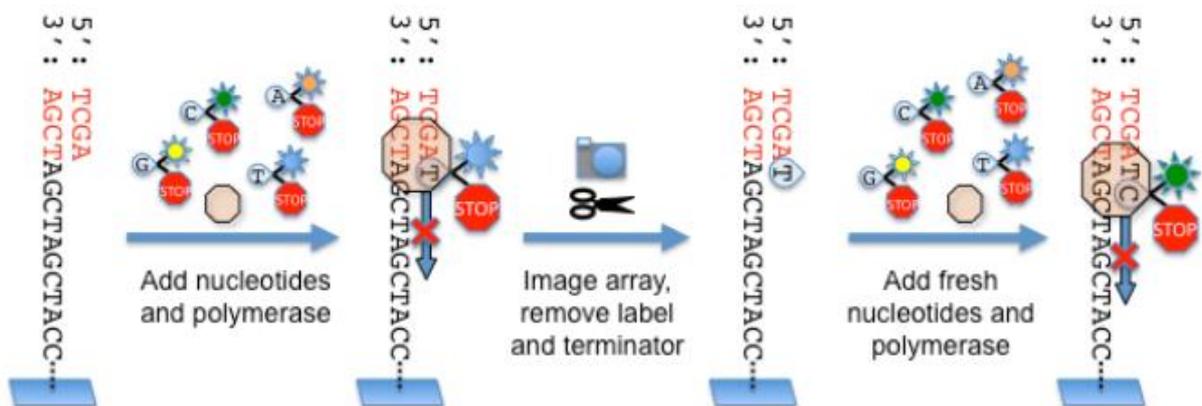


Figure 32. Principe du séquençage par synthèse. Source : Andersen and Schrijver, « Next Generation DNA Sequencing and the Future of Genomic Medicine », Genes, 2010.

Le séquençage permet ainsi d'obtenir les séquences de chacun des fragments d'ADN, appelées reads. Les données issues du séquençage sont ensuite démultiplexées à l'aide du programme bcl2fastq Conversion Software (Illumina, version 2.17), permettant ainsi, grâce aux index attribués précédemment, de réorganiser les différents reads en fonction des

individus. Ces reads sont ensuite alignés sur la séquence du génome humain (version hg19/GRCh37) à l'aide de l'outil Burrows-Wheeler Aligner. L'appel des variants, permettant de repérer les différents variants suite au séquençage, est effectué à l'aide de l'outil Genome Analysis ToolKit (GATK, version 3.3). Seuls les variants pour lesquels une couverture supérieure à 8 reads a été obtenue sont conservés. Les variants sont ensuite annotés : ceci consiste à assigner, à chaque variant, plusieurs informations à son sujet provenant généralement de bases de données, telles que la fréquence du variant dans la population, sa conséquence moléculaire, ou encore la prédiction de son effet délétère. Cette annotation est réalisée à l'aide du programme Ensembl Perl Application Program Interfaces (version 75) et des scripts de Perl custom, des bases de données dbSNP (version 135)<sup>335</sup>, et dbNSFP (version 3.0)<sup>336</sup>.

Afin de pouvoir être conservés dans notre analyse, les différents variants détectés par séquençage doivent répondre à certains critères. Ainsi, pour chaque participant, il est nécessaire de ne pas avoir plus de 5% de génotype manquant (correspondant à une couverture inférieure à 8 reads, ou un QUAL score, indiquant la qualité de l'échantillon, inférieur à 50).

#### 1.2.2 Identification des variants du gène *OPRD1* dans la UK Biobank

Dans la UK Biobank, les variants du gène *OPRD1* ont été identifiés à partir des données de séquençage d'exomes réalisé chez les participants. L'annotation des variants a été réalisée avec l'outil Ensembl Variant Effect Predictor (version 103, RefSeq). Afin d'être conservés pour la suite des analyses, ces variants doivent présenter moins de 5% de génotype manquant (correspondant à une couverture inférieure à 10 reads, ou un score QG inférieur à 20). La suite des analyses ne s'est intéressée qu'aux variants LoF, correspondant dans ce cas-ci aux variants génétiquement nuls (variants non-sens, où le changement d'un nucléotide génère un codon STOP ; frameshift, qui présentent un décalage du cadre de lecture ; variants d'épissage canoniques situés à plus ou moins 1 ou 2 nucléotides du site d'épissage ; ainsi que les variants start lost, induisant la perte du codon START), et aux variants déjà identifiés dans le projet RaDiO et définis comme étant LoF par les analyses fonctionnelles *in vitro*.

### 1.3 Etudes d'association entre les variants des gènes *OPRD1* et *GLIS3*, et divers traits métaboliques

#### 1.3.1 Etudes d'association réalisées pour l'étude RaDiO et la UK Biobank

Afin de réaliser des études d'association pour les variants rares d'*OPRD1* et de *GLIS3* identifiés dans l'étude RaDiO, et pour les variants rares d'*OPRD1* issus de la UK Biobank, ceux-ci ont été regroupés en clusters.

Pour le projet *OPRD1*, les variants rares du gène identifiés dans l'étude RaDiO ont été répartis dans 2 clusters différents selon qu'ils soient définis comme GoF ou LoF par les expériences de luciférase. Il est à noter que l'agrégation de variants LoF comprend également le variant c.1063G>T/p.E355\*, qui est considéré comme un variant nul d'un point de vue génétique. Concernant les données issues de la UK Biobank, les variants rares LoF ont également été regroupés en un seul cluster.

Pour le projet *GLIS3*, les variants définis comme LoF au cours des expériences de luciférase ont dans un premier temps été regroupés ensemble. Cette agrégation a par la suite été affinée par application des critères ACMG (incluant le critère PS3 issu des analyses fonctionnelles *in vitro*), repris dans les tableaux 5 et 6. Ainsi, seuls les variants définis comme étant P/LP selon ces critères ACMG ont été conservés pour la réalisation des études d'association.

Chacune de ces agrégations a ensuite été analysée suivant la méthode MiST<sup>337</sup>, permettant d'obtenir un score statistique  $S(\pi)$ , correspondant à l'effet moyen ( $\pi$ ) du cluster de variants, ainsi qu'un score statistique  $S(\tau)$ , correspondant à l'effet hétérogène ( $\tau$ ) du cluster (par exemple s'ils sont plus présents chez les cas ou chez les contrôles, ou si certains variants ont un effet beaucoup plus fort que d'autres). Le calcul est réalisé selon l'équation suivante :

$$Y = \alpha X + \pi GZ, \text{ où}$$

- Y correspond à la matrice des traits métaboliques d'intérêt,
- X correspond à la matrice des covariables : elles comprennent l'âge, le sexe, l'IMC, ainsi que les composantes principales PC1 à PC5 (permettant de définir l'origine ethnique des participants et calculées à partir de plus de 10 000 SNPs fréquents analysés chez les patients en commun avec le projet « 1000 Genomes » qui inclut des patients de multiples origines ethniques) lorsque l'étude d'association se porte sur l'hyperglycémie, le DT2, les niveaux de HDL- et LDL-cholestérol et la pression sanguine,

et comprennent l'âge, le sexe, et les composantes principales PC1 à PC5 lorsque l'étude s'intéresse au surpoids, à l'obésité et à l'IMC,

- G correspond à la matrice des variants du gène étudié,
- Z correspond au vecteur de la matrice des variants, répété  $n$  fois, avec  $n$  correspondant au nombre de variants rares.

Ainsi,  $\pi GZ = \sum_{i=1}^n G_i$ .

### 1.3.2 Etudes d'association réalisées pour l'AMP T2D Knowledge Portal

Des études d'association ont également été réalisées pour le variant du gène *OPRD1* c.154A>G/p.I52V, qui est fréquent dans l'ethnie africaine, à partir des données issues de l'AMP T2D Knowledge Portal. Ces analyses ont été réalisées suivant un modèle de régression linéaire additif, ajusté à l'âge, au sexe, à l'IMC et aux composantes principales PC1 à PC5 pour les études concernant le DT2, les niveaux HDL- et LDL-cholestérol, les triglycérides, et la pression sanguine, et ajusté à l'âge, au sexe, et aux composantes principales PC1 à PC5 pour les études s'intéressant à l'IMC.

Pour les études d'association concernant les niveaux de lipides sériques et la pression sanguine, seuls les individus ne présentant pas de DT2 ont été inclus. En outre, pour l'analyse du DT2, seul des contrôles de plus de 50 ans ont été inclus.

## 2 Etudes fonctionnelles

### 2.1 Génération des plasmides

Dans le cadre du projet *OPRD1*, le plasmide p*OPRD1* contenant le gène *OPRD1* WT (NM\_000911.4), ainsi que le plasmide correspondant à la condition « STOP », dans lequel un codon STOP a été inséré à la suite du premier codon du gène *OPRD1* (c.3\_4insTAA/p.E2\*), et servant de contrôle négatif, ont été désignés et commandés chez e-Zyvec (Loos, France). Pour 31 variants du gène *OPRD1*, les plasmides ont été créés par mutagenèse dirigée à partir du plasmide p*OPRD1* WT. Les 3 plasmides restants, contenant respectivement les variants c.658T>C/p.F220L, c.840C>G/p.F280L et c.1094G>A/p.G365D, et dont la création par mutagenèse dirigée s'est révélée compliquée, ont également été commandés chez E-Zyvec.

Tableau 5. Critères 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
Très forte preuve de pathogénicité (very strong)	PVS1	Variant nul (variant non-sens, frameshift, 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
Forte preuve de pathogénicité (strong)	PS1	Variant induisant un changement d'acide aminé identique à un autre, déjà établi pour un autre variant pathogénique, peu importe le changement de nucléotide
	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
Preuve modérée de pathogénicité (moderate)	PM1	Variant situé dans un point chaud (hot spot) de mutations, ou dans domaine fonctionnel critique ne contenant pas de variation bénigne
	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, ou ExAC
	PM3	Pour les maladies récessives, variant détecté en trans 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 à 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
	PM6	Variant supposé apparu <i>de novo</i> , sans confirmation paternelle et maternelle
Preuve secondaire de pathogénicité (supporting)	PP1	Coségrégation avec la maladie chez plusieurs membres de la famille atteints, 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
	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 6. 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	Combinaisons de critères possibles
Pathogène	1 critère de pathogénicité très fort (PVS1) et : <ul style="list-style-type: none"> <li>- ≥ 1 critère fort (PS1-PS4) ou</li> <li>- ≥ 2 critères modérés (PM1-PM6) ou</li> <li>- 1 critère modéré (PM1-PM6) et 1 critère secondaire (PP1-PP5) ou</li> <li>- ≥ 2 critères secondaires (PP1-PP5)</li> </ul>
	≥ 2 critères de pathogénicité forts (PS1-PSS4)
	1 critère de pathogénicité fort (PS1-PS4) et : <ul style="list-style-type: none"> <li>- ≥ 3 critères modérés (PM1-PM6) ou</li> <li>- 2 critères modérés (PM1-PM6) et ≥ 2 critères secondaires (PP1-PP5) ou</li> <li>- 1 critère modéré (PM1-PM6) et ≥ 4 critères secondaires (PP1-PP5)</li> </ul>
Probablement pathogène	1 critère très fort (PVS1) et 1 critère modéré (PM1-PM6)
	1 critère fort (PS1-PS4) et 1- 2 critères modérés (PM1-PM6)
	1 critère fort (PS1-PS4) et ≥ 2 critères secondaires (PP1-PP5)
	≥ 3 critères modérés (PM1-PM6)
	2 critères modérés (PM1-PM6) et ≥ 2 critères secondaires (PP16-PP5)
	1 critère modéré (PM1-PM6) et ≥ 4 critères secondaires (PP1-PP5)

Le plasmide pGL4.33[*luc2P/SRE/Hygro*], contenant le gène de la luciférase sous le contrôle du Serum Response Element (SRE), et qui sera dénommé pSRE-LUC dans la suite de cet écrit, provient de chez Promega (Madison, Etats-Unis, #E1340). Le plasmide contenant le gène de la β-galactosidase de la bactérie *E. Coli*, quant à lui, correspond à l'un des plasmides développés par MacGregor et Caskey<sup>338</sup> et était déjà utilisé par notre laboratoire.

Concernant le projet *GLIS3*, le plasmide pGLIS3 contenant le gène *GLIS3* WT (NM\_001042413.1), a été désigné et commandé chez Twist (San Francisco, Etats-Unis). Parmi les 105 plasmides contenant les variants du gène *GLIS3* à étudier, 100 d'entre eux ont été créés par mutagenèse dirigée à partir du plasmide pGLIS3 WT. Les 5 plasmides restants, à savoir ceux contenant les variants c.1126C>T/p.P376S, c.2051C>T/p.P684L, c.2404G>A/p.G802S, c.2636C>T/p.S879L et c.2785G>A/p.E929K, n'ont pas pu être créés par mutagenèse dirigée et ont donc été commandés chez Twist. Le plasmide pMAFA, contenant le gène *MAFA* (NM\_201589), provient de chez Origene (Rockville, Etats-Unis, #RG214782). Le

plasmide pINS-LUC, contenant le gène de la luciférase sous le contrôle de la région flanquante 5' du gène *INS* (s'étendant de -423 à -18 par rapport au site de début de transcription)<sup>339</sup>, a été fourni à notre laboratoire par le Dr. Danielle Melloul (Hadassah University Hospital, Israël). Le plasmide contenant le gène de la  $\beta$ -galactosidase est identique à celui utilisé dans le cadre du projet *OPRD1*.

Les cartes des plasmides utilisés au cours de ces projets sont reprises à l'annexe 1.

### 2.1.1 PCR de mutagenèse

Les plasmides contenant les différents variants des gènes *OPRD1* et *GLIS3* qui n'ont pas été commandés ont été générés par PCR de mutagenèse, à l'aide du Quick Change II XL Site-Directed Mutagenesis Kit (Agilent). Cette manipulation permet de générer, à partir du plasmide contenant le gène d'intérêt WT, un nouveau plasmide contenant la mutation désirée. Pour chaque plasmide à créer, on prépare, dans les puits d'une plaque PCR à 96 puits, un mix contenant 1  $\mu$ L de plasmide pOPRD1 ou pGLIS3 WT et 1,25  $\mu$ L d'amorces modifiées (portant la mutation désirée) forward et reverse, auxquels on ajoute 5  $\mu$ L de 10X reaction buffer, 1,2  $\mu$ L de dNTP mix, 3  $\mu$ L de QuickSolution reagent, ainsi que 36,3  $\mu$ L d'eau. A ce mix est ensuite ajouté 1  $\mu$ L d'enzyme *PfuUltra* High Fidelity DNA polymerase. La PCR de mutagenèse est ensuite lancée dans un thermocycleur selon le cycle repris dans le tableau 7 pour les plasmides contenant les variants d'*OPRD1*, et dans le tableau 8 pour ceux contenant les variants de *GLIS3*. Les listes des amorces modifiées utilisées pour les PCR de mutagenèse concernant les projets *OPRD1* et *GLIS3* sont reprises dans les annexes 2 et 3, respectivement.

Tableau 7. Cycle de température appliqué au cours de la PCR de mutagenèse visant à générer les plasmides exprimant les différents variants d'*OPRD1*.

<b>1X</b>	<b>18X</b>			
95°C	95°C	55°C	70°C	4°C
1 minute	1 minute	1 minutes	19 minutes	$\infty$

Tableau 8. Cycle de température appliqué au cours de la PCR de mutagenèse visant à générer les plasmides exprimant les différents variants de *GLIS3*.

<b>1X</b>	<b>18X</b>			<b>1X</b>	
95°C	95°C	55°C	70°C	68°C	4°C
1 minute	1 minute	1 minutes	22 minutes	7'	$\infty$

### 2.1.2 Dialyse

Les produits de PCR de mutagenèse sont ensuite dialysés afin d'éliminer les réactifs de PCR restants, tels que l'enzyme, les dNTPs ou les amorces. Pour cela, la totalité du produit de PCR obtenu est déposée sur un filtre qui repose sur de l'eau, permettant ainsi d'éliminer les réactifs de PCR restants tout en retenant les plasmides. Après 20 minutes, le produit de PCR purifié, composé des plasmides WT et mutés restés à la surface du filtre, est récupéré.

### 2.1.3 Digestion

Suite à la dialyse, afin de ne conserver que les plasmides mutés, il est nécessaire de se débarrasser des plasmides pOPRD1 et pGLIS3 WT. Pour cela, le produit de PCR dialysé est soumis à une digestion enzymatique afin de cliver le plasmide WT méthylé, contrairement au plasmide muté qui ne l'est pas. Pour cela, dans les puits d'une plaque PCR à 96 puits, 5 µl de produit de PCR dialysé sont mélangés avec 1 µL d'enzyme DpNI (clivant les plasmides au niveau des adénines méthylées), 1 µL de tampon Cutsmart (New England Biolabs, Ipswich, Etats-Unis), et 4,6 µL d'eau. La réaction enzymatique est ensuite lancée pour 6 heures à 37°C.

### 2.1.4 Transformation

Après l'obtention des plasmides pOPRD1 et pGLIS3 mutés purifiés, ceux-ci sont amplifiés par transformation bactérienne. Pour cela, des bactéries E. Coli One Shot TOP10 chimiquement compétentes (Invitrogen / Thermo Fisher Scientific) sont utilisées. Dans un premier temps, 5 µL de plasmide pOPRD1 ou pGLIS3 muté purifié sont ajoutés dans le tube contenant les bactéries TOP10. Ces bactéries sont ensuite placées sur glace pendant 30 minutes, puis chauffées à 42°C pendant 30 secondes et replacées sur glace 2 minutes afin de provoquer un choc thermique facilitant la transformation bactérienne. Elles sont ensuite mises en culture en suspension dans 250 µL de milieu Super Optimal Broth (SOC) (Invitrogen / Thermo Fisher Scientific) sous agitation à 225 rpm pendant une heure, à 37°C. Finalement, les bactéries transformées sontensemencées sur une gélose de milieu Lysogeny Broth (LB), enrichie en kanamycine (pour les plasmides pOPRD1) ou en ampicilline (pour les plasmides pGLIS3), et sont incubées à 37°C pendant la nuit.

### 2.1.5 Minipreps et extraction des plasmides

Suite à la transformation bactérienne, il est nécessaire d'amplifier les bactéries qui ont bien intégré les plasmides mutés : il s'agit de celles qui ont pu pousser sur la gélose, grâce à la

présence du gène de résistance à l'antibiotique présent dans les plasmides. Pour cela, 4 à 6 colonies bactériennes présentes sur la gélose sont numérotées, prélevées, puis mises en culture en suspension dans 3 mL de milieu LB enrichi en antibiotique correspondant (kanamycine ou ampicilline) sous agitation à 165 rpm à 37°C, pendant la nuit. A ce stade, 800 µL de ces bactéries sont prélevés et mélangés à 800 µL de glycérol afin d'en conserver un stock à -80°C.

L'ADN plasmidique des bactéries transformées et amplifiées est ensuite extrait afin de pouvoir récupérer le plasmide muté. Cette extraction est réalisée à l'aide du kit NucleoSpin Plasmid, Mini kit for plasmid DNA (Macherey-Nagel, Düren, Allemagne).

En pratique, pour chaque colonie amplifiée, 1,9 mL de milieu de culture contenant les bactéries est récupéré et centrifugé à 11 000 g pendant 10 minutes. Le surnageant est éliminé et seul le culot de bactéries est conservé. 250 µL de tampon A1, contenant de la RNase, y est ajouté. Le tube est vortexé jusqu'à homogénéisation complète, puis 250 µL de tampon A2, correspondant au tampon de lyse, y sont également ajoutés. La solution est mélangée doucement par retournement plusieurs fois, puis est incubée pendant 5 minutes à température ambiante. La solution est neutralisée par ajout de 300 µL de tampon A3, et agitation jusqu'à disparition de la teinte bleutée. Le lysat est alors centrifugé à 11 000 g pendant 10 minutes.

Le surnageant, contenant l'ADN plasmidique, est ensuite chargé dans une colonne contenant une membrane de silice, elle-même placée sur un tube collecteur. Cette colonne est centrifugée une première fois à 11 000g pendant 1 minute. 500 µL de tampon de lavage AW y sont ensuite ajoutés puis la colonne est à nouveau centrifugée à 11 000 g pendant 1 minute. 600 µL de tampon de lavage A4 y sont alors ajoutés, et la colonne est à nouveau centrifugée à 11 000 g pour 1 minute. La colonne est ensuite séchée par 2 centrifugations successives à 11 000 g de 1 puis 2 minutes, respectivement. Finalement, afin d'éluer les plasmides retenus par la membrane de silice, 50 µL d'eau sont ajoutés dans la colonne, puis celle-ci est placée au-dessus d'un tube propre et centrifugée pendant 1 minute à 11 000g. L'éluat est à nouveau chargé dans la colonne puis re-centrifugé 1 minute afin de récupérer la totalité des plasmides. La concentration d'ADN plasmidique récupérée à la fin du processus est dosée au Nanodrop.

### 2.1.6 Réaction de séquence et séquençage

Afin de vérifier que les nouveaux plasmides générés comportent bien la mutation souhaitée, sans insertion d'une mutation non-souhaitée à un autre endroit, un séquençage Sanger est réalisé. Le principe de ce séquençage consiste en une PCR au cours laquelle des dNTPs et des ddNTPs fluorescents, auquel aucun autre dNTP ne peut se lier à la suite, sont incorporés à l'extrémité des brins d'ADN nouvellement formés, permettant ainsi la génération d'amplicons de différentes longueurs. Ainsi, la lecture de la fluorescence émise par chaque ddNTP situé à l'extrémité de ces amplicons par le séquenceur permet de définir la séquence du brin d'ADNc.

Pour les projets *OPRD1* et *GLIS3*, la réaction de séquence a été réalisée à l'aide du BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems / Thermo Fisher Scientific). Brièvement, 3 µL de plasmide muté, à une concentration de 50 à 100 ng/µL, sont déposés dans les puits d'une plaque PCR à 96 puits, autant de fois que le nombre d'amorces qui sont utilisées pour pouvoir séquencer l'entièreté du gène. La liste des amorces de séquençage pour les plasmides pOPRD1 et pGLIS3 sont reprises dans les annexes 4 et 5, respectivement. A chaque puits est ensuite ajouté un mix composé de 0,3 µL de BigDye Terminator v3.1 Ready Reaction Mix, 0,2 µL d'amorces forward ou reverse, 3,6 µL de 5X Sequencing buffer, et 12,4 µL d'eau pure. La réaction de séquence est ensuite lancée dans un thermocycleur, selon le cycle repris dans le tableau 9.

Tableau 9. Cycle de température appliqué au cours de la réaction de séquence.

<b>1X</b>	<b>25X</b>				
96°C	96°C	50°C	60°C	4°C	
1 minute	10 secondes	5 secondes	4 minutes	∞	

Afin de se débarrasser des réactifs provenant de la réaction de séquence, les produits issus de celle-ci sont purifiés par passage à travers des colonnes de résine SEPHADEX G50, par centrifugation à 910-960 g. Le séquençage Sanger est réalisé à l'aide de l'automate 3730xl DNA Analyzer (Applied Biosystems / Thermo Fisher Scientific).

### 2.1.7 Maxipreps et extraction des plasmides

Après avoir identifié une colonie bactérienne ayant intégré la bonne mutation en miniprep, la colonie est à nouveau amplifiée en maxiprep puis l'ADN plasmidique en est extrait. Pour cela, une petite quantité des bactéries issues de la miniprep et conservées en glycérol stock est mise en culture dans 250 mL de milieu LB enrichi en antibiotique, puis est incubée sous agitation à 165 rpm à 37°C pendant la nuit. Il est possible de réaliser un nouveau glycérol stock à ce stade : pour cela, 800 µL de milieu de culture bactérien sont prélevés et ajoutés à 800 µL de glycérol, puis sont conservés à -80°C.

Le lendemain, l'ADN plasmidique est extrait de ces bactéries à l'aide du kit NucleoBond Xtra Maxi Plus (Macherey-Nagel). Concrètement, pour chaque maxiprep, le milieu de culture bactérien est distribué dans plusieurs falcons afin d'être centrifugé à 5000 g pendant 15 minutes, puis le surnageant est éliminé. Le culot de bactérie d'un premier tube est resuspendu dans 12 mL de tampon RES, contenant de la RNase, puis est transvasé dans un autre tube contenant un culot bactérien afin d'y resuspendre ces bactéries, et ainsi de suite jusqu'à resuspension de toutes les bactéries. 12 mL de tampon de lyse LYS y sont ajoutés, puis la solution est mélangée par plusieurs retournements successifs et est incubée 5 minutes à température ambiante. La solution est ensuite neutralisée par ajout de 12 mL de tampon NEU et retournement du tube jusqu'à disparition de la teinte bleutée. Le tube est alors centrifugé à 5000 g durant 10 minutes.

Le surnageant, contenant les plasmides, est ensuite introduit dans un filtre préalablement humidifié à l'aide de 25 mL de tampon d'équilibration EQU, surmontant une colonne contenant une membrane de silice. Une fois le surnageant écoulé, le filtre est de nouveau rincé par ajout de 15 mL de tampon, puis est jeté. La colonne contenant la membrane de silice, sur laquelle se sont liés les plasmides, est lavée avec 25 mL de tampon de lavage. Après écoulement, la colonne est placée au-dessus d'un nouveau tube, et les plasmides sont élués dans celui-ci grâce à l'ajout de 15 mL de tampon ELU.

Finalement, afin de récupérer l'ADN plasmidique pur, celui-ci est précipité par ajout de 10,5 mL d'isopropanol dans le tube, suivi d'une agitation douce et d'une incubation d'au moins 2 minutes à température ambiante. Cette solution est passée à travers un filtre, sur lequel les plasmides vont se lier, puis le filtre est lavé avec 4 mL d'éthanol 70% et séché. Finalement, les plasmides sont élués du filtre par passage de 400 µL d'eau à travers celui-ci. L'éluat est passé

à nouveau dans le filtre pour récupérer la totalité des plasmides retenus. La concentration en ADN plasmidique obtenue à la fin du processus est mesurée à l'aide d'un Nanodrop.

Finalement, la séquence du plasmide obtenu à la fin du processus est vérifiée comme expliqué précédemment au point 2.1.6 « Réaction de séquence et séquençage ».

## 2.2 Modèles cellulaires utilisés

### 2.2.1 Lignée cellulaire HEK293

Pour la plupart de nos expériences de génétique fonctionnelle, la lignée cellulaire Human Embryonic Kidney 293 (HEK293) a été utilisée comme modèle cellulaire. Ces cellules ont été créées en 1973, dans le laboratoire du Professeur Alex van der Eb à Leiden, aux Pays-Bas<sup>340,341</sup>. Il s'agit de cellules rénales embryonnaires, provenant d'un fœtus humain avorté, qui ont été immortalisées par transformation de celles-ci avec de l'ADN fragmenté provenant d'adénovirus de type 5<sup>341,342</sup>. Ces cellules présentent l'avantage d'être facilement cultivables, et sont aisément transfectables<sup>343</sup>.

Ces cellules sont cultivées dans du milieu Dulbecco's Modified Eagle's Medium (DMEM) high glucose, pyruvate (Gibco / Thermo Fisher Scientific), enrichi avec 10 % de sérum fœtal de veau, et 1% de pénicilline/streptomycine, dans des flasques T75 pour celles adhérentes, à une densité de 2,5 millions de cellules par flasque. Elles sont maintenues en culture à 37°C, 5% de CO<sub>2</sub>.

### 2.2.2 Lignée cellulaire EndoCβH1

Pour les tests de sécrétion d'insuline, la lignée cellulaire EndoCβH1 a été utilisée. Il s'agit d'une lignée cellulaire de cellules β-pancréatiques humaines, publiée en 2011 par Philippe Ravassard et al.<sup>344</sup>. Ces cellules ont été générées à partir de cellules β-pancréatiques issues de bourgeons pancréatiques fœtaux, et immortalisées, comme représenté à la figure 33. Brièvement, des bourgeons pancréatiques fœtaux humains ont été transduits avec un vecteur lentiviral exprimant le proto-oncogène Simian vacuolating virus 40 large T antigen (SV40LT). Ces bourgeons ont ensuite été greffés chez des souris atteintes d'un déficit immunitaire combiné sévère (SCID) afin de permettre leur maturation et la différenciation des cellules β pancréatiques, formant alors des insulinomes. Ces insulinomes ont été prélevés puis transduits avec un second vecteur lentiviral, exprimant la transcriptase inverse de la

téломérase humaine (hTERT), puis greffés chez d'autres souris immunodéficientes SCID afin de permettre leur amplification<sup>344</sup>.

Ces cellules sont cultivées dans du milieu Optiβ1 (Univercell Biosolutions, Toulouse, France), dans des flasques T25 préalablement coatées au βcoat (Univercell Biosolutions), à une densité de 2,5 millions de cellules par flasque.

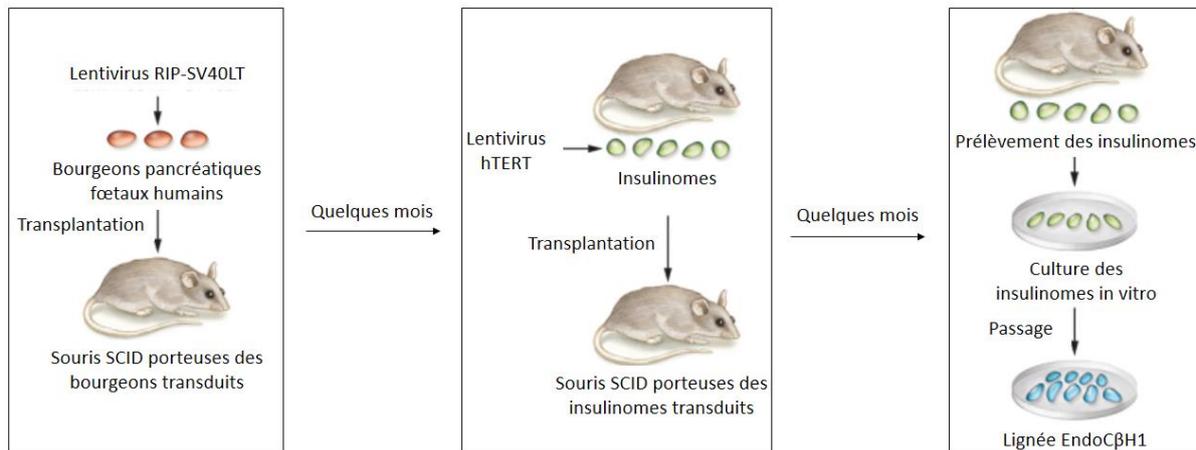


Figure 33. Processus par lequel la lignée de cellules  $\beta$  pancréatiques humaines EndoC $\beta$ H1 a été générée. **SCID** : déficit immunitaire combiné sévère. Source : adapté de Weir and Bonner-Weir, « Finally! A human pancreatic  $\beta$  cell line », *The Journal of Clinical Investigation*, 2011.

### 2.3 Expériences de luciférase

L'utilisation de la luciférase en tant que gène rapporteur est une technique largement utilisée dans le domaine de la biologie moléculaire, permettant d'évaluer l'activité de protéines telles que des facteurs de transcription ou des GPCRs, et dont le principe est repris à la figure 34A. Dans ces expériences, le gène de la luciférase est inséré dans un vecteur plasmidique, sous le contrôle d'une région promotrice ou d'un élément de réponse issu de la région promotrice de gènes ciblés par la protéine à étudier. L'activation de l'élément de réponse ou de la région promotrice par la protéine à étudier induit l'activation de la transcription du gène de la luciférase, suivie de son expression. En présence de  $Mg^{2+}$ , la luciférase est capable de métaboliser son substrat, la luciférine, en oxyluciférine, accompagnée d'un signal bioluminescent (figure 34B). La mesure de ce signal permet d'évaluer l'activité de la luciférase, correspondant à son expression, et ainsi indirectement à l'activité de la protéine à étudier en amont.

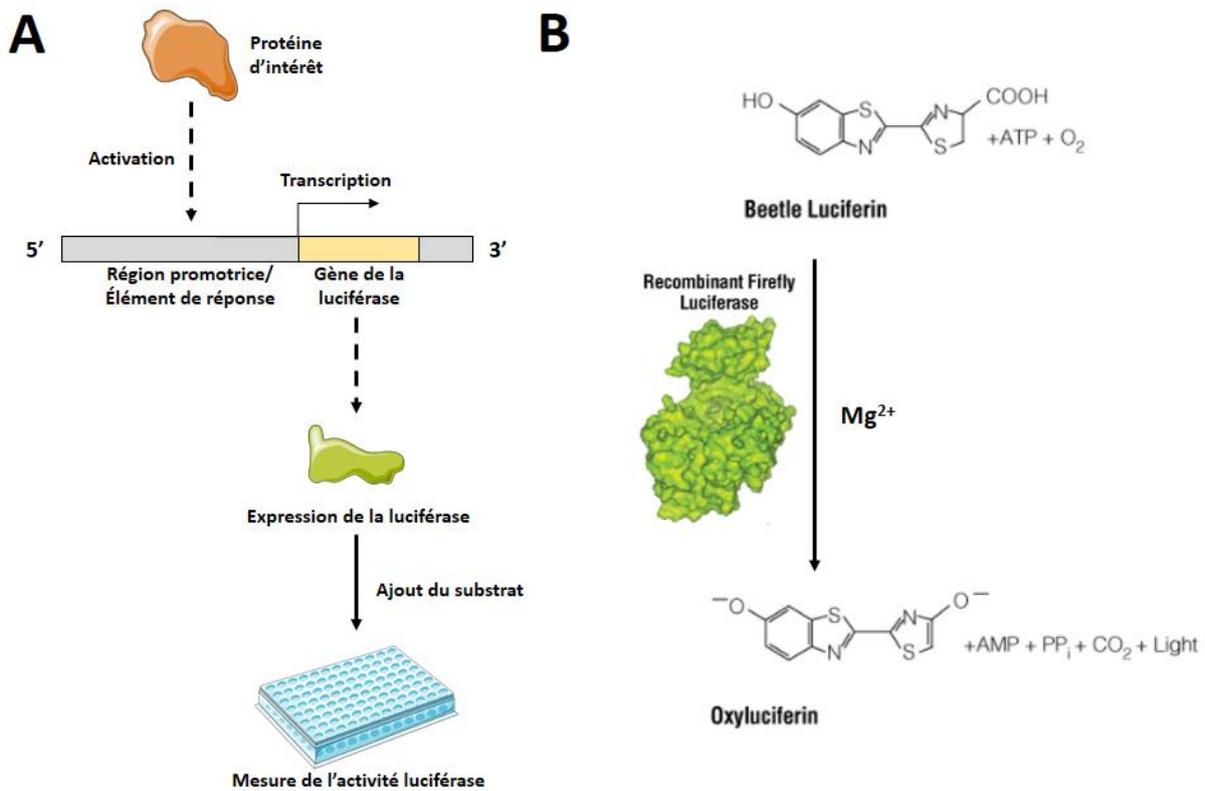


Figure 34. (A) Principe de la luciférase. Source : inspiré du site web <https://www.goldbio.com/articles/article/a-deep-dive-into-the-luciferase-assay-what-it-is-how-it-works-and-more>. (B) Réaction catalysée par la luciférase de luciole (*Photinus pyralis*). Source : adapté de <https://france.promega.com/resources/guides/cell-biology/bioluminescent-reporters>

### 2.3.1 Expériences de SRE-luciférase

Dans le cadre du projet *OPRD1*, afin de déterminer l'impact fonctionnel des différents variants de ce gène sur la fonctionnalité du récepteur DOP, des expériences de SRE-luciférase ont été réalisées. Ces expériences sont basées sur la capacité du récepteur DOP, une fois activé par un agoniste, à activer en aval le Serum Response Element (SRE), situé dans la région promotrice de gènes cibles (figure 35).

Concrètement, les cellules HEK293 sont transfectées en suspension, à l'aide de FuGENE HD (Promega), avec 225 ng/mL de plasmide pOPRD1 contenant le gène *OPRD1* WT ou un des variants à étudier, 150 ng/mL de plasmide pSRE-LUC comportant le gène de la luciférase sous le contrôle de l'élément de réponse SRE, ainsi que 375 ng/mL de plasmide pβgal permettant de normaliser les résultats en fonction de l'efficacité de la transfection. Deux conditions contrôles négatifs sont également réalisées. Pour la première, appelée condition « non-transfectée » (NT), les plasmides pSRE-LUC et pβgal sont transfectés, mais pas le plasmide pOPRD1 ; pour la seconde, appelée condition « STOP », le plasmide pOPRD1 WT est remplacé par le plasmide pOPRD1 STOP, contenant le variant c.3\_4insTAA/p.E2\* générant un récepteur

tronqué. 100 µL de ces cellules, concentrées à  $1 \times 10^6$  cellules/mL, sont ensuite plantées dans les puits d'une plaque 96 puits pré-coatée à la poly-L-lysine.

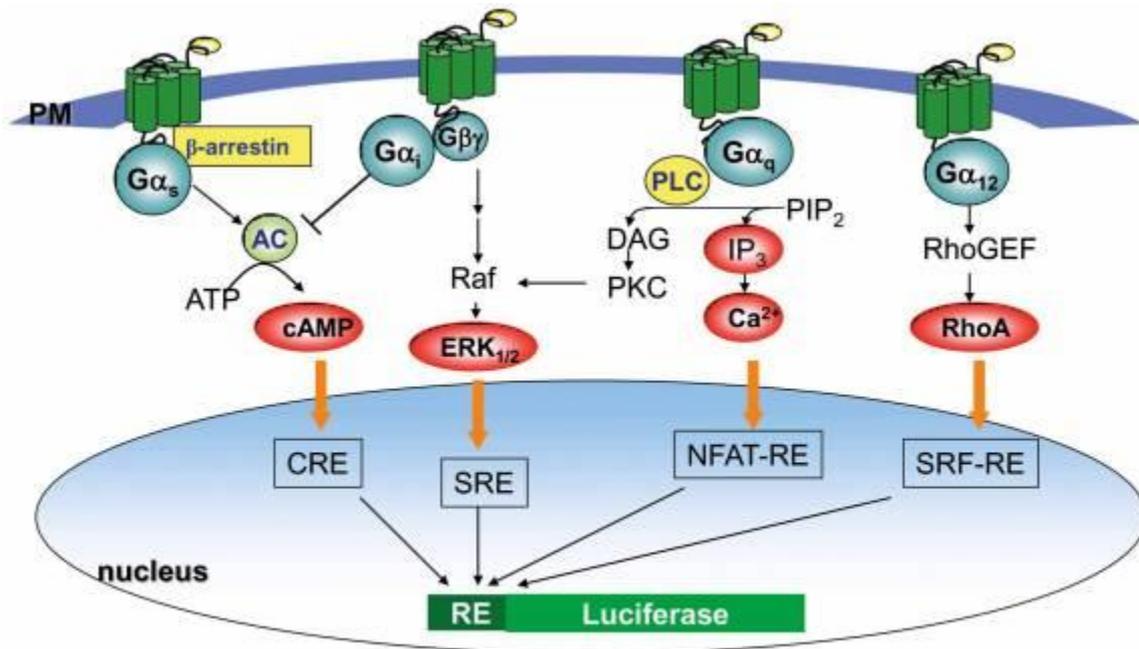


Figure 35. Schéma reprenant les différents éléments de réponse pouvant être couplés à la luciférase en tant que gène rapporteur, en fonction du type de GPCR que l'on souhaite étudier. Source : Cheng et al., « Luciferase Reporter Assay System for Deciphering GPCR Pathways », *Current Chemical Genomics*, 2010.

Le lendemain soir, les cellules sont rincées au Phosphate buffered saline (PBS) 1X puis mises en starving pour la nuit dans du milieu de culture contenant seulement 0,1% de sérum fœtal de veau. 48 heures après la transfection, les cellules sont traitées durant 5 heures avec du diméthylsulfoxyde (DMSO) 0,1%, servant de condition contrôle afin de déterminer l'activité basale de la luciférase, ou avec des concentrations croissantes de 1, 10 et 100 nM d'agonistes sélectifs de DOP [D-Ala<sup>2</sup>]-Deltorphin II (DII) et [D-Pen<sup>2</sup>,D-Pen<sup>5</sup>]Enkephalin (DPDPE) (Tocris, Bristol, Royaume-Uni) dilués dans du milieu de starving. Après 5 heures de traitement, les cellules sont rincées au PBS 1X, puis lysées dans 100 µL de Luciferase Cell Culture Lysis Buffer (Promega) sous agitation pendant 30 minutes. Afin de mesurer l'activité luciférase pour chacune des conditions, 40 µL de lysat sont prélevés, et 25 µL de Luciferase Assay System Reagent (Promega) sont ajoutés à ceux-ci. La luminescence émise suite à la réaction de la luciférase avec son substrat est lue à l'aide d'un luminomètre GloMax (Promega), à une longueur d'onde de 550-570 nm. En parallèle, afin de mesurer l'activité de la β-galactosidase, 100 µL de tampon de β-galactosidase (composé de 0,479 g de NaHPO<sub>4</sub>, 1,060 g de NAH<sub>2</sub>PO<sub>4</sub>–

2 H<sub>2</sub>O, 0,5 mL de KCl 2M, 0,1 mL de MgCl<sub>2</sub> 1M, et complété jusqu'à 100 mL d'eau, auxquels sont ensuite ajoutés du β-mercaptoéthanol 50 mM (3,25 μL/mL) et de l'ortho-nitrophenyl-β-galactoside (ONPG) 4 mM (250 μL/mL)) sont ajoutés à 40 μL de lysat cellulaire. L'activité de la β-galactosidase peut alors être lue à une longueur d'onde de 450 nm à l'aide d'un spectrophotomètre GloMax (Promega).

Chaque expérience de luciférase a été réalisée au minimum 4 fois, et pour chacune d'entre elle, chaque condition a été testée en triplicat technique. Pour chaque point, les mesures de l'activité luciférase, exprimées en unités de luminescence relative, ont été normalisées par les mesures de l'activité β-galactosidase associées. Suite à cette normalisation, pour chaque variant (comprenant également la condition WT), le fold change correspondant à chaque traitement a été calculé en divisant l'activité de la luciférase par la moyenne de l'activité basale de la luciférase, correspondant au traitement au DMSO 0,1%.

L'effet de chaque variant d'*OPRD1*, ainsi que l'effet du gène WT, ont été analysés par un test ANOVA, selon l'équation suivante :  $FC = \beta_0 + \beta_1C + \delta_1V + \delta_2C \times V + \beta_2E + \epsilon$ , où :

- *C* correspond à la concentration d'agoniste DII ou DPDPE,
- *V* correspond au variant étudié, et a pour référence la condition WT,
- *E* correspond au numéro de l'expérience.

### 2.3.2 Expériences de INS-luciférase

Dans le cadre du projet *GLIS3*, afin d'étudier l'effet des différents variants du gène sur la fonctionnalité du facteur de transcription, des expériences de INS-luciférase ont été réalisées. Celles-ci sont basées, d'une part, sur la capacité du facteur de transcription GLIS3 à se lier aux domaines GLISBS situés dans la région promotrice du gène *INS*, et à activer ainsi sa transcription. D'autre part, elles se basent également sur la capacité de GLIS3 à stabiliser la liaison du facteur de transcription MAFA, un second facteur de transcription impliqué dans l'activation de la transcription du gène *INS*, sur la région promotrice de ce gène.

Concrètement, les cellules HEK293 sont transfectées en suspension, à l'aide de la Lipofectamine 2000 (Invitrogen / Thermo Fisher Scientific), avec 200 ng/mL de plasmide pGLIS3 contenant le gène *GLIS3* WT ou un des variants à étudier, 200 ng/mL de plasmide pINS-LUC, contenant le gène de la luciférase sous le contrôle de la région promotrice proximale du

gène *INS*, ainsi que 50 ng/mL de plasmide p $\beta$ gal, contenant le gène de la  $\beta$ -galactosidase dont l'activité nous permettra de normaliser nos résultats selon l'efficacité de la transfection. A ces plasmides sont également ajoutés ou non 50 ng/mL de plasmide pMAFA, contenant le gène *MAFA* : la stabilisation de l'interaction de MAFA sur la région promotrice du gène de l'insuline par GLIS3 permet d'augmenter l'activité luciférase. Deux conditions servant de contrôles négatifs à la transfection sont réalisées : une première condition contrôle « NT », où les cellules sont transfectées avec tous les plasmides sauf le plasmide pGLIS3, ainsi qu'une seconde condition contrôle « p.Y627\* », correspondant à la transfection du plasmide pGLIS3 contenant le variant c.1998TA/p.Y627\*, engendrant l'apparition d'un codon STOP et ainsi l'expression d'une protéine GLIS3 tronquée non-fonctionnelle. 100  $\mu$ L de ces cellules transfectées, à une concentration de  $1 \times 10^6$  cellules/mL, sont ensuite placées dans les puits d'une plaque 96 puits pré-coatée à la poly-L-lysine.

48 heures après la transfection, ces cellules sont lysées, et l'activité de la luciférase et de la  $\beta$ -galactosidase est mesurée de la même façon que décrit précédemment pour le projet *OPRD1* (2.3.1 « Expériences de SRE-luciférase »).

Chaque expérience de luciférase a été réalisée au minimum 4 fois, avec un triplicat technique réalisé à chaque fois pour chacune des conditions. Pour chaque point, les mesures de l'activité luciférase, exprimées en unités de luminescence relative, ont été normalisées par les mesures de l'activité  $\beta$ -galactosidase associées. L'effet de chaque variant de *GLIS3* (comprenant la condition WT), en présence ou en absence de *MAFA*, a été analysé par un test ANOVA.

## 2.4 Extraction d'ARN, rétrotranscription et qPCR

Dans le cadre du projet *GLIS3*, afin de s'assurer que la technique de transfection permette de surexprimer correctement ce gène, des expériences de PCR quantitative (qPCR) ont été réalisées. Cette technique se base sur l'amplification exponentielle d'une portion d'ADN, correspondant au gène d'intérêt présent dans un échantillon, par plusieurs cycles. Le nombre de cycles d'amplification nécessaires pour détecter un signal permet de mesurer de manière semi-quantitative l'expression du gène d'intérêt dans l'échantillon. Ce type d'expérience a également été réalisé dans le cadre du projet *OPRD1*, mais nous y reviendrons plus loin.

#### 2.4.1 Transfection des cellules HEK293

Brièvement, des cellules HEK293 sont transfectées en suspension, à l'aide de Lipofectamine 2000, avec 200 ng/mL de plasmide pGLIS3 contenant le gène *GLIS3* WT ou un des variants à étudier. De manière similaire aux expériences de luciférase, les conditions contrôles négatifs « NT », où le plasmide pGLIS3 n'a pas été transfecté, et « p.Y627\* », engendrant une protéine tronquée non-fonctionnelle, sont réalisées. 2 mL de ces cellules, concentrées à  $1 \times 10^6$  cellules/mL, sont platées dans les puits d'une plaque 6 puits pré-coatée à la poly-L-lysine. 48 heures après la transfection, les cellules sont rincées au PBS 1X, puis grattées dans 1 mL de PBS 1X et centrifugées à 2000 rpm pendant 5 minutes à 4°C afin d'obtenir des culots de cellules. Ces culots sont ensuite stockés à -20°C jusqu'à leur utilisation.

#### 2.4.2 Extraction d'ARN

L'ARN des cellules est ensuite extrait. Pour cela, les cellules sont resuspendues dans 1 mL de TRIzol (Invitrogen / Thermo Fisher Scientific). 150 µL de chloroforme y sont ensuite ajoutés, puis les tubes sont vortexés et incubés sur glace durant 5 minutes. Ils sont ensuite centrifugés à 14 000 rpm pendant 20 minutes, à 4°C. La phase supérieure, contenant l'ARN, est alors récupérée. Un volume équivalent d'isopropanol 100%, permettant de précipiter l'ARN, y est ajouté. Les tubes sont mélangés 2 à 3 fois par retournement, puis sont incubés à -20°C pendant minimum 2 heures. Ces tubes sont ensuite centrifugés durant 20 minutes à 14 000 rpm, à 4°C, afin de précipiter l'ARN. Le surnageant de chaque tube est éliminé, puis 500 µL d'éthanol 75% sont ajoutés aux culots d'ARN afin de les laver, et les tubes sont centrifugés 15 minutes à 14 000 rpm, à 4°C. Après aspiration de l'éthanol 75%, les culots d'ARN sont à nouveau lavés avec de l'éthanol 100%, puis les tubes sont centrifugés 15 minutes à 14 000 rpm, à 4°C. Finalement, l'éthanol 100% est éliminé. Les culots d'ARN sont resuspendus dans de l'eau ultrapure, et la concentration en ARN pour chaque échantillon est mesurée à l'aide d'un Nanodrop.

#### 2.4.3 Rétrotranscription

Afin de pouvoir mesurer l'expression du gène *GLIS3* à partir de l'ARN extrait des cellules, il est nécessaire de le rétrotranscrire en ADNc. Cette rétrotranscription (RT) est réalisée à l'aide du kit High Capacity cDNA Reverse Transcription Kit (Applied Biosystems / Thermo Fisher Scientific). Brièvement, 400 ng d'ARN préalablement extrait sont placés dans les puits d'une plaque PCR à 96 puits, et sont complétés jusqu'à 10 µL avec de l'eau ultrapure. Dans chaque

puits, 10 µL de mix contenant 2 µL de 10X RT Buffer, 0,8 µL de 25X dNTP Mix, 2 µL de 10X RT Random Primers, 1µL de MultiScribe Reverse Transcriptase et 4,2 µL d'eau sont ajoutés. La RT est ensuite lancée selon le cycle suivant : 10 minutes à 25°C, 2 heures à 37°C, et finalement 5 minutes à 85°C. Les ADNc obtenus sont ensuite dilués au 1/8<sup>ème</sup> dans de l'eau, et sont conservés à -20°C.

#### 2.4.4 qPCR

Grâce à l'ADNc obtenu, il est possible de mesurer l'expression des gènes d'intérêt par qPCR. Pour cela, 2 µL d'ADNc dilué au 1/8<sup>ème</sup> sont distribués 2 fois en triplicat dans les puits d'une plaque qPCR 384 puits : un dépôt en triplicat permettra de mesurer l'expression du gène d'intérêt *OPRD1* ou *GLIS3* en fonction du projet, et le second permettra de mesurer l'expression du gène de ménage *RPLP0*. Dans chacun des puits est également ajouté un mix composé de 5 µL d'agent intercalant SYBR Green (Brilliant III Ultra-Fast SYBR Green qPCR Master Mix with Low Rox, Agilent), 0,6 µL d'amorces forward et reverse spécifiques du gène à détecter (tableau 10), et 2,4 µL d'eau. La qPCR est alors réalisée à l'aide de l'appareil ViiA7 (Applied Biosystems / Thermo Fisher Scientific), selon le cycle repris dans le tableau 11.

Tableau 10. Amorces utilisées lors des expériences de qPCR

Gène ciblé	Référence	Séquence des amorces
<i>OPRD1</i>	Integrated DNA Technologies (Coralville, Etats-Unis)	PrimeTime qPCR primers, référence Hs.PT.58.14581911
<i>GLIS3</i>	Integrated DNA technologies	Forward : AAAGGCAGCTGCAACAATCT Reverse : ACCACTCCTGCTTCATGCTT
<i>RPLP0</i>	Integrated DNA technologies	Forward : TACACCTTCCCCTTGCTGA Reverse : CTCTTCCTTGGCTTCAACCT

Tableau 11. Cycle de température appliqué pour les expériences de qPCR.

Hold stage	PCR stage		Melting curve		
1x	40x		1x		
50°C	95°C	60°C	95°C	60°C	95°C
2 minutes	15 secondes	1 minute	15 secondes	1 minute	15 secondes

Comme représenté à la figure 36, lors du premier cycle d'amplification, les 2 brins d'ADN appariés sont séparés par dénaturation. Les amorces forward et reverse peuvent alors s'apparier sur leurs brins complémentaires, permettant l'élongation d'un nouveau brin d'ADN par amorce. A partir du deuxième cycle, uniquement le fragment d'ADN correspondant au

gène à étudier est amplifié. Ainsi, à chaque cycle, la quantité d'ADN correspondant au fragment d'intérêt est doublée. A partir d'une certaine quantité d'ADN amplifiée, le seuil de détection de la fluorescence émise par l'agent intercalant SYBR Green lors de son insertion entre les brins d'ADN appariés est atteint. Le nombre de cycles nécessaires pour atteindre ce seuil, appelé cycle threshold (Ct) permet ainsi d'estimer le niveau d'expression du gène d'intérêt pour chaque échantillon.

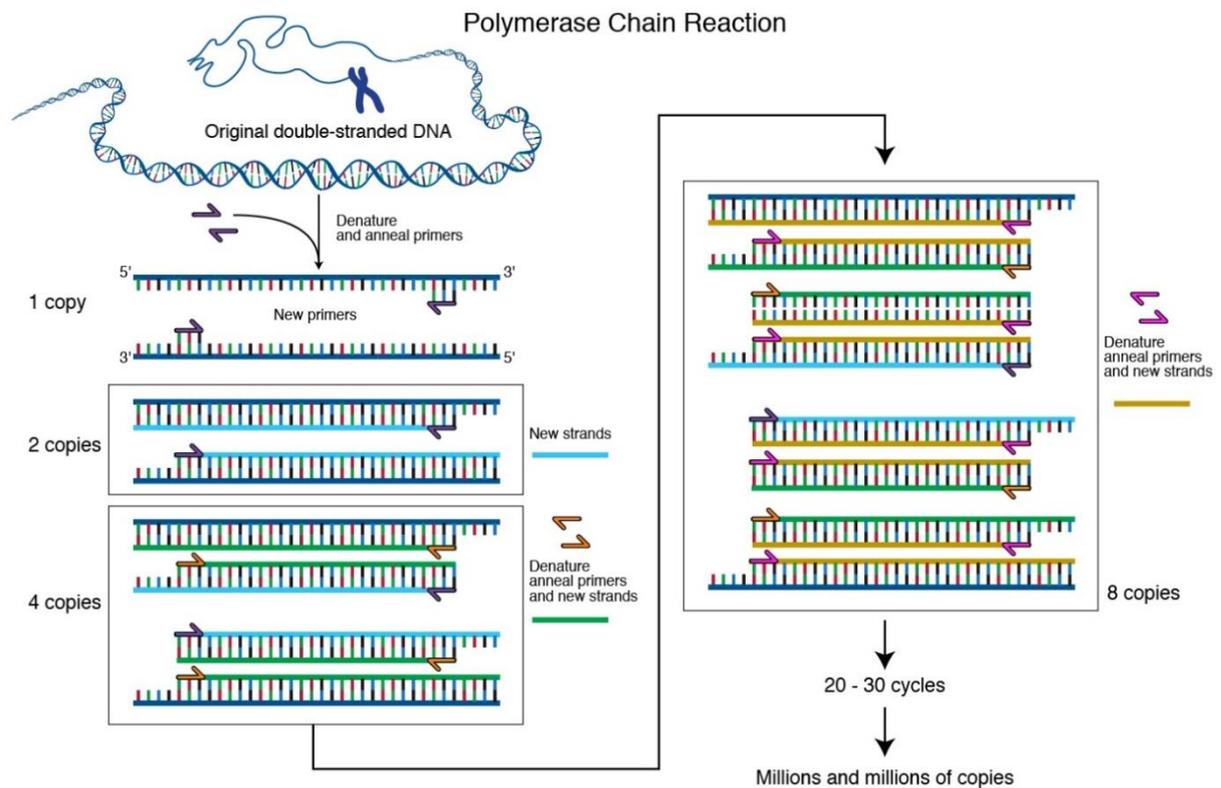


Figure 36. Principe de la PCR. Source : <https://www.genome.gov/genetics-glossary/Polymerase-Chain-Reaction>

A chaque expérience réalisée, les valeurs de Ct obtenues pour le gène d'intérêt (*OPRD1* ou *GLIS3*) pour chacune des conditions étudiées ont été comparées entre elles. Les valeurs obtenues pour le gène de ménage *RPLPO* ont également été comparées entre elles afin de s'assurer que la même quantité d'ADN ait été utilisée dans chacune des conditions.

## 2.5 Western blots

Le western blot est une technique de biologie moléculaire couramment utilisée afin d'évaluer de manière semi-quantitative l'expression d'une protéine d'intérêt. Le principe de cette technique se base sur la séparation des protéines de l'échantillon selon leur poids moléculaire, puis leur transfert sur une membrane, permettant le marquage de la protéine d'intérêt à l'aide d'un anticorps primaire spécifique sur lequel peut venir se lier un anticorps secondaire marqué, suivi de la quantification de son expression (figure 37).

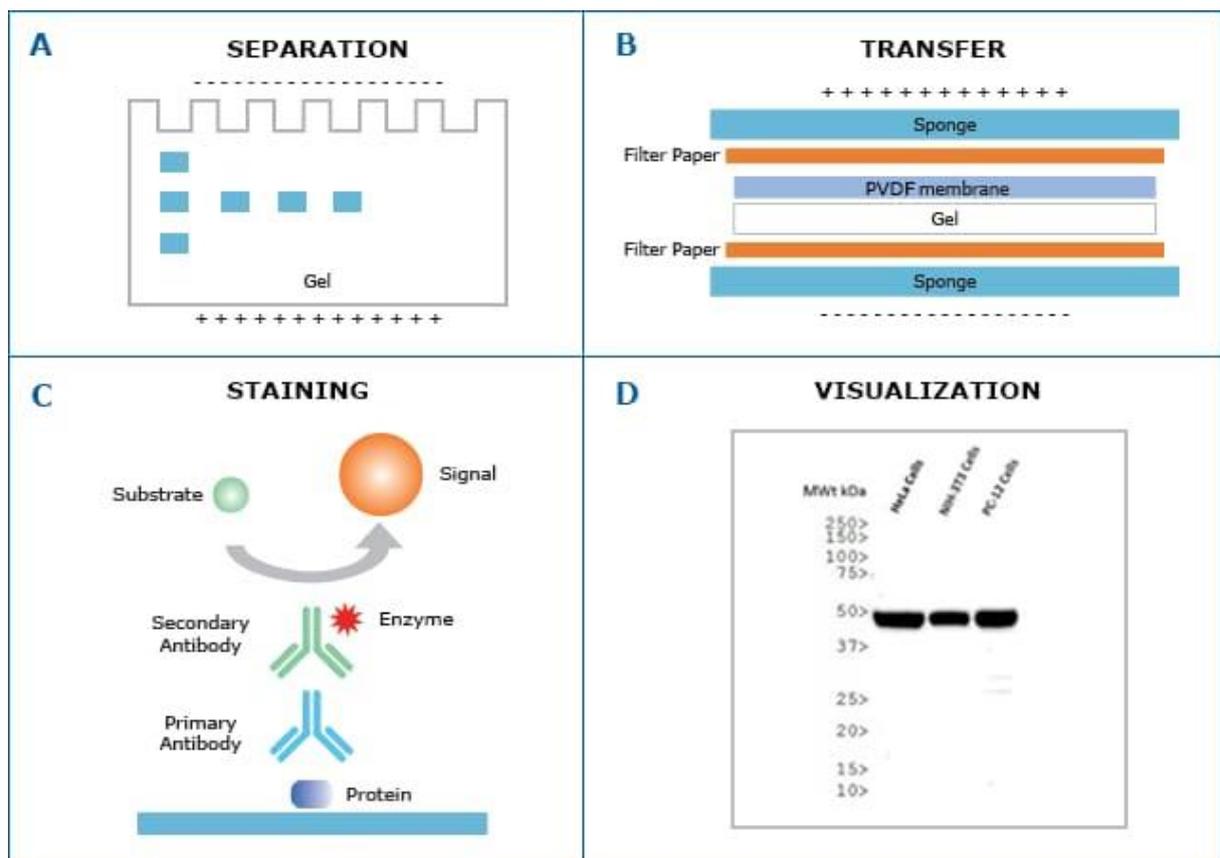


Figure 37. Principe du western blot. Source : site web <https://www.novusbio.com/application/western-blotting>

### 2.5.1 Mesure de l'expression du récepteur DOP

Afin de déterminer si les différents variants du gène *OPRD1* impactent l'expression du récepteur DOP, des expériences de western blotting ont été réalisées. Pour cela, chaque variant du gène *OPRD1* a été surexprimé dans des cellules HEK293.

Brièvement, des cellules HEK293 sont transfectées en suspension, à l'aide de FuGENE HD, avec 225 ng/mL de plasmide pOPRD1 contenant le gène *OPRD1* WT ou un des variants à étudier. Comme pour les expériences de luciférase, une condition « NT », où le plasmide pOPRD1 n'a pas

été transfecté, ainsi qu'une condition « STOP » où le plasmide contenant le variant c.3\_4insTAA/p.E2\* est transfecté, sont également réalisées. 2 mL de cellules transfectées, à une concentration de  $1 \times 10^6$  cellules/mL, sont platées dans les puits d'une plaque 6 puits pré-coatée à la poly-L-lysine. 48 heures après la transfection, ces cellules sont rincées au PBS 1X, puis grattées dans 1 mL de PBS 1X et centrifugées pendant 5 minutes à 2000 rpm, à 4°C afin d'obtenir un culot de cellules. Les culots secs sont conservés à -20°C jusqu'à leur utilisation.

Afin d'en extraire les protéines, les culots cellulaires sont lysés dans 85 µL de tampon RIPA (Thermo Fisher Scientific, Waltham, Etats-Unis), supplémenté en dithiothréitol (DTT) (Invitrogen / Thermo Fisher Scientific), ainsi qu'en inhibiteurs de protéases et de phosphatases (Roche), puis incubés sur glace durant 30 minutes. Les cellules lysées sont ensuite centrifugées pendant 15 minutes à 14 000 rpm, à 4°C. Le surnageant, contenant les protéines, est récupéré.

Suite à l'extraction de protéines, celles-ci sont dosées à l'aide du kit Pierce BCA Protein Assay (ThermoFisher) afin de déterminer la concentration en protéines pour chacun des échantillons. Il s'agit d'un test colorimétrique, basé sur la réduction de l'ion cuivre  $\text{Cu}^{2+}$  en  $\text{Cu}^+$  par les protéines en condition alcaline, suivie de la chélation de l'ion  $\text{Cu}^+$  ainsi formé avec l'acide bicinchoninique (BCA), formant un complexe violet (figure 38). La mesure de l'absorbance suite à la formation de ce complexe permet ainsi de déterminer la concentration en protéines présentes dans l'échantillon.

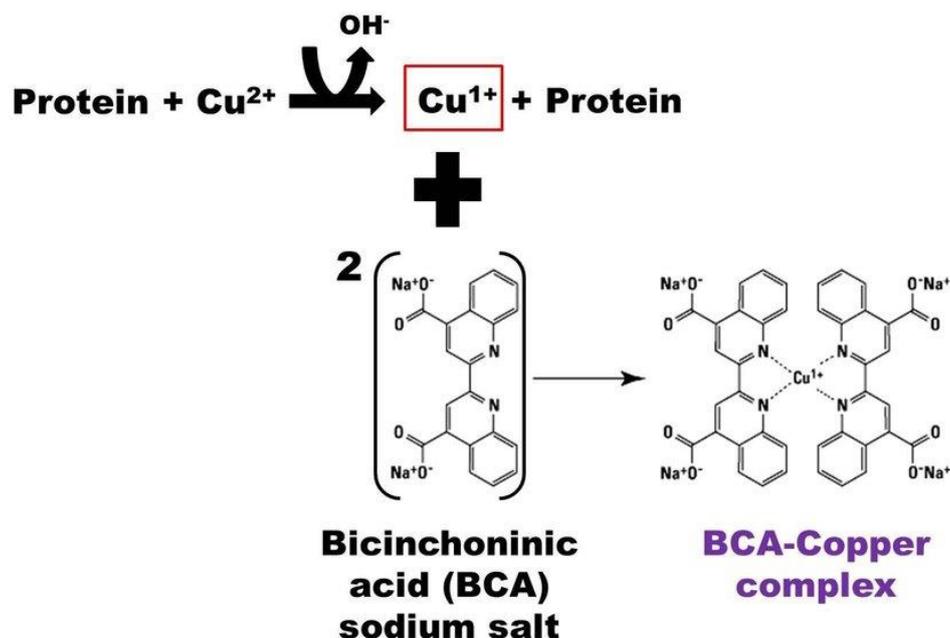


Figure 38. Réaction enzymatique à l'origine du dosage des protéines par la méthode BCA. Source : Thèse d'Angelique Lewies, « An in vitro evaluation of the antibacterial and anticancer properties of the antimicrobial peptide nisin Z », 2018.

Brièvement, dans une plaque 96 puits, 25 µL d'échantillon dilué au 1/10<sup>ème</sup> est chargé en duplicat. En parallèle, 25 µL de chaque point d'une gamme d'étalonnage, composée de concentrations croissantes d'albumine de sérum bovin (BSA) allant de 0 à 2000 mg/mL et réalisée dans du tampon RIPA complet dilué au 1/10<sup>ème</sup>, sont également chargés en duplicat dans la plaque. Dans chaque puits, 200 µL de réactif BCA, composé de 50 volumes de solution A pour 1 volume de solution B, sont ensuite ajoutés. La plaque est alors agitée quelques secondes puis incubée à 37°C pendant 30 minutes. L'absorbance de chaque échantillon est finalement lue à une longueur d'onde de 562 nm, et la concentration en protéines pour chaque échantillon est calculée à partir de la gamme d'étalonnage.

Suite au dosage, pour chaque échantillon, 40 µg de protéines sont mélangés à 7,5 µL de tampon Laemmli 4x réducteur (Alfa Aesar / Thermo Fisher Scientific, Waltam, Etats-Unis), puis les solutions sont complétées avec de l'eau afin d'atteindre un volume final de 30 µL. Les protéines sont dénaturées par incubation de ce mix à 95°C pendant 5 minutes. 25 µL de chaque échantillon, ainsi que 5 µL de marqueur de poids moléculaire PageRuler Prestained Protein Ladder (Thermo Fisher Scientific) sont alors chargés dans les puits d'un gel de polyacrylamide 12% à 15 puits, baigné dans un tampon Tris-Glycine-SDS (Eurobio Scientific, Les Ulis, France). La migration est réalisée durant environ 1 heure 30, d'abord à 75V, afin que les échantillons puissent migrer du gel de concentration vers le gel de séparation, puis à 125V une fois les échantillons passés dans le gel de séparation.

A la suite de la migration, les protéines qui ont migré dans le gel sont transférées sur une membrane de nitrocellulose (GE Healthcare, Chicago, Etats-Unis), à un ampérage constant de 0,25A, durant 1 heure 30. La migration et le transfert corrects des protéines sont vérifiés par coloration de la membrane au Rouge Ponceau. Les sites aspécifiques de la membrane sont ensuite bloqués dans un tampon composé de Tris buffered saline (TBS) 1X (Euromedex, Souffelweyersheim, France), de Tween 20 0,1% (Fisher Scientific / Thermo Fisher Scientific), et de lait 5%, sous agitation, durant une heure à température ambiante, puis la membrane est incubée avec l'anticorps primaire de lapin anti-DOP (Abcam [Cambridge, Royaume-Unis], #ab176324, dilution au 1/1000<sup>ème</sup>) durant la nuit à 4°C. Le lendemain, la membrane est rincée 3 fois au TBS-Tween durant 15 minutes, puis est incubée avec l'anticorps secondaire fluorescent correspondant (Invitrogen / Thermo Fisher Scientific, #SA5-35571, dilution au 1/5000<sup>ème</sup>). Après 3 rinçages, la protéine d'intérêt est révélée à l'aide de l'appareil

Odyssey CLx (LI-COR, Lincoln, Etats-Unis), à une longueur d'onde de 700 nm. Enfin, la membrane est incubée avec de l'anticorps primaire de souris anti- $\beta$ -actine, servant de contrôle de charge (Cell Signaling Technology [Leiden, Pays-Bas], #3700, dilution au 1/5000<sup>ème</sup>) pendant 2 heures à température ambiante ou toute la nuit à 4°C. La membrane est ensuite rincée 3 fois, incubée avec de l'anticorps secondaire fluorescent correspondant (Invitrogen / Thermo Fisher Scientific, #35518, dilution au 1/5000<sup>ème</sup>) ou un anticorps secondaire couplé à la peroxydase de raifort (HRP) (Merck / Sigma-Aldrich [Darmstadt, Allemagne], #NA931, dilution au 1/5000<sup>ème</sup>) puis rincée à nouveau 3 fois. La  $\beta$ -actine est finalement révélée à l'aide de l'Odyssey CLx, à une longueur d'onde de 800 nm. Après conversion des images fluorescentes en noir et blanc, l'intensité de chaque bande est quantifiée à l'aide du programme ImageJ.

Pour chaque variant, au moins 3 expériences indépendantes ont été réalisées. Pour chacun d'entre eux, l'intensité des bandes correspondant au récepteur DOP a été normalisée par l'intensité de la  $\beta$ -actine associée. Des comparaisons deux à deux entre chaque variant et la condition WT ont été réalisées par un test de la somme des rangs de Wilcoxon.

#### 2.5.2 Mesure de l'expression du facteur de transcription GLIS3

Dans le cadre du projet *GLIS3*, afin de s'assurer que les transfections réalisées permettent une surexpression correcte de la protéine, des western blots ont également été réalisés.

Brièvement, les cellules HEK293 sont transfectées en suspension, à l'aide de Lipofectamine 2000, avec 200 ng/mL de plasmide pGLIS3 contenant le gène *GLIS3* WT ou un des variants à étudier. De manière similaire aux expériences de luciférase, les conditions contrôles négatifs « NT » et « p.Y627\* » sont également réalisées. 2 mL de ces cellules, concentrées à  $1 \times 10^6$  cellules/mL, sont plâtées dans les puits d'une plaque 6 puits pré-coatée à la poly-L-lysine. Les cellules sont ensuite récoltées, et les protéines en sont extraites, dosées, dénaturées puis chargées dans un gel de polyacrylamide 10% à 10 puits, comme décrit au point précédent. La suite du western blot est également identique.

Afin de révéler la protéine GLIS3, la membrane est incubée avec de l'anticorps primaire de lapin anti-GLIS3 (Thermo Fisher Scientific, #PA5-41677, dilution au 1/1000<sup>ème</sup>), puis est rincée et incubée avec de l'anticorps secondaire fluorescent ciblant l'anticorps primaire de lapin, et

révélee à l'Odyssey CLx à une longueur d'onde de 700 nm. La  $\beta$ -actine est marquée et révélée comme expliqué au point précédent.

Comme pour le projet *OPRD1*, les western blots ont été réalisés 3 fois afin de confirmer nos observations. L'intensité des bandes correspondant à la protéine GLIS3 a été normalisée par l'intensité des bandes correspondant à la  $\beta$ -actine qui y est associée.

## 2.6 Expériences d'immunofluorescence

L'immunofluorescence est une technique d'immunomarquage, permettant de mettre en évidence l'expression et la localisation d'une protéine d'intérêt. Cette technique repose sur la liaison d'un anticorps spécifique de la protéine d'intérêt sur celle-ci, sur lequel peut venir se lier un second anticorps fluorescent, ciblant le premier. De cette façon, il est possible de localiser la protéine d'intérêt au sein d'une cellule ou d'un tissu.

Dans le cadre du projet *OPRD1*, des immunofluorescences ont été réalisées afin de déterminer si les différents variants du gène induisent une altération de l'expression et/ou de la localisation cellulaire du récepteur DOP. Pour cela, les différents variants d'*OPRD1* ont été surexprimés dans des HEK293.

Brièvement, ces cellules sont transfectées en suspension à l'aide de FuGENE HD avec 225 ng/mL de plasmide pOPRD1 contenant le gène *OPRD1* WT ou un des variants à étudier. Les conditions « NT » et « STOP » sont également réalisées. 2 mL de ces cellules, concentrées à  $5 \times 10^5$  cellules/mL, sont ensuite platées sur des lamelles en verre pré-coatées à la poly-L-lysine, déposées dans les puits d'une plaque 6 puits.

48 heures après la transfection, les cellules transfectées sont fixées au paraformaldéhyde 4% (Fisher Scientific / Thermo Fisher Scientific) durant 20 minutes, rincées au PBS 1X pendant 5 minutes, puis perméabilisées dans une solution composée de PBS 1X – Triton X-100 0,1% pendant 20 minutes. Après 3 rinçages au PBS 1X, les sites aspécifiques des lamelles sont bloqués dans un tampon de blocage composé de PBS 1X, de sérum d'âne 1% (Merck / Sigma Aldrich) et de Triton X-100 0,1% pendant 1 heure. Les cellules sont ensuite incubées avec l'anticorps primaire de lapin anti-DOP (Abcam, #ab176324, dilué au 1/100<sup>ème</sup>) toute la nuit à 4°C. Suite à cela, les cellules sont rincées 3 fois 5 minutes dans la solution de PBS-Triton X-100 0,1%, et incubées avec l'anticorps secondaire correspondant (Invitrogen / Thermo Fisher Scientific, #A-21206, dilué au 1/500<sup>ème</sup>) pendant 1 heure à l'obscurité. Après 3 rinçages de 5

minutes au PBS-Triton X-100 0,1%, l'actine F, permettant de démarquer le contour cellulaire, est marquée à l'aide de phalloïdine conjuguée à un fluorophore (Thermo Fisher Scientific, #21836, dilué au 1/300<sup>ème</sup>). En parallèle, le noyau est marqué au DAPI (Invitrogen / Thermo Fisher Scientific, #R37606). Les cellules sont incubées avec ces solutions pendant 30 minutes à l'abri de la lumière. Enfin, les cellules sont rincées 2 fois pendant 5 minutes au PBS-Triton X-100 0,1%, puis les lames sont montées à l'aide du milieu de montage Prolong Gold Antifade Mountant (Invitrogen / ThermoFisher). La fluorescence émise par chacun des composants est observée à l'aide du microscope Cell Observer Spinning Disk (Zeiss, Oberkochen, Allemagne).

Ces expériences ont été réalisées une seule fois, sauf pour les conditions NT, WT, STOP, et pour les variants ayant présenté un effet au cours de la première expérience, qui ont été reproduites 3 fois.

## 2.7 Tests de stimulation-sécrétion

### 2.7.1 Electroporation des cellules EndoC $\beta$ H1

Afin de déterminer si le gène *OPRD1* est impliqué dans la régulation de l'insulinosécrétion, ce gène a été surexprimé dans les cellules EndoC $\beta$ H1 par électroporation du plasmide pOPRD1 WT à l'aide du Cell Line Nucleofactor Kit V (Lonza, Basel, Suisse).

Concrètement, 2,5 millions de cellules EndoC $\beta$ H1 sont électroporées avec 250 ng/mL de plasmide pOPRD1 WT. Pour cela, les cellules sont resuspendues dans 100  $\mu$ L de solution d'électroporation, à laquelle sont ajoutés 625 ng de plasmide pOPRD1 WT, et les cellules sont ensuite électroporées à l'aide de l'appareil Nucleofactor 2b (Lonza), suivant le programme G-016. En parallèle, une condition contrôle, pour laquelle les cellules ont subi le processus d'électroporation en absence du plasmide pOPRD1 WT, est également réalisée. Suite à cela, les cellules sont resuspendues dans 2,5 mL de milieu Opti $\beta$ 1 (Univercell Biosolutions). 100  $\mu$ L de ces cellules, concentrées à  $1 \times 10^6$  cellules/mL, sont platées dans les puits d'une plaque 96 puits préalablement coatée au  $\beta$ coat (Univercell Biosolutions). 6 heures après l'électroporation, le milieu des cellules est remplacé par un milieu composé à 50% de milieu Opti $\beta$ 1 frais, et à 50% de milieu Opti $\beta$ 1 conditionné (c'est-à-dire provenant de cellules en culture) puis par du milieu Opti $\beta$ 1 frais le lendemain.

En parallèle, afin de vérifier que le gène *OPRD1* soit bien surexprimé dans les cellules, 2 mL de ces cellules électroporées sont platées dans les puits d'une plaque 6 puits

préalablement coatée au  $\beta$ coat. L'extraction d'ARN, la rétrotranscription, ainsi que la qPCR sont réalisées comme décrit précédemment au point 2.4 « Extraction d'ARN, rétrotranscription et qPCR ».

#### 2.7.2 Test de stimulation-sécrétion

Le test de stimulation-sécrétion, réalisé sur les cellules EndoC $\beta$ H1, est illustré à la figure 39. Concrètement, 48 heures après la transfection, les cellules sont mises en starving pour la nuit dans du milieu DMEM no glucose (Gibco / Thermo Fisher Scientific) enrichi avec de la BSA 2% (Roche Molecular Diagnostics), du  $\beta$ -mercaptoéthanol à 50  $\mu$ M, du nicotinamide à 10 mM (Calbiochem / Merck), de la transferrine humaine à 5,5 mg/mL (Sigma-Aldrich), du sélénite de sodium à 6,7 ng/mL (Sigma-Aldrich), de la pénicilline-streptomycine à 200 U/mL (Gibco / Thermo Fisher Scientific), et du glucose à 2,8 mM (Sigma-Aldrich).

Le lendemain, soit 72 heures après la transfection, les cellules sont rincées 2 fois avec du tampon  $\beta$ Krebs (Univercell Biosolution) enrichi avec 0,1% de BSA. Elles sont ensuite incubées en triplicat avec 100  $\mu$ L de solution  $\beta$ Krebs-BSA 0,1% contenant 0,5 mM de glucose (considéré comme la condition « basale »), dans laquelle a été rajouté ou non l'agoniste de DOP DII à 100 nM, durant 1 heure à 37°C. Ce milieu est récupéré, puis les cellules sont incubées avec 100  $\mu$ L de solution  $\beta$ Krebs-BSA 0,1% contenant cette fois 16,7 mM de glucose (considéré comme la condition « stimulée »), contenant ou non l'agoniste DII à 100 nM, durant 1 heure à 37°C. Ce milieu est également récupéré. Finalement, les cellules sont incubées sur glace avec 50  $\mu$ L de tampon de lyse TETG (composé de Tris-HCl pH 8,0 20 mM, de NaCl 137 mM, de Triton X-100 1%, de glycérol 10%, et d'acide egtazique (EGTA) 2 mM), enrichi en inhibiteurs de protéases (Roche), pendant 5 minutes, puis le lysat cellulaire est récupéré. Le milieu des conditions basales et stimulées, ainsi que le lysat cellulaire pour chacune des conditions sont centrifugés pendant 5 minutes à 700 rcf, à 4°C. Le surnageant de chaque solution (70  $\mu$ L pour les milieux des conditions basales et stimulées, 30  $\mu$ L pour le lysat cellulaire) est alors récupéré et conservé à -20°C jusqu'au dosage de l'insuline.

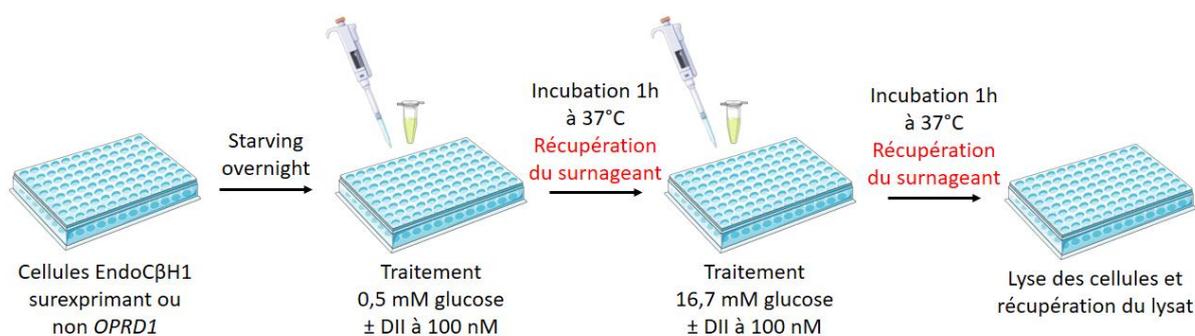


Figure 39. Illustration des différentes étapes du test de stimulation-sécrétion réalisé sur les cellules EndoCβH1.

### 2.7.3 Quantification des cellules électroporées

En parallèle du test de stimulation-sécrétion, il est également important de vérifier que le nombre de cellules présentes dans chaque puits soit approximativement le même pour toutes les conditions. Pour cela, la quantité de cellules est estimée par marquage de celles-ci à l'aide de sondes fluorescentes Draq5 (Thermo Scientific / Thermo Fisher Scientific). Concrètement, juste après la lyse des cellules à la fin du test de stimulation-sécrétion, les cellules présentes dans la plaque dans laquelle s'est déroulée l'expérience sont rincées au PBS 1X, puis sont fixées au paraformaldéhyde 4% pendant 20 à 30 minutes, et rincées 3 fois au PBS 1X. Suite à cela, un mix composé de 125  $\mu$ L de PBS 1X et 0,13  $\mu$ L de Draq5 pour chaque puits est préparé, puis les cellules sont incubées avec 100  $\mu$ L de cette solution de Draq5 durant 30 minutes à température ambiante, à l'abri de la lumière. Finalement, la solution est éliminée des puits, et la fluorescence est lue à l'aide de l'Odyssey CLx, à une longueur d'onde de 800 nm.

### 2.7.4 Dosage de l'insuline par ELISA

Afin de déterminer si l'activation du récepteur DOP a un effet sur la sécrétion d'insuline, la quantité d'insuline sécrétée pour chacune des conditions est dosée par la méthode de dosage immuno-enzymatique ELISA (Enzyme-Linked ImmunoSorbent Assay), à l'aide de l'ELISA Human Insulin Kit (Merckodia, Uppsala, Suède) basé sur le principe de l'ELISA en sandwich (figure 40) : brièvement, l'échantillon contenant la protéine d'intérêt à quantifier est chargé dans une plaque recouverte d'anticorps spécifique de cette protéine. Un second anticorps ciblant la protéine, et sur lequel est lié une enzyme, est ajouté. L'ajout de substrat enzymatique dans le milieu permet de générer un produit coloré. La lecture de l'absorbance pour chaque condition permet ainsi de déterminer la concentration de la protéine d'intérêt présente dans l'échantillon.

Suite au test de stimulation-sécrétion, les surnageants correspondant aux conditions basales et stimulées sont dilués au 1/8<sup>ème</sup> dans de l'eau, tandis le lysat cellulaire est dilué au 1/200<sup>ème</sup>. 25 µL de chaque échantillon dilué, ainsi qu'une gamme d'insuline, sont alors chargés en duplicat dans une plaque 96 puits dont le fond est recouvert d'anticorps anti-insuline humaine. 100 µL d'enzyme peroxydase conjuguée à de l'anticorps ciblant l'insuline humaine sont alors ajoutés à chaque puits, puis la plaque est incubée sous agitation pendant 1 heure à température ambiante. Après l'incubation, chaque puits de la plaque est rincé 6 fois avec 350 µL de tampon de lavage. 200 µL de 3, 3', 5, 5'-tétraméthylbenzidine (TMB), correspondant au substrat de la peroxydase, est alors ajouté à chaque puits, et la plaque est incubée durant 15 minutes sous agitation. Finalement, 50 µL de solution STOP, permettant d'arrêter la réaction enzymatique, est ajouté à chacun des puits. La densité optique de chaque puits est alors lue à l'aide d'un spectrophotomètre GloMax, à une longueur d'onde de 450 nm.

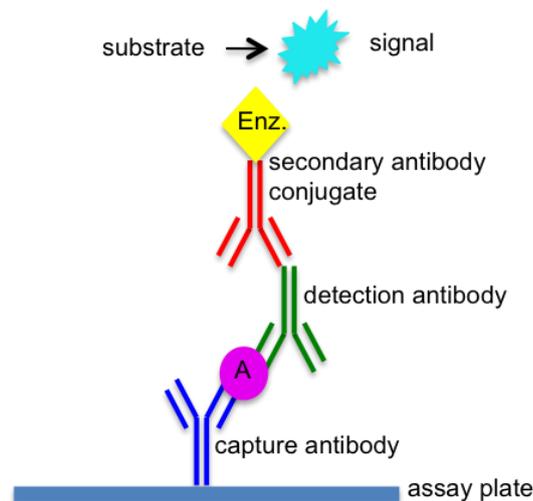


Figure 40. Principe du test ELISA en sandwich. A : protéine d'intérêt, Enz : enzyme conjuguée à l'anticorps secondaire. Source : <https://nordicbiosite.com/blog/elisa-principles-101>

La sécrétion d'insuline en conditions basales, correspondant au pourcentage du contenu intracellulaire en insuline sécrété à bas glucose, a été calculée comme suit :

$$\frac{\text{insuline sécrétée en conditions basales}}{\text{insuline sécrétée en conditions basales} + \text{sécrétée en conditions stimulées} + \text{contenue dans le lysat}} \times 100$$

La sécrétion d'insuline en conditions stimulées, correspondant au pourcentage du contenu intracellulaire en insuline sécrété à haut glucose, a quant à elle été calculée comme suit :

$$\frac{\text{insuline sécrétée en conditions stimulées}}{\text{insuline sécrétée en conditions stimulées} + \text{contenue dans le lysat}} \times 100$$

Le fold change entre la sécrétion stimulée et la sécrétion basale d'insuline a ensuite été calculé pour chacune des conditions.

Les analyses statistiques de ces tests de stimulation-sécrétion ont été réalisées à l'aide du package R, appelé « INsulin Secretion ANalyzEr » (INSANE, R package version 0.1.0). En plus de permettre des analyses statistiques, ce package permet également de réaliser des étapes de contrôle qualité. Le seuil de sécrétion d'insuline, permettant de définir si les cellules EndoCβH1 sécrètent correctement, a été établi à 1,5. Ainsi, les expériences au cours desquelles ce seuil n'a pas été atteint pour les cellules contrôles ont été rejetées. D'autre part, les expériences pour lesquelles le blanc de la gamme d'étalonnage est 1,5 fois supérieur à l'écart interquartile supérieur au 75<sup>ème</sup> percentile, ou est 1,5 fois inférieur à l'écart interquartile inférieur au 25<sup>ème</sup> percentile ont également été rejetées. Le fold change sécrétion stimulée/sécrétion basale pour chaque condition a été analysé suivant un modèle de régression linéaire ajusté aux conditions expérimentales. 5 expériences indépendantes ont été réalisées.

# RESULTATS

# Résultats

## 1 Projet OPRD1

### 1.1 Identification de variants du gène OPRD1 dans l'étude RaDiO

Le séquençage des régions codantes du gène *OPRD1* dans une étude incluant 6971 individus (étude appelée RaDiO) a permis d'identifier 34 variants codants (tableau 12). Parmi ceux-ci, 31 sont des variants rares avec une MAF inférieure à 1%, et ont été identifiés chez 36 individus au total. 14 d'entre eux n'avaient jamais été rapportés auparavant dans la base de données Genome Agregation Database browser (gnomAD, version 2.1.1), et sont donc considérés comme nouveaux. Les 3 autres variants identifiés dans l'étude RaDiO ont une MAF supérieure à 1% dans au moins une des populations répertoriées dans gnomAD, et sont considérés comme des variants fréquents. Tous les variants identifiés dans cette étude sont des variants faux-sens, sauf le variant p.E355\*, qui est un variant non-sens générant un codon STOP prématuré.

### 1.2 Prédiction *in silico* de l'effet fonctionnel des variants du gène OPRD1

L'effet délétère de ces variants a été évalué dans un premier temps *in silico*, à l'aide de la méthode Rare Exome Variant Ensemble Learner (REVEL). Cette méthode est basée sur la compilation de 18 scores de prédiction de pathogénicité grâce à l'utilisation de pas moins de 13 outils de prédiction différents<sup>345</sup>. Un score de pathogénicité peut alors être attribué à chaque variant : on considère qu'en dessous du seuil de 50%, un variant est probablement bénin, tandis qu'au-dessus de ce seuil, le variant est probablement délétère. En appliquant cette méthode pour les variants du gène *OPRD1*, 6 variants ont été prédits comme étant délétères : il s'agit des variants p.D95N, p.Q105L, p.N131D, p.M132I, p.M142L et p.R160G (tableau 12). Si ce seuil est augmenté à 75%, seul les variants rares p.D95N et p.R160G sont considérés comme délétères.

### 1.3 Etudes d'association entre les variants rares du gène OPRD1 et les traits métaboliques associés au DT2

Afin de déterminer si les variants rares du gène *OPRD1* sont associés au DT2, des études d'association ont été réalisées entre l'agrégation de variants rares (c'est-à-dire l'ensemble des variants rares identifiés dans ce projet) et le DT2 et plusieurs traits

métaboliques associés, parmi lesquels l'hyperglycémie, le surpoids, l'obésité, l'IMC, le niveau de HDL-cholestérol et de triglycérides, et la pression artérielle systolique et diastolique.

Tableau 12. Variants du gène *OPRD1* (NM\_000911.4) identifiés dans l'étude RaDiO. Chr : chromosome, MAF : fréquence d'allèle mineur, gnomAD : Genome Agregation Database Browser (version 2.1.1), MAC : minor allele count (correspondant au nombre d'individus de la cohorte porteurs de la mutation).

Chr	Position (Hg19)	Mutation	MAF dans gnomAD	MAC dans RaDiO	Score de prédiction de pathogénicité <i>in silico</i>
1	29138936	c.41C>G, p.P14R	0	1	0,016
1	29138936	c.41C>T, p.P14L	0,00018	1	0,067
1	29138965	c.70C>T, p.P24S	0,000032	1	0,071
1	29138975	c.80G>T, p.C27F	0,91	commun	0,10
1	29138990	c.95C>G, p.A32G	0	1	0,070
1	29139002	c.107G>A, p.G36E	0,0000063	1	0,075
1	29139021	c.126G>C, p.S42R	0	3	0,037
1	29139049	c.154A>G, p.I52V	0,0066 (0,074 chez les Africains)	28	0,16
1	29139079	c.184G>C, p.V62L	0	1	0,19
1	29185481	c.243G>T, p.K81N	0	1	0,23
1	29185521	c.283G>A, p.D95N	0,0000080	1	0,75
1	29185525	c.287C>T, p.A96V	0,000016	1	0,20
1	29185552	c.314A>T, p.Q105L	0	1	0,54
1	29185629	c.391A>G, p.N131D	0	1	0,51
1	29185634	c.396G>C, p.M132I	0,0000040	1	0,66
1	29185662	c.424A>T, p.M142L	0,000012	1	0,53
1	29185716	c.478C>G, p.R160G	0	1	0,76
1	29185785	c.547A>G, p.I183V	0,00018	1	0,035
1	29185813	c.575G>A, p.R192Q	0,000045	1	0,083
1	29189334	c.658T>C, p.F220L	0,000028	1	0,16
1	29189337	c.661G>T, p.A221S	0,000028	1	0,37
1	29189385	c.709C>A, p.L237M	0	1	0,21
1	29189401	c.725G>A, p.S242N	0	1	0,096
1	29189419	c.743G>T, p.G248V	0	1	0,45
1	29189425	c.749A>G, p.K250R	0,00013	1	0,15
1	29189457	c.781C>T, p.R261C	0,0000081	1	0,44
1	29189516	c.840C>G, p.F280L	0	1	0,39
1	29189601	c.925G>A, p.A309T	0,015	3	0,11
1	29189642	c.966C>G, p.D322E	0	1	0,27
1	29189665	c.989G>T, p.R330L	0,000032	1	0,33
1	29189739	c.1063G>T, p.E355*	0	1	NA
1	29189758	c.1082C>A, p.T361N	0,000057	1	0,075
1	29189770	c.1094G>A, p.G365D	0,00021	3	0,15
1	29189773	c.1097C>T, p.P366L	0,000039	3	0,066

Ces analyses ont été réalisées suivant la méthode MiST, et ont été ajustées à l'âge, au sexe, à l'origine ethnique et à l'IMC pour les analyses associées au DT2, à l'hyperglycémie, aux niveaux de HDL-cholestérol et de triglycérides, et à la pression artérielle systolique et diastolique, et ajustées à l'âge, au sexe, et à l'origine ethnique pour les analyses associées au

surpoids, à l'obésité, et à l'IMC. Certains critères d'exclusion ont été appliqués : ainsi, les analyses associées aux niveaux de HDL-cholestérol et de triglycérides ont été réalisées uniquement chez les individus qui ne sont pas sous traitement hypolipémiant, tandis que les analyses associées à la pression artérielle systolique et diastolique n'ont été réalisées que chez les patients qui ne sont pas traités pour l'hypertension, afin d'éviter tout biais associé aux traitements.

La méthode MiST permet de calculer un score  $\pi$ , correspondant à l'effet moyen du cluster de variants, et un score  $\tau$ , correspondant à l'hétérogénéité du cluster. L'analyse ayant démontré que le cluster de variants rares du gène *OPRD1* présente peu d'hétérogénéité, seule la valeur *P* du score  $\pi$  a été prise en compte pour déterminer s'il existe des associations significatives ou non avec certains traits métaboliques.

Comme représenté dans le tableau 13, l'agrégation de variants rares ne présente aucune association avec le DT2 ni avec les différents traits métaboliques qui y sont associés.

Tableau 13. Etudes d'association réalisées entre le burden de variants rares du gène *OPRD1* et le DT2 et les traits métaboliques associés, suivant la méthode MiST. **N** : nombre d'individus inclus dans l'analyse, **OR** : odds ratio, **CI** : intervalle de confiance,  **$\pi$**  : effet moyen des variants, **SE** : erreur standard,  **$P_{\pi}$**  : valeur *P* du score  $\pi$ , **DT2** : diabète de type 2, **IMC** : indice de masse corporelle, **HDL** : HDL-cholestérol, **TG** : triglycérides, **PAS** : pression artérielle systolique, **PAD** : pression artérielle diastolique.

Trait	N	OR [CI]	$\pi \pm SE$	$P_{\pi}$
Hyperglycémie ou DT2	6745	0,80 [0,37-1,7]	-	0,56
Surpoids ou obésité	6211	0,93 [0,48-1,9]	-	0,83
IMC (kg/m <sup>2</sup> )	6971	-	0,13 $\pm$ 0,10	0,21
HDL (mmol/L)	6008	-	0,12 $\pm$ 0,085	0,16
TG (mmol/L)	6051	-	-0,098 $\pm$ 0,092	0,29
PAS	5257	-	5,0 $\pm$ 0,5	0,19
PAD	4693	-	0,21 $\pm$ 2,5	0,93

#### 1.4 Identification des variants GoF et LoF du gène *OPRD1*

Afin de pouvoir affiner les études d'association réalisées entre les variants rares du gène *OPRD1* et les différents traits métaboliques étudiés, nous avons décidé de réaliser des expériences de génétique fonctionnelle permettant de caractériser l'effet fonctionnel de chacun des variants, c'est-à-dire s'ils causent une augmentation (GoF) ou une diminution (LoF) de l'activité du récepteur DOP en réponse à ses agonistes.

A cette fin, des expériences de luciférase, basées sur la capacité du récepteur DOP à activer la voie de signalisation ERK et l'élément de réponse SRE en aval de celle-ci ont été réalisées. Pour cela, des cellules HEK293 ont été transfectées avec le plasmide pOPRD1 contenant le gène *OPRD1* WT ou un des variants à étudier, ainsi qu'avec les plasmides pSRE-LUC, dans lequel le gène de la luciférase est sous le contrôle de l'élément SRE, et pβgal, contenant le gène de la β-galactosidase permettant de normaliser les résultats selon l'efficacité de la transfection. Deux conditions contrôles, appelées « NT » où le plasmide pOPRD1 n'a pas été transfecté (figure 41A), ainsi que la condition « STOP », où le plasmide pOPRD1 contient une version mutée du gène (p.E2\*) générant un codon STOP prématuré (figure 41B), ont également été réalisées. 48 heures plus tard, ces cellules surexprimant le gène *OPRD1* WT ou muté ont été traitées avec des concentrations croissantes des agonistes spécifiques de DOP DII ou DPDPE, allant de 1 à 100 nM. Après 5 heures de traitement, les cellules ont été lysées et l'activité de la luciférase a été révélée, et normalisée par les résultats de la β-galactosidase correspondants.

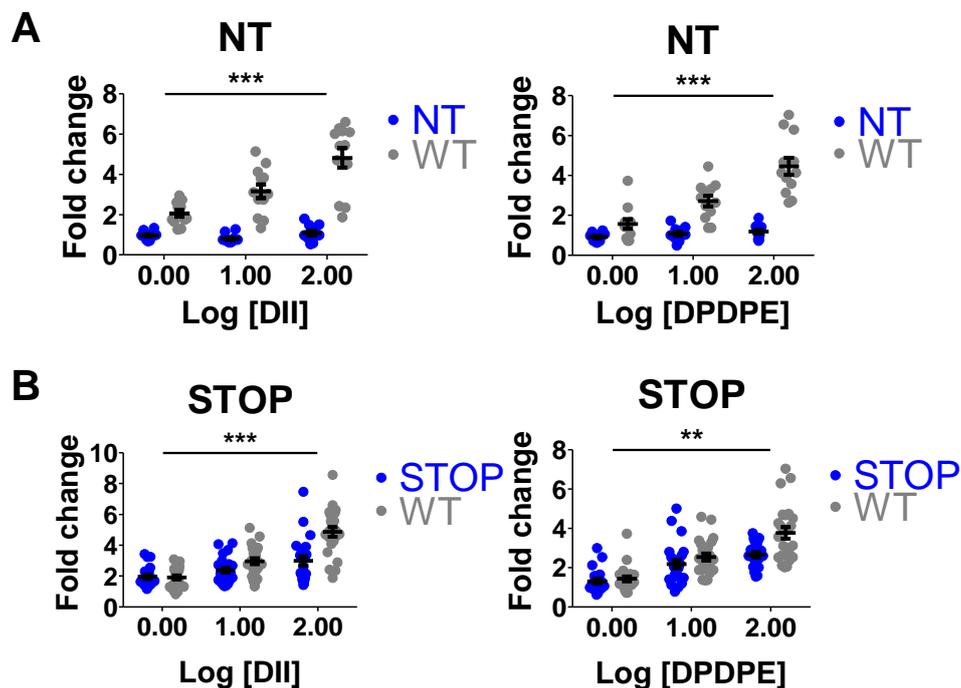


Figure 41. Activité luciférase mesurée en conditions « NT » (A) et « STOP » (B), en réponse aux agonistes de DOP DII et DPDPE. Ces graphiques représentent, pour chacune des conditions, le fold change de l'activité de la SRE-luciférase (normalisée par les valeurs de β-galactosidase) en réponse à des concentrations en DII et en DPDPE de 1, 10 et 100 nM, par rapport à l'activité basale du récepteur DOP WT. L'effet des conditions « NT » et « STOP » sur l'activité luciférase a été analysé selon le modèle ANOVA suivant :  $FC = \theta_0 + \theta_1C + \delta_1V + \delta_2C \times V + \theta_2E + \epsilon$ , avec C correspondant à la concentration de l'agoniste, V correspondant au variant étudié, et E correspondant au numéro d'identification de l'expérience. Ces graphiques représentent la moyenne  $\pm$  SEM des fold change d'au moins 4 expériences indépendantes. \*\* $P \leq 0,01$ , \*\*\*  $P \leq 0,001$ . DII : [D-Ala<sup>2</sup>]-Deltorphin II, DPDPE : [D-Pen<sup>2</sup>,D-Pen<sup>5</sup>]-Enkephalin, NT : non-transfecté, WT : wild-type.

Ces expériences de luciférase ont mis en évidence que parmi les 34 variants identifiés dans l'étude RaDiO, 7 d'entre eux entraînent une augmentation de l'activité luciférase en réponse aux agonistes DII et DPDPE, par rapport à l'activité de pOPRD1 WT : il s'agit des variants rares p.P24S, p.A32G, p.I183V, p.K250R, p.T361N et p.P366L, ainsi que du variant p.I52V, qui est fréquent chez les individus d'origine africaine selon gnomAD. Ceux-ci ont ainsi été définis comme étant des variants GoF (figure 42). En parallèle, les 12 variants rares p.P14R, p.G36E, p.K81N, p.D95N, p.A96V, p.N131D, p.M132I, p.R160G, p.R192Q, p.A221S, p.G248V et p.D322E entraînent une diminution de l'activité luciférase par rapport à la condition WT, en réponse aux agonistes DII et DPDPE, et ont donc été définis comme étant des variants LoF (figure 43). Les variants restants, n'ayant pas démontré d'activité statistiquement différente de celle du récepteur DOP WT en réponse aux agonistes DPDPE ou DII, ou présentant une différence seulement pour l'un des 2 traitements agonistes, ont été considérés comme neutres (annexe 6).

Les résultats obtenus par expérience de luciférase ont ensuite été comparés aux prédictions de pathogénicité réalisées pour chaque variant par l'outil REVEL. En utilisant un seuil de pathogénicité de 50%, REVEL présente une faible sensibilité, par prédiction d'un effet délétère seulement pour 33% des variants LoF, et aucune prédiction concernant les variants GoF. Cependant, cet outil présente une bonne spécificité, qui s'élève à 86%.

### 1.5 Impact des variants du gène *OPRD1* sur l'expression et la localisation cellulaire du récepteur DOP

En plus de l'étude de l'activité du récepteur par expériences de luciférase, nous avons également souhaité déterminer si les différents variants du gène *OPRD1* pouvaient être associés à une altération de l'expression ou de la localisation cellulaire du récepteur DOP. A cette fin, des expériences de western blot et d'immunofluorescence ont été réalisées.

Dans un premier temps, l'expression du récepteur DOP a été mesurée semi-quantitativement par western blot. Pour ce faire, la surexpression du gène *OPRD1* WT ou muté, ainsi que les conditions contrôles « NT » et « STOP », ont été réalisées dans des HEK293 (figure 44). 48 heures plus tard, ces cellules ont été lysées, et des western blots ont été réalisés à partir des extraits protéiques obtenus afin de quantifier l'expression du récepteur DOP et de la  $\beta$ -actine, utilisée comme contrôle de charge.

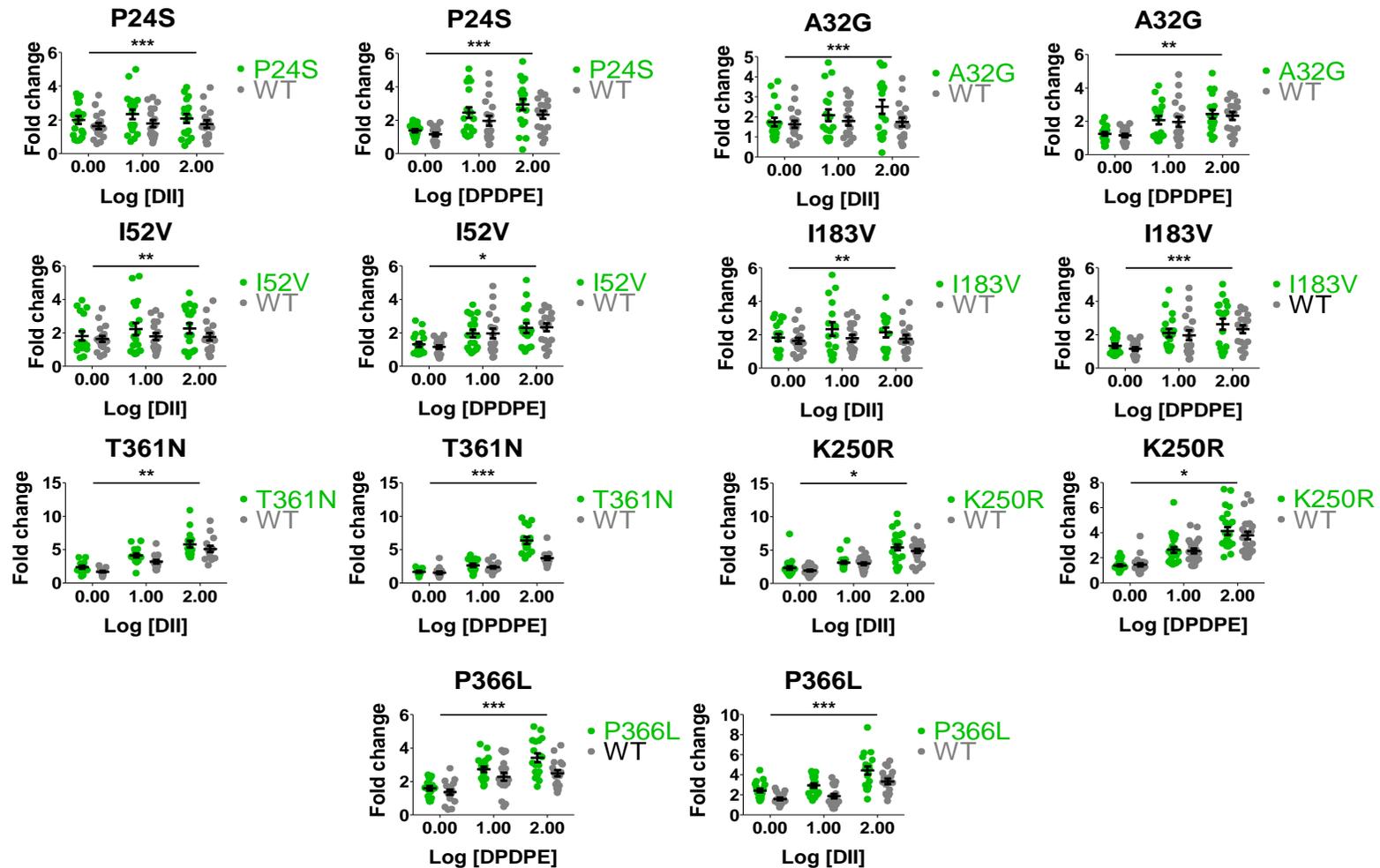


Figure 42. Activité luciférase des variants GoF du gène OPRD1, en réponse aux agonistes de DOP DII et DPDPE. Ces graphiques représentent, pour chaque variant, le fold change de l'activité de la SRE-luciférase (normalisée par les valeurs de  $\beta$ -galactosidase) en réponse à des concentrations en DII et en DPDPE de 1, 10 et 100 nM, par rapport à l'activité basale du récepteur DOP WT. L'effet de chaque variant a été analysé selon le modèle ANOVA suivant :  $FC = \beta_0 + \beta_1C + \delta_1V + \delta_2C \times V + \beta_2E + \epsilon$ , avec C correspondant à la concentration de l'agoniste, V correspondant au variant étudié, et E correspondant au numéro d'identification de l'expérience. Les graphiques représentent la moyenne  $\pm$  SEM des fold change d'au moins 4 expériences indépendantes. \* $P \leq 0,05$ , \*\* $P \leq 0,01$ , \*\*\* $P \leq 0,001$ . DII : [D-Ala<sup>2</sup>]-Deltorphin II, DPDPE : [D-Pen<sup>2</sup>,D-Pen<sup>5</sup>]-Enkephalin, WT : wild-type.

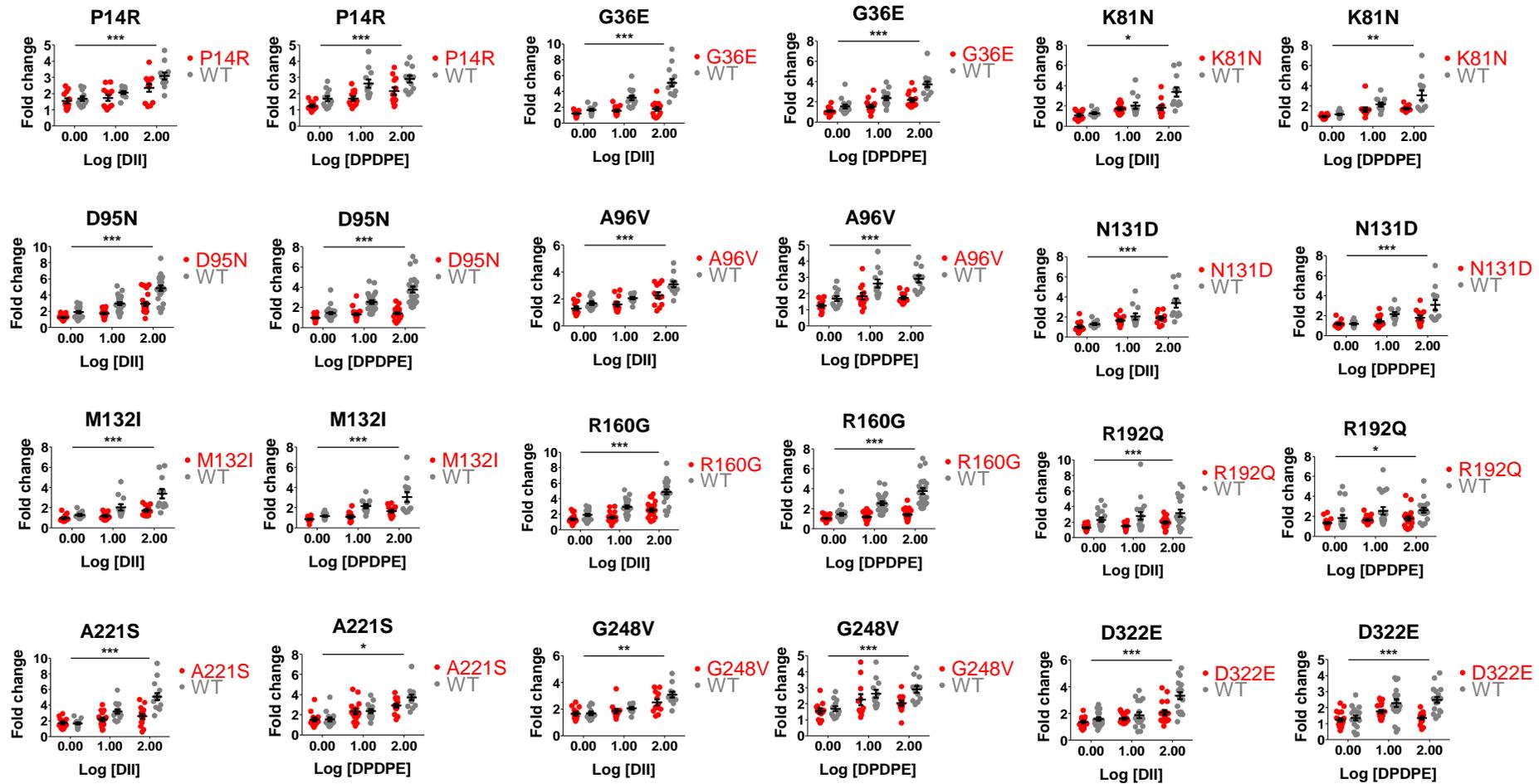


Figure 43. Activité luciférase des variants LoF du gène OPR1, en réponse aux agonistes de DOP DII et DPDPE. Ces graphiques représentent, pour chaque variant, le fold change de l'activité de la SRE-luciférase (normalisée par les valeurs de  $\beta$ -galactosidase) en réponse à des concentrations en DII et en DPDPE de 1, 10 et 100 nM, par rapport à l'activité basale du récepteur DOP WT. L'effet de chaque variant a été analysé selon le modèle ANOVA suivant :  $FC = \beta_0 + \beta_1C + \delta_1V + \delta_2C \times V + \beta_2E + \epsilon$ , avec C correspondant à la concentration de l'agoniste, V correspondant au variant étudié, et E correspondant au numéro d'identification de l'expérience. Les graphiques représentent la moyenne  $\pm$  SEM des fold change d'au moins 4 expériences indépendantes. \* $P \leq 0,05$ , \*\* $P \leq 0,01$ , \*\*\* $P \leq 0,001$ . DII : [D-Ala<sup>2</sup>]-Deltorphin II, DPDPE : [D-Pen<sup>2</sup>,D-Pen<sup>5</sup>]-Enkephalin, WT : wild-type.

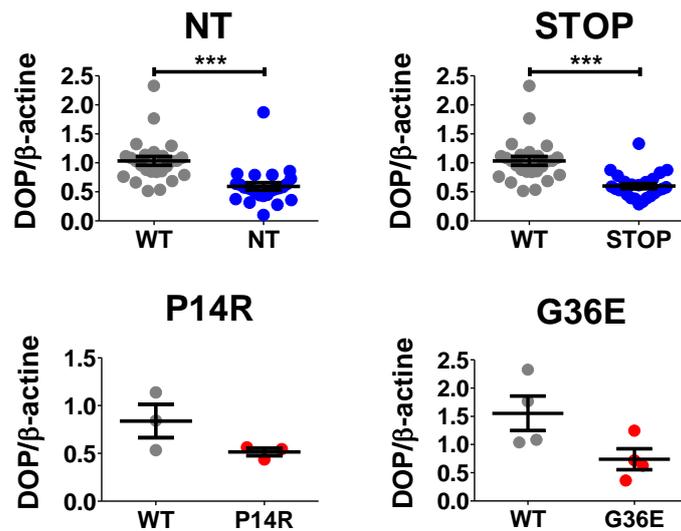


Figure 44. Expression du récepteur DOP en conditions « NT » et « STOP », ainsi que les variants p.P14R et p.G36E. Pour chaque condition, l'intensité obtenue pour chaque bande correspondant au récepteur DOP a été normalisée par celle de la  $\beta$ -actine associée, puis comparée à l'expression du récepteur DOP en condition WT par un test de la somme des rangs de Wilcoxon. Ces graphiques représentent la moyenne de l'expression normalisée du récepteur DOP  $\pm$  SEM d'au moins 3 expériences indépendantes. \*\*\* $P \leq 0,001$ . NT : non-transfecté, WT : wild-type.

Ces expériences de western blotting ont permis de démontrer que la grande majorité des variants sont exprimés à un niveau similaire au récepteur DOP WT (annexe 7). Cependant, les variants rares LoF p.P14R et p.G36E semblent avoir un niveau d'expression plus faible que le récepteur DOP WT, sans atteindre cependant un effet significatif (figure 44).

En parallèle, afin d'investiguer la localisation cellulaire des différents mutants, mais également pour vérifier l'expression du récepteur, des expériences d'immunofluorescence ont été réalisées. Brièvement, les différents variants du gène *OPRD1* ont été surexprimés dans des cellules HEK293. Les conditions contrôles « NT » et « STOP » ont également été réalisées (figure 45). 48 heures plus tard, le récepteur DOP, la F-actine permettant de délimiter les cellules, ainsi que les noyaux cellulaires ont été marqués et observés par immunofluorescence.

Ces expériences d'immunofluorescence ont ainsi permis de démontrer que la plupart des variants du gène *OPRD1*, identifiés dans l'étude RaDiO, sont exprimés à la membrane cellulaire à un niveau similaire au récepteur DOP WT (annexe 8). Cependant, concernant les variants rares LoF p.G14R et p.G36E, ces expériences ont permis de confirmer visuellement leur expression plus faible. De plus, ces 2 variants semblent ne pas être exprimés à la membrane de la cellule, mais plutôt retenus dans la région intracellulaire (figure 45). La perte de fonction de ces variants pourrait donc être expliquée, au moins en partie, par leur expression plus faible et par l'altération de leur localisation cellulaire.

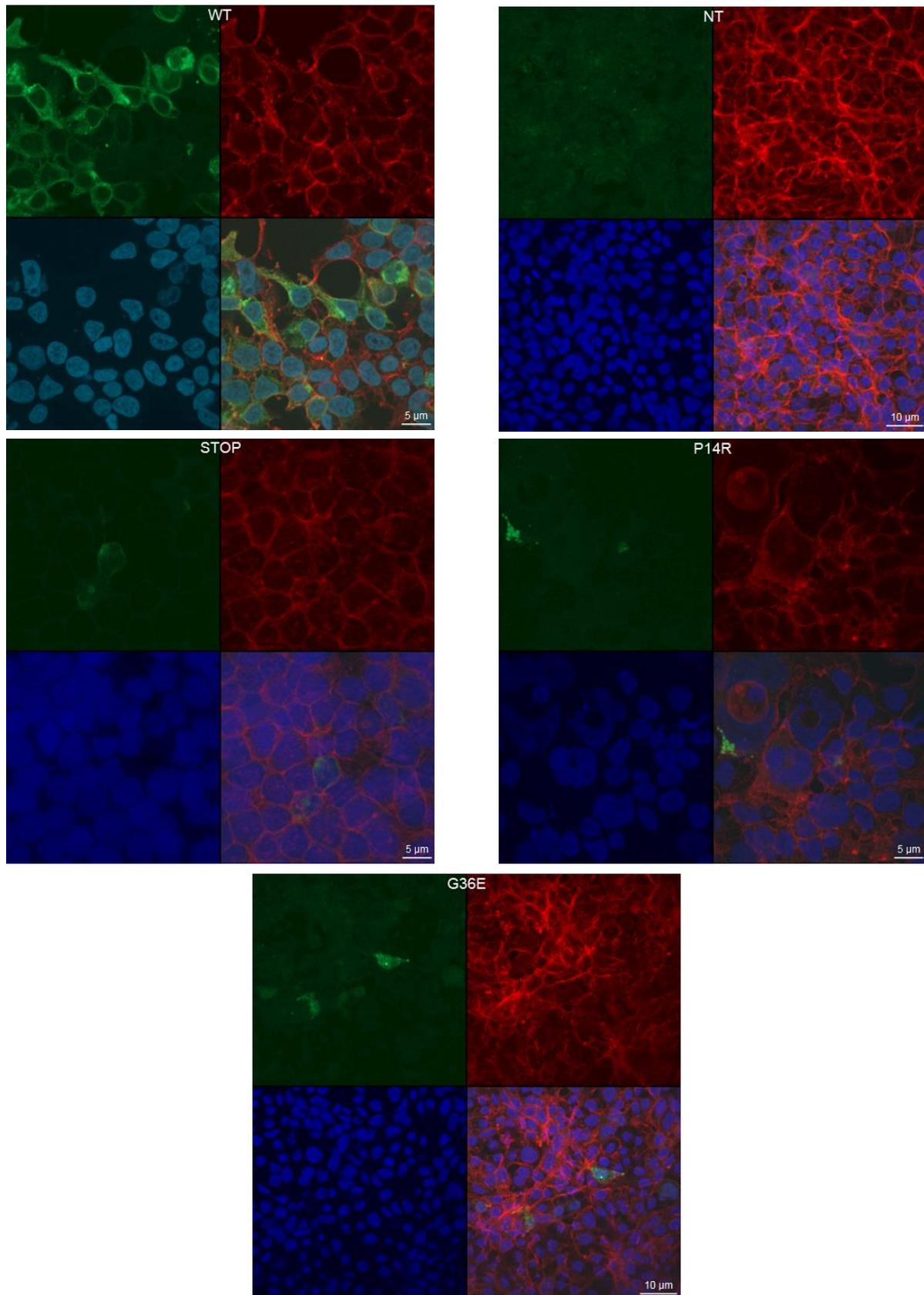


Figure 45. Immunofluorescence du récepteur DOP dans des cellules HEK293 surexprimant (ou non) le gène OPRD1 WT, ou les mutations « STOP », p.P14R et p.G36E. Le récepteur DOP est marqué en vert, l'actine F en rouge, et les noyaux cellulaires en bleu. Ces images sont représentatives des résultats obtenus pour 3 expériences indépendantes. NT : non-transfecté, WT : wild-type.

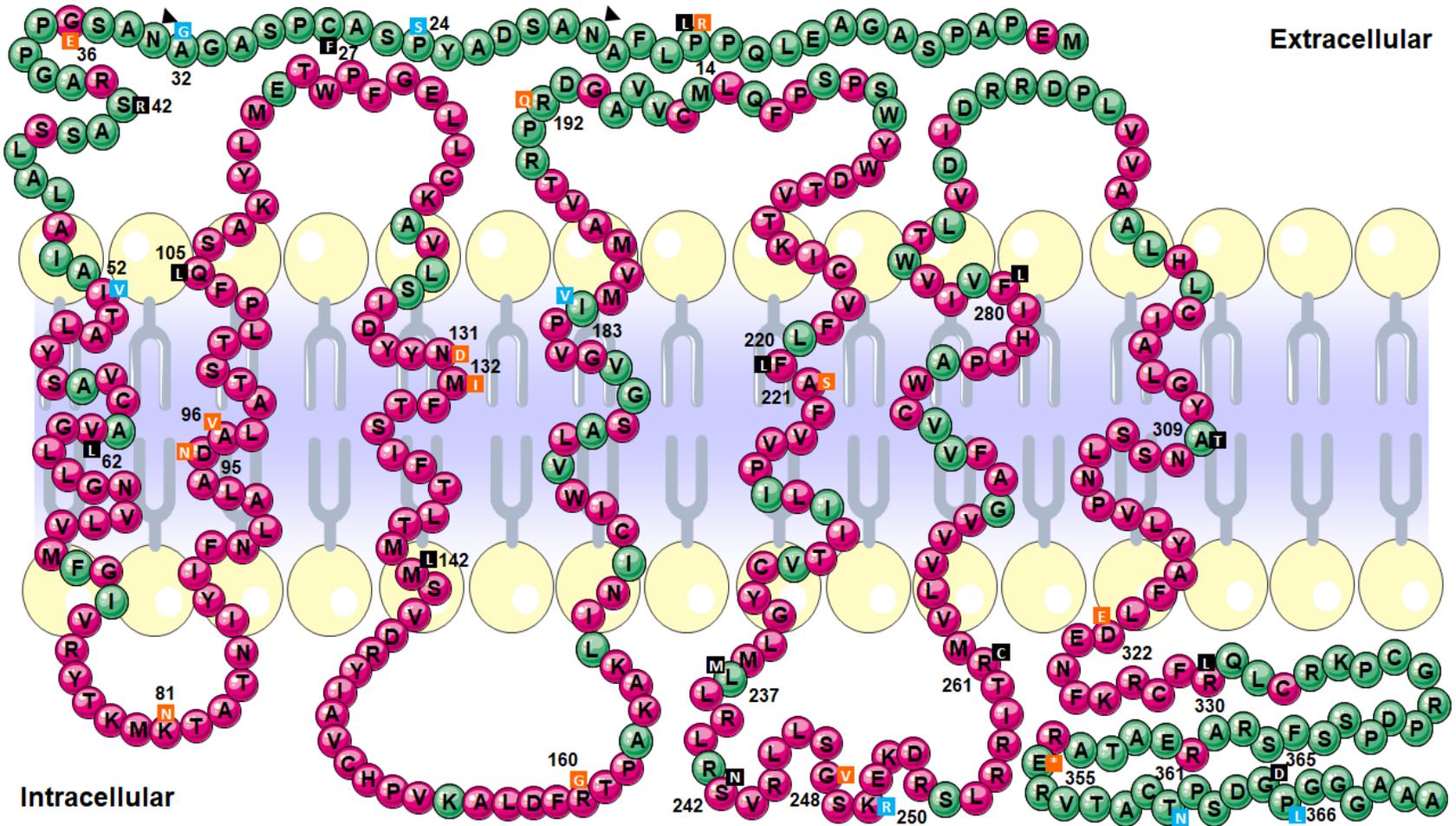


Figure 46. Représentation de la structure du récepteur DOP, ainsi que de la position des différents variants du gène OPRD1 identifiés dans l'étude RaDiO. Les bulles roses correspondent aux acides aminés conservés, c'est-à-dire présentant une forte homologie de séquence interspèce selon Ensembl (humain versus chimpanzé, gibbon à joues blanches, babouin olive, rat, souris, chien, platy, poule, grenouille, poisson-zèbre), les bulles vertes correspondent aux acides aminés non-conservés. Les triangles noirs indiquent les sites putatifs de glycosylation. Les variants sont représentés par les petits carrés, dont la couleur correspond à leur effet fonctionnel selon les expériences de luciférase : bleu pour les variants GoF, orange pour les variants LoF et génétiquement nuls, noir pour les variants neutres. La lettre au centre de ces carrés correspond au changement d'acide aminé, tandis que le numéro à proximité correspond à la position de l'acide aminé modifié dans la séquence du récepteur DOP.

## 1.6 Position des 34 variants du gène *OPRD1* dans la structure du récepteur DOP

La figure 46 représente la structure du récepteur DOP, la conservation ou non de chaque acide aminé, ainsi que la position de chaque variant et son effet sur l'activité du récepteur. Les acides aminés touchés par les variants LoF sont significativement plus conservés que les acides aminés touchés par les variants GoF : 83% des variants LoF sont retrouvés dans des résidus conservés *versus* 29% pour les variants GoF ( $P = 0,027$ ). L'agrégation de variants GoF, quant à elle, tend à être plus importante dans le domaine N-terminal extracellulaire et le domaine C-terminal intracellulaire de la protéine ( $P = 0,085$ ), où les acides aminés sont significativement moins conservés que dans les régions transmembranaires et les boucles (16% de conservation dans les régions transmembranaires et les boucles *versus* 79% dans les régions terminales intra- et extracellulaires,  $P = 5,7 \times 10^{-22}$ ).

## 1.7 Etudes d'association entre l'agrégation de variants GoF ou LoF du gène *OPRD1* et les traits métaboliques associés au DT2

### 1.7.1 Effet des variants rares du gène *OPRD1* identifiés dans l'étude RaDiO

Grâce aux résultats des expériences de luciférase, les différents variants rares du gène *OPRD1* identifiés dans le cadre de l'étude RaDiO ont été classés selon leur effet fonctionnel. Ainsi, un cluster composé des 6 variants rares définis comme étant GoF, et un cluster composé des 12 variants rares définis comme étant LoF ainsi que du variant p.E355\* considéré comme génétiquement nul (car induisant un codon STOP prématuré) ont été formés. Les caractéristiques cliniques des individus porteurs de variants GoF ou LoF sont reprises à l'annexe 9. Des études d'association ont ensuite été réalisées entre chacun des clusters de variants rares GoF et LoF, et le DT2, l'hyperglycémie, le surpoids, l'obésité, l'IMC, les niveaux de HDL-cholestérol et de triglycérides, et la pression artérielle systolique et diastolique, selon la méthode MiST. Comme précédemment, ces analyses ont été ajustées ajustée à l'âge, au sexe, à l'origine ethnique et à l'IMC pour les analyses s'intéressant au DT2, à l'hyperglycémie, aux niveaux de HDL-cholestérol et de triglycérides, ainsi que pour la pression artérielle systolique et diastolique, et ajustées à l'âge, au sexe, et à l'origine ethnique pour les analyses portant sur le surpoids, l'obésité, et l'IMC. Les mêmes critères d'exclusion ont également été appliqués : les analyses associées aux niveaux de HDL-cholestérol et de triglycérides ont été réalisées seulement chez les individus qui ne sont pas sous traitement hypolipémiant, tandis que les analyses associées à la pression artérielle systolique et diastolique n'ont été menées

que chez les individus non traités pour l'hypertension. Seule la valeur  $P$  du score  $\pi$  a été prise en compte pour déterminer si une association significative entre les variants et le trait étudié est présente car nous n'avons pas identifié d'hétérogénéité au sein des clusters.

#### 1.7.1.1 Effets des variants rares GoF

Dans l'étude RaDiO, hormis un participant, tous les individus porteurs de mutations rares GoF du gène *OPRD1* sont non-obèses. Les individus non-obèses sont donc proportionnellement plus nombreux à être porteurs de mutations GoF que les individus obèses (0,083% chez les individus obèses *versus* 0,19% chez les individus non-obèses). Cependant, étant donné le faible nombre d'individus porteurs de ce type de variants, la puissance statistique n'est pas assez importante pour pouvoir établir une association significative entre les variants GoF et l'obésité. Les études d'association ont en revanche permis de mettre en évidence une association significative entre l'agrégation de variants rares GoF et une amélioration des niveaux de lipides sériques : ces variants sont associés à une augmentation des niveaux de HDL-cholestérol ( $\pi = 0,33 \pm 0,17$  ;  $P = 0,049$  ; tableau 14) et à une diminution des niveaux de triglycérides ( $\pi = -0,37 \pm 0,18$  ;  $P = 0,046$  ; tableau 14).

#### 1.7.1.2 Effet des variants rares LoF

Dans l'étude RaDiO, hormis un participant, tous les individus porteurs de mutations rares LoF du gène *OPRD1* sont en surpoids ou obèses. Les individus obèses porteurs d'une mutation LoF sont même significativement plus nombreux que les individus non-obèses (0,31% chez les individus obèses *versus* 0,038% chez les individus non-obèses ; OR = 11 ;  $P = 0,0054$  ; tableau 15). En outre, ces variants LoF sont significativement associés à un IMC plus important ( $\pi = 0,35 \pm 0,17$  ;  $P = 0,044$  ; tableau 15). D'autre part, l'agrégation des variants LoF tend à diminuer le risque de DT2 : alors que 0,063% d'individus diabétiques sont porteurs de ce type de mutations, 0,25% des individus normoglycémiques en sont porteurs (OR = 0,23 ;  $P = 0,054$  ; tableau 15).

#### 1.7.2 Effets du variant fréquent p.I52V dans l'AMP T2D Knowledge Portal

Afin de répliquer les résultats obtenus concernant les variants rares GoF du gène *OPRD1* de l'étude RaDiO, des études d'association ont également été réalisées pour le variant p.I52V, fréquent dans la population d'origine africaine, à partir des données issues de la base de données AMP T2D Knowledge Portal qui permet d'analyser plusieurs milliers d'individus

africains atteints de DT2 ou témoins (ce qui n'était pas possible dans RaDiO). Les études d'association se sont ainsi intéressées à l'effet de ce variant dans la population d'origine africaine ou toutes ethnies confondues du Portal, sur plusieurs traits métaboliques globalement similaires à ceux étudiés pour l'étude RaDiO : le DT2, l'IMC, les niveaux de HDL- et LDL-cholestérol et de triglycérides, et la pression sanguine. Ces études d'association ont été réalisées selon un modèle de régression logistique pour le DT2 ou linéaire additif pour les autres traits métaboliques étudiés. Ces études ont été ajustées à l'âge, au sexe, à l'IMC et à l'ethnie concernant le DT2 et les niveaux de lipides sériques, et ajustées à l'âge, au sexe et à l'ethnie concernant l'IMC. Des critères d'exclusion ont également été appliqués : notamment, afin de ne pas inclure de biais lié aux traitements, les niveaux de lipides sériques et la pression sanguine ont été analysés uniquement chez les individus ne souffrant pas de DT2, et seuls des individus de plus de 50 ans ont été compris dans l'étude cas-contrôles du DT2.

Ces études d'association ont permis de démontrer que le variant fréquent GoF p.I52V est associé à une élévation significative du risque de DT2 chez les 34 812 individus étudiés, toutes ethnies confondues (OR = 1,7 ;  $P = 6,1 \times 10^{-5}$  ; tableau 14). Cet effet est probablement dirigé par les résultats obtenus pour les individus d'origine africaine (OR = 2,0 ;  $P = 3,6 \times 10^{-6}$  ; tableau 14). De plus, ce variant GoF est aussi associé à une diminution de l'IMC ( $\beta = -0,027 \pm 0,0093$  ;  $P = 0,0038$  ; tableau 14), ainsi qu'à une amélioration du profil lipidique, par une élévation nominale des niveaux de HDL-cholestérol chez les individus d'origine africaine ( $\beta = 0,095 \pm 0,049$  ;  $P = 0,055$  ; tableau 14), et une diminution des niveaux de LDL-cholestérol ( $\beta = -0,37 \pm 0,12$  ;  $P = 0,0026$  dans la population générale ;  $\beta = -0,48 \pm 0,16$  ;  $P = 0,0026$  dans la population africaine ; tableau 3) et de triglycérides ( $\beta = -0,13 \pm 0,054$  ;  $P = 0,018$  dans la population générale ;  $\beta = -0,11 \pm 0,060$  ;  $P = 0,069$  dans la population africaine ; tableau 14).

### 1.7.3 Effets des variants rares LoF identifiés dans la UK Biobank

Dans le but de répliquer les résultats obtenus concernant les variants LoF issus de l'étude RaDiO, des études d'association ont été réalisées dans la UK Biobank, regroupant les données de séquençage d'exons réalisé chez plusieurs milliers d'individus. Ainsi, 12 variants rares génétiquement nuls du gène *OPRD1* (c'est-à-dire non-sens ou décalants), ainsi que 7 variants déjà identifiés dans l'étude RaDiO comme LoF, ont été détectés chez 187 242 participants (tableau 16). Des études d'association ont ensuite été réalisées entre l'agrégation de ces 19 variants considérés comme LoF et le DT2 et les traits métaboliques associés.

Tableau 14. Etudes d'association réalisées entre l'agrégation de 6 variants rares GoF du gène OPRD1 et le DT2 et les traits métaboliques associés issus de l'étude RaDiO, suivant la méthode MiST, ou entre le variant p.I52V et le DT2 et les traits métaboliques associés issus de l'AMP T2D Knowledge Portal, suivant un modèle de régression logistique (pour le DT2) ou linéaire additif (pour les autres traits étudiés). **N** : nombre d'individus inclus dans l'analyse, **SD** : déviation standard, **IQR** : écart interquartile, **OR** : odds ratio, **CI** : intervalle de confiance,  **$\pi$  ou  $\beta$**  : effet moyen des variants, **SE** : erreur standard,  **$P_\pi$  et  $P_\beta$**  : valeur P du score  $\pi$  ou  $\beta$ , **DT2** : diabète de type 2, **IMC** : indice de masse corporelle, **HDL** : HDL-cholestérol, **TG** : triglycérides, **LDL** : LDL-cholestérol, **PAS** : pression artérielle systolique, **PAD** : pression artérielle diastolique.

Trait	Etude	N	%, moyenne $\pm$ SD ou médiane [IQR] du trait chez les porteurs	%, moyenne $\pm$ SD ou médiane [IQR] du trait chez les non-porteurs	OR [CI]	$\pi$ ou $\beta \pm$ SE	$P_\pi$ ou $P_\beta$
DT2 ou hyperglycémie	RaDiO		50%	47%	0,89 [0,15–4,7]	-	0,89
	AMP-Tous	34 812	67%	57%	1,7 [1,3–2,2]	-	$6,1 \times 10^{-5}$
	AMP-Africains	4966	69%	54%	2,0 [1,5–2,6]	-	$3,6 \times 10^{-6}$
IMC (kg/m <sup>2</sup> )	RaDiO		25 [24–27]	26 [23–29]	-	$0,040 \pm 0,22$	0,86
	AMP-Tous	17 073	-	-	-	$-0,027 \pm 0,0093$	0,0038
	AMP-Africains	1988	-	-	-	$-0,011 \pm 0,013$	0,41
HDL (mmol/L)	RaDiO		$1,8 \pm 0,39$	$1,5 \pm 0,45$	-	$0,33 \pm 0,17$	0,049
	AMP-Tous	7427	-	-	-	$0,048 \pm 0,046$	0,30
	AMP-Africains	492	-	-	-	$0,095 \pm 0,049$	0,055
TG (mmol/L)	RaDiO		0,55 [0,54–0,92]	1,1 [0,76–1,6]	-	$-0,37 \pm 0,18$	0,046
	AMP-Tous	7974	-	-	-	$-0,13 \pm 0,054$	0,018
	AMP-Africains	492	-	-	-	$-0,11 \pm 0,060$	0,069
LDL (mmol/L)	RaDiO	-	-	-	-	-	-
	AMP-Tous	5288	-	-	-	$-0,37 \pm 0,12$	0,0026
	AMP-Africains	489	-	-	-	$-0,48 \pm 0,16$	0,0026
PAS (mmHg)	RaDiO		$120 \pm 17$	$130 \pm 15$	-	$-1,3 \pm 8,3$	0,88
	AMP-Tous	6551	-	-	-	$-0,91 \pm 2,3$	0,70
	AMP-Africains	493	-	-	-	$-0,79 \pm 3,3$	0,81
PAD (mmHg)	RaDiO		$70 \pm 11$	$78 \pm 9,4$	-	$-2,3 \pm 5,7$	0,69
	AMP-Tous	6552	-	-	-	$2,6 \pm 1,4$	0,066
	AMP-Africains	493	-	-	-	$2,5 \pm 2,0$	0,22

Tableau 15. Etudes d'association réalisées entre l'agrégation de 13 variants rares LoF du gène OPRD1 identifiés dans l'étude RaDiO ou l'agrégation de 17 variants rares LoF identifiés dans la UK Biobank, et le DT2 et les traits métaboliques associés, suivant la méthode MiST. **N** : nombre d'individus inclus dans l'analyse, **SD** : déviation standard, **IQR** : écart interquartile, **OR** : odds ratio, **CI** : intervalle de confiance,  $\pi$  : effet moyen des variants, **SE** : erreur standard,  $P_\pi$  : valeur P du score  $\pi$ , **DT2** : diabète de type 2, **IMC** : indice de masse corporelle, **HDL** : HDL-cholestérol, **TG** : triglycérides, **PAS** : pression artérielle systolique, **PAD** : pression artérielle diastolique.

Trait	Etude	N	%, moyenne $\pm$ SD ou médiane [IQR] du trait chez les porteurs	%, moyenne $\pm$ SD ou médiane [IQR] du trait chez les non-porteurs	OR [CI]	$\pi \pm$ SE	$P_\pi$
DT2 ou hyperglycémie	RaDiO	6745	15%	47%	0,23 [0,033–0,95]	-	0,054
	UK Biobank	171 667	0%	3,4%	$3,2 \times 10^{-5}$ [ $2,0 \times 10^{-14}$ – $9,5 \times 10^{-4}$ ]	-	0,13
Surpoids ou obésité	RaDiO	6211	92%	58%	11 [2,1–194]	-	0,0054
	UK Biobank	187 242	85%	67%	2,6 [1,4–5,5]	-	0,0047
IMC (kg/m <sup>2</sup> )	RaDiO	6971	27 [25–30]	26 [23–29]	-	0,35 $\pm$ 0,17	0,044
	UK Biobank	187 242	28 [26–31]	27 [24–30]	-	0,081 $\pm$ 0,027	0,0028
HDL (mmol/L)	RaDiO	6008	1,3 $\pm$ 0,23	1,5 $\pm$ 0,45	-	-0,029 $\pm$ 0,16	0,86
	UK Biobank	164 358	1,5 $\pm$ 0,37	1,5 $\pm$ 0,38	-	0,10 $\pm$ 0,058	0,073
TG (mmol/L)	RaDiO	6051	0,90 [0,81–1,1]	1,1 [0,76–1,6]	-	-0,13 $\pm$ 0,17	0,44
	UK Biobank	178 797	1,5 [1,1–1,9]	1,5 [1,0–2,1]	-	-0,0052 $\pm$ 0,027	0,85
PAS (mmHg)	RaDiO	5257	129 $\pm$ 20	130 $\pm$ 15	-	3,7 $\pm$ 6,4	0,56
	UK Biobank	180 474	136 $\pm$ 18	140 $\pm$ 20	-	1,9 $\pm$ 3,0	0,51
PAD (mmHg)	RaDiO	4693	77 $\pm$ 11	78 $\pm$ 9,4	-	1,4 $\pm$ 4,4	0,75
	UK Biobank	180 514	81 $\pm$ 8,2	82 $\pm$ 11	-	2,3 $\pm$ 2,0	0,25

L'agrégation de variants rares LoF identifiés dans la UK Biobank est associée à une augmentation significative du risque de surpoids ou d'obésité (OR = 2,6 ;  $P = 0,0047$  ; tableau 15), ainsi qu'à une élévation significative de l'IMC ( $\pi = 0,081 \pm 0.027$  ;  $P = 0,0028$  ; tableau 15). D'autre part, parmi tous les individus porteurs de mutations rares LoF parmi les individus de la UK Biobank, aucun d'entre eux n'est atteint de DT2 (tableau 15).

Tableau 16. Variants LoF du gène *OPRD1* identifiés dans la UK Biobank. **Chr** : chromosome, **MAF** : fréquence d'allèle mineur, **gnomAD** : Genome Agregation Database Browser (version 2.1.1), **MAC** : minor allele count (correspondant au nombre d'individus portant la mutation dans la cohorte), **LoF** : perte de fonction.

Chr	Position (Hg38)	Mutation	MAF dans gnomAD	MAC dans UK Biobank	Déteecté dans RaDiO en tant que LoF
1	28812384	c.1A>T, p.M1?	0	2	Non
1	28812411	c.28G>T, p.E10*	0	1	Non
1	28812424	c.41C>G, p.P14R	0	1	Oui
1	28812514	c.131C>A, p.S44*	0,0000053	13	Non
1	28812555	c.174del, p.V59Cfs*22	0,0000065	12	Non
1	28859009	c.283G>A, p.D95N	0,0000080	1	Oui
1	28859013	c.287C>T, p.A96V	0,000016	9	Oui
1	28859041	c.316del, p.S106Vfs*5	0	1	Non
1	28859066	c.342dup, p.P115Afs*31	0	4	Non
1	28859067	c.341G>A, p.W114*	0,0000040	1	Non
1	28859122	c.396G>C, p.M132I	0,0000040	1	Oui
1	28859301	c.575G>A, p.R192Q	0,000045	10	Oui
1	28862855	c.692dup, p.C232Vfs*80	0,0000040	1	Non
1	28862863	c.699T>G, p.Y233*	0,0000040	1	Non
1	28862989	c.828_832del, p.I277Hfs*33	0	1	Non
1	28863120	c.959del, p.F320Sfs*158	0	1	Non
1	28863130	c.966C>G, p.D322E	0	2	Oui
1	28863227	c.1063G>T, p.E355*	0	2	Oui
1	28863228	c.1067_1076del, p.R356Pfs*119	0,000011	1	Non

## 1.8 Effets de l'activation du récepteur DOP sur la sécrétion d'insuline

Etant donné que la consommation d'opioïdes entraîne une augmentation de la glycémie<sup>346</sup>, et que le gène *OPRD1* est exprimé dans les îlots pancréatiques humains, et plus spécifiquement dans les cellules  $\beta$  pancréatiques, nous avons supposé que ce gène pourrait être directement impliqué dans la régulation de l'insulinosécrétion.

Afin de vérifier notre hypothèse, nous avons travaillé sur la lignée cellulaire  $\beta$  pancréatique EndoC $\beta$ H1. Ces cellules n'expriment pas le gène *OPRD1* (figure 47), en lien avec le fait que les cellules  $\beta$  pancréatiques fœtales n'expriment pas ce gène<sup>347</sup>. Le gène *OPRD1* y a donc été dans un premier temps surexprimé. Pour cela, le plasmide pOPRD1 contenant le gène *OPRD1* WT a été électroporé dans ces cellules. 72 heures après l'électroporation, un test de stimulation-sécrétion par le glucose, en présence ou non de l'agoniste de DOP DII à 100

nM, a été réalisé. Brièvement, les cellules sont stimulées à bas glucose (0,5 mM) en présence ou non de l'agoniste de DOP DII pendant 1 heure, puis à haut glucose (16,7 mM) en présence ou non de l'agoniste DII pendant 1 heure. La quantité d'insuline sécrétée dans chacune des conditions a ensuite été dosée à l'aide d'un test ELISA, puis analysée suivant un modèle de régression linéaire ajusté aux conditions expérimentales. En parallèle, l'efficacité de l'électroporation a été vérifiée par mesure de l'expression du gène *OPRD1* dans ces cellules par qPCR. D'autre part, afin de vérifier qu'un même nombre de cellules ait été correctement platé dans chaque puits, les cellules ont été marquées au Draq5 puis révélées.

Par expériences de qPCR, nous avons pu démontrer que le gène *OPRD1* est significativement plus exprimé dans les cellules qui ont été électroporées avec le plasmide pOPRD1 WT par rapport aux cellules qui ne l'ont pas été ( $Ct = 29,88 \pm 0,47$  dans les cellules contrôles et  $Ct = 17,17 \pm 0,17$  dans les cellules surexprimant *OPRD1* ;  $P = 0,0006$  ; figure 47). La valeur du Ct pour le gène *RPLP0* est similaire dans les 2 groupes de cellules ( $Ct = 20,73 \pm 0,44$  dans les cellules contrôles, et  $Ct = 20,34 \pm 0,37$  dans les cellules surexprimant *OPRD1* ;  $P = 0,5314$  ; figure 47). D'autre part, par le marquage des cellules au Draq5, une quantité satisfaisante de cellules a été mise en évidence dans chacun des puits, confirmant le platage correct des cellules (figure 48).

#### Expression des gènes *OPRD1* et *RPLP0*

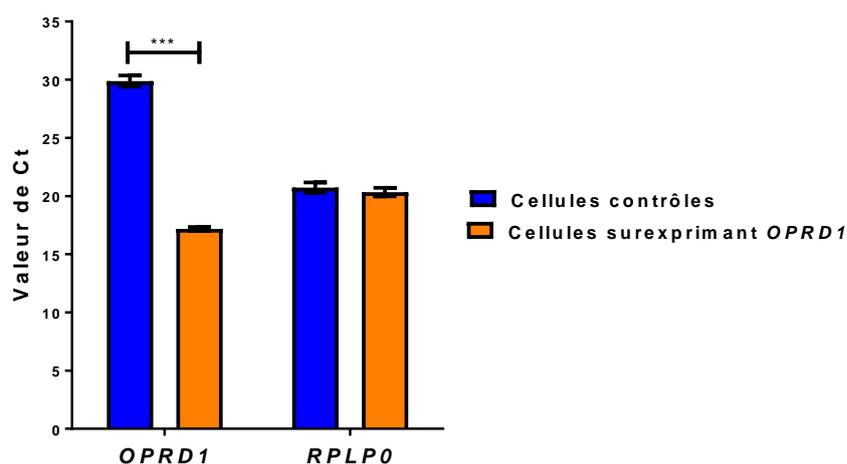


Figure 47. Expression des gènes *OPRD1* et *RPLP0* (utilisé comme gène de ménage) dans les cellules EndoC8H1 électroporées avec le plasmide contenant le gène *OPRD1* WT (en orange), et dans les cellules contrôles qui ne l'ont pas été (en bleu). Pour chaque gène, la différence de Ct entre chaque condition a été analysée par un t test non-pairé. Ces données représentent la moyenne des Ct  $\pm$  SEM pour les gènes *OPRD1* et *RPLP0* de 5 expériences indépendantes. Une diminution du nombre de Ct correspond à une augmentation de l'expression du gène. \*\*\*  $P \leq 0,001$ . Ct : cycle threshold.

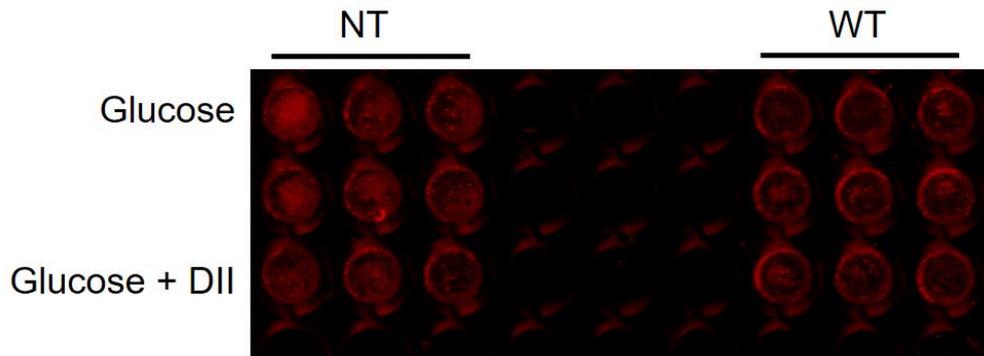


Figure 48. Marquage au Draq5 des cellules EndoCβH1 utilisées au cours des tests de stimulation-sécrétion. Cette image est représentative de 5 expériences indépendantes. **NT** : non-transfecté, **WT** : wild-type, **DII** : [D-Ala<sup>2</sup>]-Deltorphin II.

La réalisation de tests de stimulation-sécrétion au glucose sur ces cellules, couplés ou non à la présence de l'agoniste de DOP DII, a mis en évidence plusieurs éléments intéressants. D'une part, lorsque les cellules EndoCβH1 surexprimant le gène *OPRD1* sont traitées au glucose en présence de l'agoniste DII, on observe une inhibition significative de la sécrétion d'insuline par rapport au traitement de ces cellules uniquement avec du glucose ( $P = 0,018$  ; figure 49). D'autre part, le traitement des cellules EndoCβH1 surexprimant le gène *OPRD1* avec du glucose, en présence de l'agoniste DII, entraîne une inhibition significative de la sécrétion d'insuline par rapport au même traitement réalisé dans les cellules contrôles ( $P = 0,0076$  ; figure 49). DOP jouerait donc un rôle inhibiteur sur la sécrétion d'insuline par les cellules β pancréatiques humaines quand il est activé. Les niveaux d'insuline sécrétée en conditions basales et stimulées sont représentés séparément à l'annexe 10.

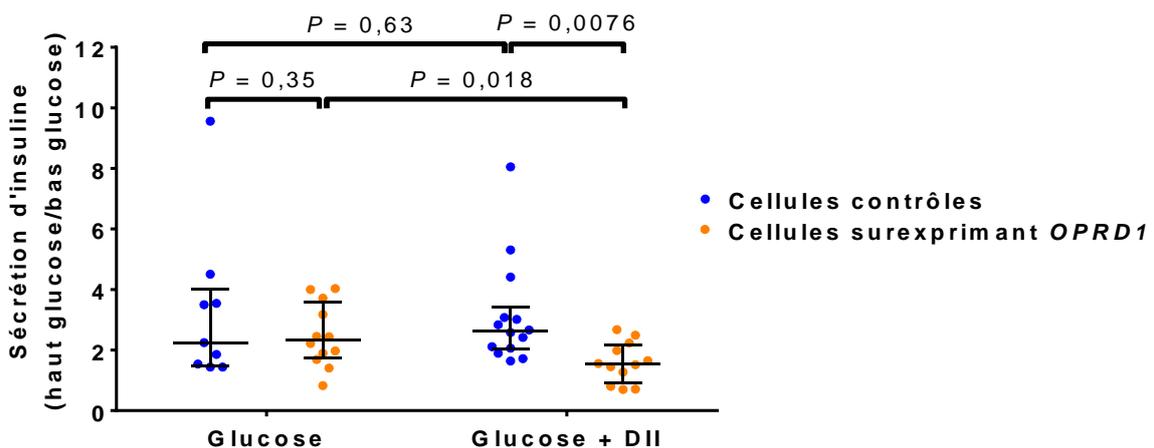


Figure 49. Inhibition de la sécrétion d'insuline par l'activation du récepteur DOP. Ces données représentent le fold change (haut glucose/bas glucose) de la sécrétion d'insuline mesurée dans des cellules EndoCβH1 surexprimant le gène *OPRD1* (représentées en orange) ou non (représentées en bleu), en présence ou non de l'agoniste de DOP DII. Ce graphique représente la médiane ± les écarts interquartiles de 5 expériences indépendantes.

## 2 Projet GLIS3

### 2.1 Identification de variants rares du gène GLIS3 dans l'étude RaDiO

Par séquençage d'exomes du gène *GLIS3* chez 5471 individus issus de l'étude RaDiO, 105 variants hétérozygotes rares de ce gène, ayant une MAF inférieure à 1%, ont été identifiés (tableau 17). 32 de ces variants n'avaient jamais été rapportés auparavant dans la base de données gnomAD (version 2.1.1), et sont donc considérés comme nouveaux. Tous les variants identifiés dans cette étude sont des variants faux-sens, sauf le variant p.Y627\* qui est un variant non-sens, ainsi que le variant p.V916del, qui correspond à une délétion d'un acide aminé.

### 2.2 Prédiction *in silico* de l'effet fonctionnel des variants du gène GLIS3

L'effet délétère de chacun des variants du gène *GLIS3*, identifiés dans l'étude RaDiO, a ensuite été prédit *in silico* selon la méthode REVEL. En établissant un seuil de pathogénicité à 50%, seuls 4 variants (p.M156R, p.S323L, p.E572V et p.T592R) sont prédits comme étant délétères.

Tableau 17. Variants du gène *GLIS3* (NM\_001042413.1) identifiés dans l'étude RaDiO. **Chr** : chromosome, **MAF** : fréquence d'allèle mineur, **gnomAD** : Genome Agregation Database Browser (version 2.1.1), **MAC** : minor allele count (correspondant au nombre d'individus de la cohorte porteurs de la mutation).

Chr	Position (Hg19)	Mutation	MAF dans gnomAD	MAC dans RaDiO	Score de prédiction de pathogénicité <i>in silico</i>
9	4286392	c.34C>T, p.R12W	0,000	1	0,080
9	4286362	c.64A>G, p.M22V	0,0000040	1	0,055
9	4286344	c.82A>G, p.I28V	0,00026	2	0,018
9	4286340	c.86C>T, p.P29L	0	1	0,050
9	4286320	c.106G>A, p.G36R	0,00012	1	0,079
9	4286313	c.113C>G, p.P38R	0,0000040	1	0,22
9	4286224	c.202A>G, p.M68V	0,0000040	1	0,027
9	4286194	c.232C>G, p.R78G	0,00055	3	0,026
9	4286193	c.233G>A, p.R78H	0,000028	1	0,013
9	4286174	c.252A>T, p.L84F	0,00019	10	0,19
9	4286170	c.256C>T, p.P86S	0	1	0,070
9	4286151	c.275C>T, p.T92I	0,000050	1	0,093
9	4286139	c.287C>T, p.P96L	0,00020	3	0,15
9	4286068	c.358G>C, p.G120R	0,000021	3	0,17
9	4286056	c.370C>T, p.P124S	0,0000080	2	0,040
9	4125917	c.413G>A, p.C138Y	0	1	0,18
9	4125908	c.422T>C, p.I141T	0,00023	7	0,071
9	4125905	c.425G>A, p.G142E	0	1	0,32
9	4125863	c.467T>G, p.M156R	0	1	0,57
9	4125858	c.472G>T, p.V158F	0,000067	1	0,27
9	4125834	c.496C>G, p.P166A	0,00045	2	0,37
9	4125827	c.503C>G, p.A168G	0	1	0,24
9	4125815	c.515C>T, p.S172F	0	1	0,36
9	4118765	c.713C>G, p.S238C	0,0000040	1	0,17

9	4118720	c.758G>C, p.S253T	0	1	0,052
9	4118685	c.793G>T, p.V265F	0,00027	1	0,070
9	4118664	c.814T>A, p.Y272N	0,0000040	2	0,40
9	4118651	c.827C>T, p.T276M	0,000012	1	0,090
9	4118645	c.833G>A, p.S278N	0	1	0,031
9	4118636	c.842C>T, p.S281F	0	1	0,14
9	4118634	c.844C>G, p.P282A	0,0016	15	0,10
9	4118588	c.890G>A, p.R297H	0,0000040	1	0,097
9	4118585	c.893C>A, p.S298Y	0,0021	52	0,18
9	4118541	c.937G>A, p.G313R	0,000044	2	0,45
9	4118510	c.968C>T, p.S323L	0,000008	1	0,61
9	4118504	c.974C>T, p.T325M	0,000028	4	0,23
9	4118466	c.1012C>T, p.P338S	0	2	0,21
9	4118415	c.1063C>T, p.R355C	0,0000042	1	0,24
9	4118405	c.1073G>A, p.C358Y	0	1	0,18
9	4118393	c.1085C>A, p.P362Q	0,0012	4	0,063
9	4118388	c.1090C>T, p.P364S	0,0015	31	0,068
9	4118367	c.1111G>T, p.G371C	0	2	0,082
9	4118361	c.1117C>G, p.L373V	0,0012	3	0,043
9	4118352	c.1126C>T, p.P376S	0,0025	10	0,023
9	4118333	c.1145C>A, p.P382Q	0,000014	1	0,049
9	4118324	c.1154G>T, p.G385V	0,00020	5	0,052
9	4118293	c.1185G>A, p.M395I	0	1	0,097
9	4118287	c.1191G>C, p.Q397H	0,0021	24	0,058
9	4118279	c.1199A>G, p.H400R	0,00080	9	0,078
9	4118278	c.1200C>G, p.H400Q	0,00056	1	0,017
9	4118268	c.1210C>A, p.Q404K	0,000032	1	0,12
9	4118264	c.1214C>T, p.P405L	0,000058	2	0,036
9	4118262	c.1216G>T, p.G406C	0,0028	13	0,13
9	4118251	c.1227C>A, p.N409K	0,000075	1	0,063
9	4118246	c.1232T>A, p.M411K	0	1	0,29
9	4118207	c.1271C>G, p.S424W	0	1	0,045
9	4118183	c.1295G>A, p.R432H	0	1	0,077
9	4118160	c.1318A>G, p.T440A	0,00099	3	0,011
9	4118153	c.1325A>C, p.D442A	0	1	0,097
9	4118142	c.1336G>A, p.A446T	0	1	0,025
9	4118090	c.1388A>C, p.H463P	0,000016	2	0,19
9	4118061	c.1417C>T, p.L473F	0,000088	3	0,19
9	4118048	c.1430C>T, p.A477V	0,00074	2	0,047
9	4117998	c.1480G>A, p.D494N	0,000028	1	0,20
9	4117995	c.1483G>A, p.G495S	0,0000040	1	0,083
9	4117960	c.1518C>G, p.D506E	0,000028	2	0,47
9	4117933	c.1545G>C, p.E515D	0,0029	59	0,24
9	4117769	c.1709C>T, p.T570M	0,000018	2	0,15
9	3937185	c.1715A>T, p.E572V	0,0000040	1	0,71
9	3937125	c.1775C>G, p.T592R	0,0000040	1	0,54
9	3937057	c.1843G>A, p.A615T	0,0000040	1	0,35
9	3932462	c.1881T>A, p.Y627*	0	1	NA
9	3898825	c.1994G>A, p.S665N	0,000085	1	0,052
9	3898768	c.2051C>T, p.P684L	0,000039	1	0,035
9	3898759	c.2060C>T, p.S687F	0,00026	5	0,23
9	3898752	c.2067A>T, p.R689S	0	1	0,14
9	3898730	c.2089G>A, p.V697M	0,00080	4	0,039
9	3898723	c.2096G>A, p.R699H	0,00024	2	0,024
9	3898712	c.2107C>T, p.P703S	0,000028	3	0,040
9	3879583	c.2141C>G, p.S714C	0,000018	2	0,10
9	3879443	c.2281G>A, p.D761N	0,000008	2	0,07

9	3856177	c.2305C>A, p.P769T	0,000085	2	0,019
9	3856170	c.2312C>A, p.A771D	0	1	0,10
9	3856078	c.2404G>A, p.G802S	0,000004	1	0,18
9	3856051	c.2431G>C, p.E811Q	0,000004	1	0,033
9	3856043	c.2439C>G, p.N813K	0	1	0,013
9	3856021	c.2461A>G, p.I821V	0	1	0,11
9	3829472	c.2494A>G, p.K832E	0,00016	1	0,071
9	3829448	c.2518G>A, p.D840N	0,000078	1	0,20
9	3829435	c.2531T>C, p.I844T	0,000043	2	0,024
9	3829417	c.2549C>G, p.S850C	0	1	0,26
9	3829357	c.2609G>A, p.S870N	0,000032	3	0,074
9	3829352	c.2614G>C, p.D872H	0,0000071	1	0,22
9	3829343	c.2623C>T, p.H875Y	0	1	0,36
9	3829330	c.2636C>T, p.S879L	0,00013	2	0,27
9	3829330	c.2636C>G, p.S879W	0	1	0,37
9	3829318	c.2648G>C, p.G883A	0	1	0,28
9	3828360	c.2705G>A, p.R902H	0,000036	1	0,032
9	3828355	c.2710G>A, p.G904R	0,00021	18	0,17
9	3828324	c.2741G>A, p.S914N	0	1	0,074
9	3828319	c.2746G>A, p.V916M	0,00013	3	0,026
9	3828316	c.2747_2749del, p.V916del	0	1	NA
9	3828316	c.2749G>A, p.D917N	0	1	0,21
9	3828295	c.2770T>G, p.S924A	0	1	0,23
9	3828280	c.2785G>A, p.E929K	0,000011	1	0,24

### 2.3 Etude d'association entre l'agrégation de variants rares du gène *GLIS3* et le risque de DT2

Afin de déterminer s'il existe une association entre les variants rares du gène *GLIS3* à l'état hétérozygote et le risque de développer un DT2, une étude d'association a été réalisée entre l'agrégation de variants rares du gène et le risque de DT2. Cette étude a été ajustée à l'âge, au sexe, à l'IMC et à l'ethnie. Il est à noter que tous les individus ayant une glycémie à jeun comprise entre 5,6 mmol/L et 7 mmol/L, considérés comme prédiabétiques, ont été exclus de l'étude cas-contrôles.

Comme décrit précédemment, cette analyse a été réalisée suivant la méthode MiST, qui permet de calculer un score  $\pi$ , correspondant à l'effet moyen du cluster de variants, et un score  $\tau$ , correspondant à l'hétérogénéité du cluster. Contrairement au projet *OPRD1*, pour lequel il n'y avait pas d'hétérogénéité du cluster et pour lequel seule la valeur *P* du score  $\pi$  a été prise en compte, une hétérogénéité est présente pour l'agrégation de variants rares du gène *GLIS3*. Ainsi, la valeur *P* considérée pour les études d'association au cours de ce projet sera la valeur *P* globale, reprenant les valeurs *P* du score  $\pi$  et du score  $\tau$ .

Cette analyse a démontré une tendance d'association entre le cluster composé des 105 variants rares du gène *GLIS3* et une diminution du risque de DT2, bien que celle-ci ne soit pas significative (OR = 0,85 ;  $P = 0,055$  ; tableau 18).

Tableau 18. Etude d'association entre l'agrégation des 105 variants rares du gène *GLIS3* identifiés dans l'étude RaDiO et le DT2, réalisée selon la méthode MiST. **N** : nombre d'individus inclus dans l'analyse, **OR** : odds ratio, **CI** : intervalle de confiance,  $\pi$  : effet moyen des variants,  $P_{\pi}$  : valeur P du score  $\pi$ ,  $\tau$  : hétérogénéité du cluster,  $P_{\tau}$  : valeur P du score  $\tau$ .

Variants étudiés	N	OR [CI]	$\pi$	$P_{\pi}$	$\tau$	$P_{\tau}$	$P$ globale
Variants rares (105)	5356	0,85 [-0,41–0,08]	1,66	0,198	49,39	0,050	0,055

## 2.4 Identification des variants LoF par expériences de luciférase

Afin de déterminer l'effet fonctionnel de chacun des variants du gène *GLIS3*, l'activité de chacun d'entre eux a été évaluée par des expériences de luciférase suite à la surexpression de ceux-ci dans le modèle cellulaire HEK293.

### 2.4.1 Vérification de l'efficacité de la transfection dans les cellules HEK293

Dans un premier temps, l'efficacité de la technique de transfection utilisée a été vérifiée. Pour cela, des cellules HEK293 ont été transfectées avec le plasmide pGLIS3 contenant le gène *GLIS3* WT. En parallèle, les conditions contrôles « NT », correspondant à l'absence de transfection du plasmide, ainsi que la condition « p.627\* », correspondant à la transfection d'un plasmide contenant le variant p.Y627\* identifié dans l'étude RaDiO et générant une protéine tronquée dont on suggère fortement la perte totale de fonction, ont également été réalisées. 48 heures après la transfection, ces cellules ont été récupérées. Pour une partie d'entre elles, les protéines en ont été extraites, et des western blots ont été réalisés afin d'évaluer l'expression protéique du facteur de transcription *GLIS3* dans chaque condition. Concernant la seconde partie de ces cellules, l'ARN en a été extrait, rétrotranscrit, puis des expériences de qPCR ont été menées afin de mesurer l'expression du gène *GLIS3* dans chaque condition. Les différences entre les valeurs de Ct obtenues pour chacune des conditions ont été analysées selon un test ANOVA suivi d'un test post-hoc de Tukey.

#### 2.4.1.1 Expression du gène *GLIS3*

Les expériences de qPCR ont permis de mettre en évidence la surexpression significative du gène *GLIS3* dans les cellules transfectées avec le plasmide pGLIS3 WT par rapport aux cellules non-transfectées (Ct = 15,15 ± 0,1463 en condition WT *versus* Ct = 27 ±

0,2415 en condition NT ;  $P < 0,0001$  ; figure 50A). D'autre part, la transfection du plasmide p.Y627\* permet également une surexpression significative du gène *GLIS3* par rapport à la condition non-transfectée (Ct =  $17,17 \pm 0,2269$  en condition p.Y627\* versus Ct =  $27 \pm 0,2415$  en condition NT ;  $P < 0,0001$  ; figure 50A). Cependant, la surexpression de *GLIS3* dans les cellules transfectées avec le plasmide contenant le variant p.Y627\* est significativement plus faible que celle des cellules transfectées avec le plasmide contenant le gène *GLIS3* WT (Ct =  $15,15 \pm 0,1463$  en condition WT versus Ct =  $17,17 \pm 0,2269$  en condition p.Y627\* ;  $P = 0,0012$  ; figure 50A). L'expression du gène *RPLP0*, utilisé comme gène de ménage, est similaire entre les 3 conditions étudiées (figure 50A).

#### 2.4.1.2 Expression de la protéine GLIS3

En parallèle, les expériences de western blots réalisées sur ces cellules ont confirmé la surexpression du gène *GLIS3* dans les cellules HEK293 transfectées avec le plasmide contenant le gène *GLIS3* WT par rapport aux cellules qui n'ont pas été transfectées, et pour lesquelles aucune bande n'est visible, lié probablement à sa faible expression endogène (figure 50B). Concernant les cellules transfectées avec le plasmide contenant le variant p.Y627\*, une bande correspondant au facteur de transcription est également présente, mais à un poids moléculaire plus faible (figure 50B). Ceci est en corrélation avec le remplacement de la tyrosine en position 627 de la protéine par un codon STOP, générant de ce fait une protéine tronquée pour laquelle une partie de la région C-terminale est manquante. Le niveau d'expression de la  $\beta$ -actine, utilisée comme contrôle de charge, est similaire pour les 3 conditions.

#### 2.4.2 Identification des variants LoF du gène *GLIS3* par expériences de luciférase

Afin d'identifier quels variants, parmi les 105 identifiés dans l'étude RaDiO, ont un effet sur la fonctionnalité du récepteur, des expériences de luciférase basées sur la capacité du facteur de transcription *GLIS3* à se lier sur le promoteur du gène de l'insuline afin d'en activer la transcription ont été menées. Pour cela, des cellules HEK293 ont été transfectées avec le plasmide p*GLIS3* contenant le gène *GLIS3* WT ou un des variants à étudier, ainsi qu'avec les plasmides pSRE-LUC dans lequel le gène de la luciférase se trouve sous le contrôle de la région promotrice proximale du gène de l'insuline, et p $\beta$ gal contenant le gène de la  $\beta$ -galactosidase permettant de normaliser les résultats de luciférase selon l'efficacité de la transfection. Un 4<sup>ème</sup> plasmide appelé pMAFA, contenant le gène *MAFA* permettant d'augmenter l'activité

luciférase suite à son recrutement par la protéine GLIS3, a été ajouté ou non à chacune des conditions. Les conditions contrôles « NT » et « p.Y627\* » ont également été réalisées (figure 51). 48 heures après la transfection, les cellules transfectées ont été lysées, puis l'activité de la luciférase a été mesurée et normalisée par celle de la  $\beta$ -galactosidase correspondante.

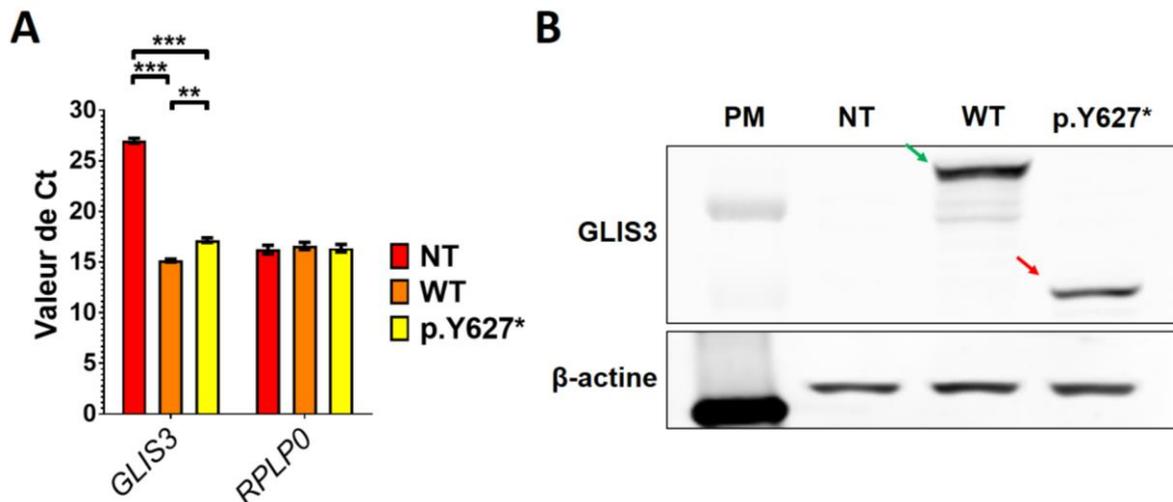


Figure 50. Expression du gène et de la protéine GLIS3 dans des cellules HEK293 transfectées (ou non) avec le plasmide contenant le gène GLIS3 WT ou le variant p.Y627\*. A) Mesure de l'expression des gènes GLIS3 et RPLP0 (utilisé comme gène de ménage) en conditions NT, WT ou p.Y627\* par qPCR. Pour chaque gène, la différence de Ct entre chaque condition a été analysée selon un modèle ANOVA, suivi d'un test Post-Hoc de Tukey. \*\*  $P \leq 0,01$  ; \*\*\*  $P \leq 0,001$ . Ct : cycle threshold, NT : non-transfecté, WT : wild-type. B) Mesure de l'expression des protéines GLIS3 et  $\beta$ -actine (utilisée comme contrôle de charge). La bande correspondant à la protéine GLIS3 WT est indiquée en vert, celle correspondant à la version tronquée de la protéine est indiquée en rouge. PM : marqueur de poids moléculaire, NT : non-transfecté, WT : wild-type.

Les expériences de luciférase ont permis d'identifier 49 variants LoF, c'est-à-dire associés à une diminution significative de l'activité luciférase en absence et/ou en présence de MAFA dans le système par rapport à la condition WT (figure 52). Les 56 variants restants n'ont pas été associés à une diminution de l'activité luciférase, et ont donc été considérés comme neutres (annexe 11).

En comparaison avec les résultats issus des analyses fonctionnelles *in vitro*, l'outil REVEL permettant la prédiction *in silico* de l'effet délétère des variants est peu sensible : seuls 4% des variants LoF identifiés par les expériences de luciférase ont été prédits comme étant délétères par cet outil. Cependant, cet outil est très spécifique, étant donné que tous les variants qui avaient été prédits délétères par REVEL ont été confirmés comme tels par les expériences fonctionnelles.

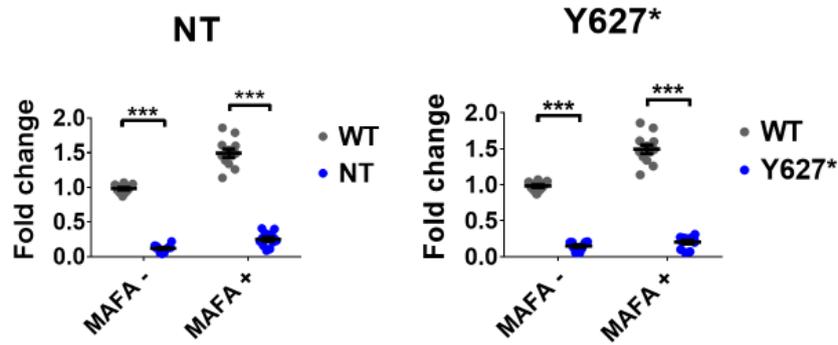


Figure 51. Activité luciférase en conditions « NT » et « Y627\* ». Ces graphiques représentent, pour chaque condition, le fold change de l'activité de la SRE-luciférase (normalisée par les valeurs de  $\beta$ -galactosidase), en présence ou en absence de MAFA, par rapport à l'activité du facteur de transcription GLIS3 WT. Pour chaque condition, l'effet a été analysé selon un modèle ANOVA. Les graphiques représentent la moyenne  $\pm$  SEM des fold change d'au moins 4 expériences indépendantes. \* $P \leq 0,05$ , \*\* $P \leq 0,01$ , \*\*\* $P \leq 0,001$ . NT : non-transfecté, WT : wild-type.

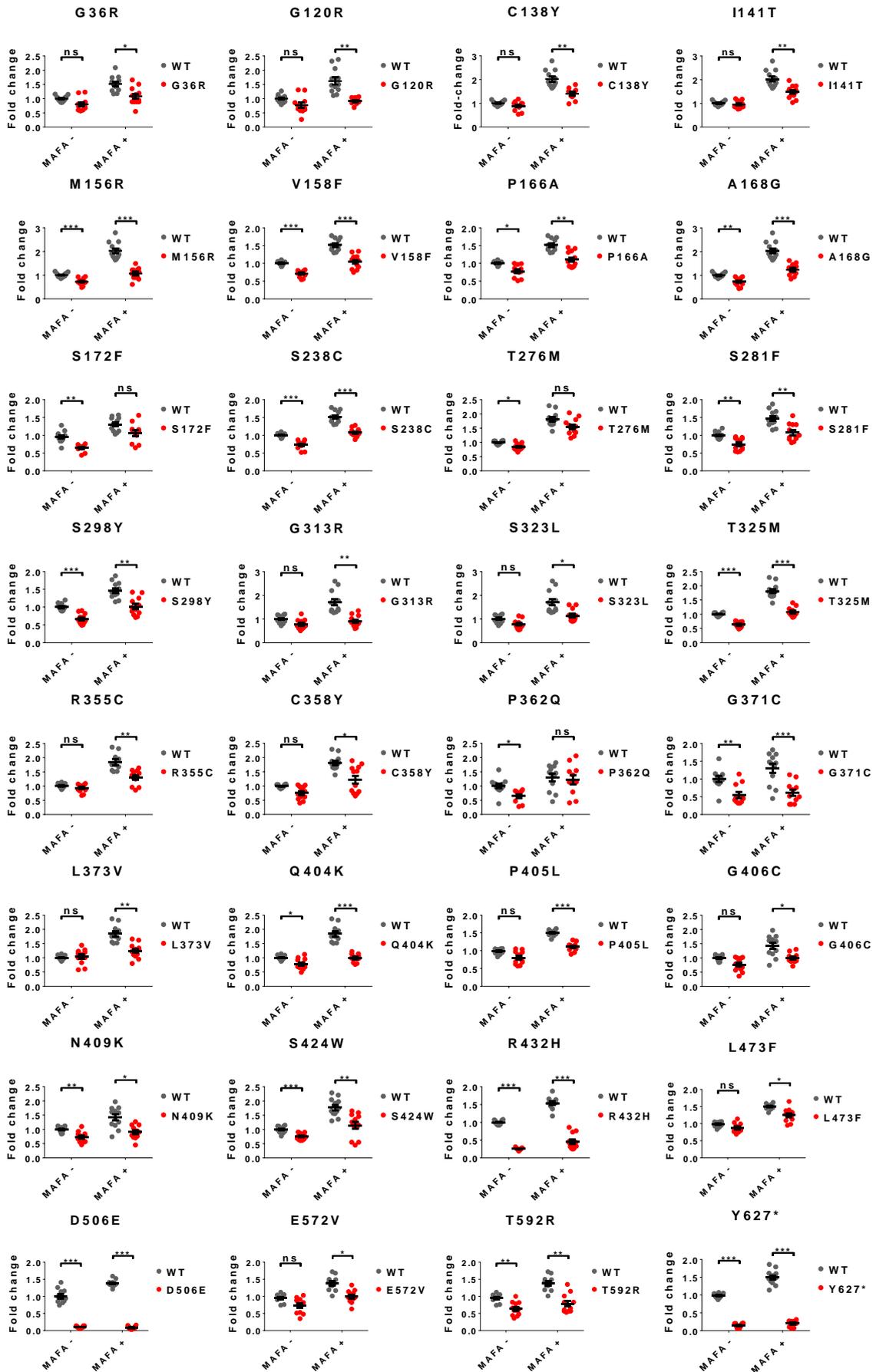
## 2.5 Etude d'association entre l'agrégation des variants rares LoF du gène GLIS3 et le risque de DT2

Suite aux expériences de luciférase permettant d'identifier les variants LoF parmi les 105 variants identifiés dans l'étude RaDiO, une étude d'association a ensuite été réalisée entre l'agrégation des 49 variants LoF et le DT2. Comme pour l'étude du burden de variants rares, cette étude a été ajustée à l'âge, au sexe, à l'IMC et à l'ethnie, et les individus ayant une glycémie à jeun comprise entre 5,6 mmol/L et 7 mmol/L ont été exclus de l'étude cas-contrôles.

Cependant, cette étude n'a établi aucune association significative entre l'agrégation de variants rares LoF et le risque de DT2 (OR = 1,06 ;  $P = 0,495$  ; tableau 19). L'identification des variants LoF par expériences de génétique fonctionnelle, suivie de leur agrégation pour réaliser des études d'association avec le DT2, n'est donc pas suffisant pour pouvoir observer une association statistiquement significative avec le risque de développer un DT2.

## 2.6 Identification des variants P/LP du gène GLIS3 selon les critères ACMG

Pour augmenter la stringence de la sélection des variants délétères, les critères ACMG ont été appliqués à chacun des 105 variants du gène GLIS3 identifiés dans l'étude RaDiO, sachant que GLIS3 est connu pour être impliqué dans le diabète monogénique. La combinaison de ces différents critères permet de déterminer si un variant est pathogène ou non, au-delà des analyses fonctionnelles *in vitro* qui ne constituent qu'un seul critère [PS3] parmi tous les critères ACMG. Ainsi, les résultats obtenus par expériences de luciférase ont été combinés à cette analyse en répondant au critère PS3.



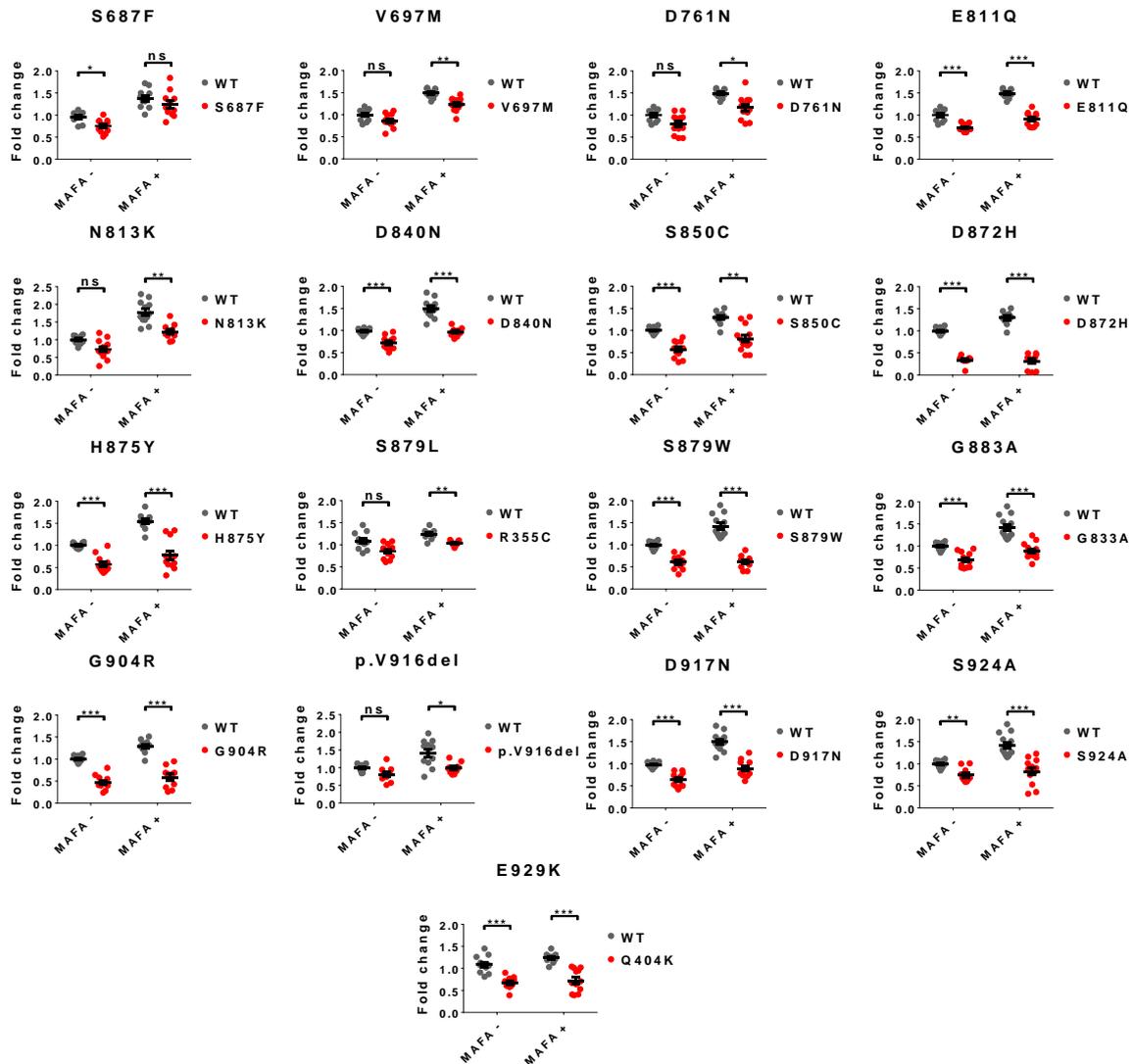


Figure 52. Activité luciférase des variants LoF du gène *GLIS3*. Ces graphiques représentent, pour chaque variant, le fold change de l'activité de la SRE-luciférase (normalisée par les valeurs de  $\beta$ -galactosidase), en présence ou en absence de MAFA, par rapport à l'activité du facteur de transcription *GLIS3* WT. L'effet de chaque variant a été analysé selon un modèle ANOVA. Les graphiques représentent la moyenne  $\pm$  SEM des fold change d'au moins 4 expériences indépendantes. \* $P \leq 0,05$ , \*\* $P \leq 0,01$ , \*\*\*  $P \leq 0,001$ . WT : wild-type.

Tableau 19. Etude d'association entre l'agrégation des 49 variants rares LoF du gène *GLIS3* (identifiés par les expériences de luciférase) et le risque de DT2, réalisée selon la méthode MiST. N : nombre d'individus inclus dans l'analyse, OR : odds ratio, CI : intervalle de confiance,  $\pi$  : effet moyen des variants,  $P_{\pi}$  : valeur P du score  $\pi$ ,  $\tau$  : hétérogénéité du cluster,  $P_{\tau}$  : valeur P du score  $\tau$ , LoF : perte de fonction.

Variants étudiés	N	OR [CI]	$\pi$	$P_{\pi}$	$\tau$	$P_{\tau}$	P globale
Variants LoF (49)	5393	1,06 [-0,31-0,42]	0,10	0,753	14,06	0,244	0,495

Par application de tous les critères ACMG pour les 105 variants identifiés dans l'étude RaDiO (repris à l'annexe 12), seuls 18 d'entre eux ont été définis comme étant P/LP : il s'agit des variants p.C138Y, p.M156R, p.A168G, p.S172F, p.S281F, p.C358Y, p.G371C, p.S424W, p.R432H, p.Y627\*, p.N813K, p.S850C, p.H875Y, p.S879W, p.G883A, p.V916Del, p.D917N, et p.S924A. Les caractéristiques cliniques des individus porteurs de mutations P/LP sont reprises à l'annexe 13.

## 2.7 Enrichissement en variants P/LP dans le domaine C-terminal de la protéine GLIS3

Afin de déterminer si certaines régions du facteur de transcription GLIS3 sont plus enrichies en variants P/LP, celui-ci a été divisé en 5 sections différentes contenant chacune 2 exons. Pour chaque section, le rapport entre le nombre de variants P/LP qui s'y trouvent et le nombre d'acides aminés contenus a été calculé. Les rapports obtenus pour chaque section ont ensuite été comparés entre eux. Ainsi, un enrichissement significatif en variants P/LP a été identifié dans la dernière section de la protéine, contenant les exons 9 et 10 (effet  $[\beta] = 1,65$  ;  $P = 8,37 \times 10^{-3}$  ; figure 15 et tableau 20). Par homologie avec la protéine Glis3 murine, cette région pourrait contenir le domaine de transactivation nécessaire à la fonction de la protéine<sup>348</sup>, pouvant expliquer la plus grande proportion de variants ayant un effet pathogène dans cette région.

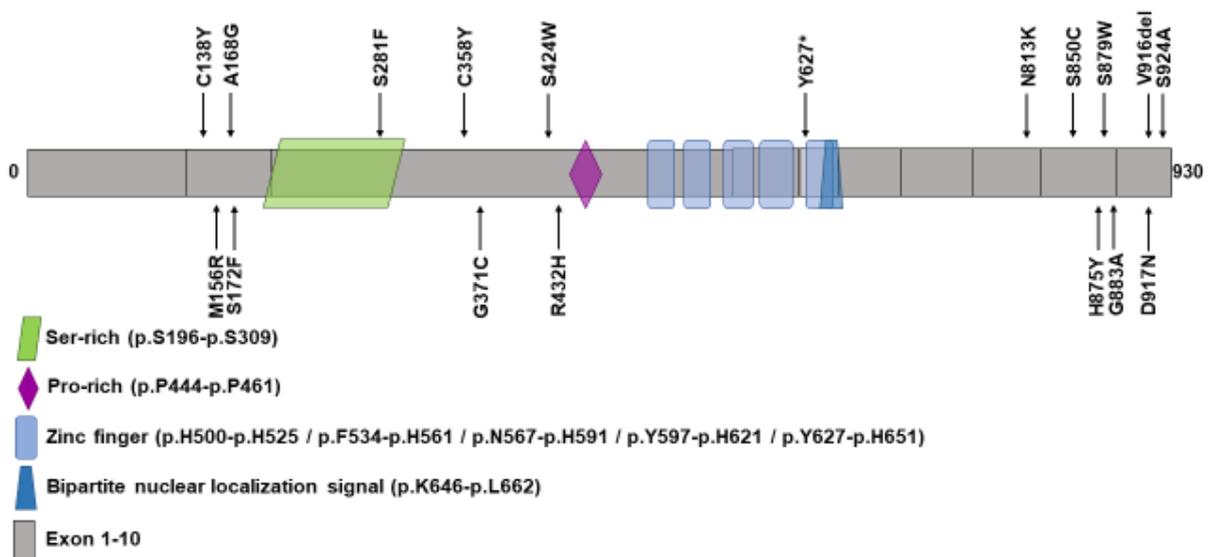


Figure 53. Position des 18 variants P/LP dans la structure du facteur de transcription GLIS3, composé au total de 930 acides aminés. La position des différents mutants P/LP est indiquée par les flèches noires. **Ser-rich** : domaine riche en sérine, **Pro-rich** : domaine riche en proline, **Zinc finger** : motifs en doigts de zinc, **Bipartite nuclear localization signal** : signal bipartite de localisation nucléaire.

Tableau 20. Enrichissement en variants P/LP dans chacune des sections de la protéine GLIS3, contenant chacune 2 exons. La ligne « rapport » correspond au rapport entre le nombre de variants P/LP et le nombre d'acides aminés présents dans la section. P/LP : variants pathogènes/probablement pathogènes.

	Exon 1-2	Exon 3-4	Exon 5-6	Exon 7-8	Exon 9-10
Nombre d'acides aminés	199	425	85	115	106
Nombre de variants P/LP	4	5	1	1	7
Rapport	2,0%	1,2%	1,2%	0,9%	6,6%

## 2.8 Etude d'association entre l'agrégation de variants rares P/LP du gène GLIS3 et le risque de DT2

Suite à l'identification des 18 variants P/LP du gène GLIS3, une nouvelle étude d'association a été réalisée entre l'agrégation de ces variants et le risque de DT2. Comme précédemment, l'analyse a été ajustée à l'âge, au sexe, à l'IMC et à l'ethnie, et les individus ayant une glycémie à jeun comprise entre 5,6 mmol/L et 7 mmol/L ont été exclus.

Cette fois-ci, une association statistiquement significative a été établie entre l'agrégation de variants P/LP et le risque de DT2, avec un risque quasiment 4 fois plus élevé de développer la maladie chez les porteurs de ce type de variants (OR = 3,91 ;  $P = 0,003$  ; tableau 21). D'autre part, tous les individus porteurs de variants P/LP du gène GLIS3 et atteints de DT2 sont traités avec des sulfonyles, suggérant que les médecins ont pu ressentir une altération significative de la fonction des cellules  $\beta$  pancréatiques. En outre, l'âge d'apparition de la maladie est similaire chez les individus porteurs et non-porteurs de ces mutations (annexe 13). De plus, l'IMC moyen des individus porteurs de ces mutations est similaire à celui des individus non-porteurs (annexe 13).

Tableau 21. Etude d'association entre l'agrégation des 18 variants rares P/LP du gène GLIS3 (définis par application des critères ACMG) et le DT2, réalisée selon la méthode MiST. N : nombre d'individus inclus dans l'analyse, OR : odds ratio, CI : intervalle de confiance,  $\pi$  : effet moyen des variants,  $P_n$  : valeur P du score  $\pi$ ,  $\tau$  : hétérogénéité du cluster,  $P_\tau$  : valeur P du score  $\tau$ , P/LP : variants pathogènes/probablement pathogènes.

Variants étudiés	n	OR [CI]	$\pi$	$P_n$	$\tau$	$P_\tau$	P globale
Variants P/LP (18)	5471	3,91 [0,36-2,48]	7,02	0,008	2,80	0,041	0,003

## **DISCUSSION & PERSPECTIVES**

# Discussion et perspectives

## 1 Projet OPRD1

L'identification de variants du gène *OPRD1* par séquençage NGS, suivie de l'investigation de chacun de ces variants par génétique fonctionnelle, a permis de mettre en évidence une association entre ce gène et des effets opposés sur le diabète et l'adiposité : alors que les mutations GoF sont associées à une augmentation du risque de DT2, mais à une diminution du risque d'obésité et une amélioration du profil lipidique, les mutations LoF sont quant à elles associées à une diminution du risque de DT2 mais à une augmentation du risque d'obésité et une détérioration des niveaux de lipides sériques.

L'effet opposé du gène *OPRD1* sur le DT2 et l'adiposité est intrigant, étant donné qu'il est généralement considéré que le DT2 et l'adiposité sont deux troubles métaboliques associés, notamment à cause de l'insulinorésistance provoquée par l'accumulation de graisse qui favorise le développement du DT2. Cependant, ces effets opposés, quoique surprenants, ne sont pas inconnus. Ainsi, dans un article de Mahajan et al. publié en 2018, une méta-analyse GWAS réalisée chez 898 130 individus d'origine européenne a mis en évidence que sur 403 signaux associés au risque DT2, 41 d'entre eux présentent un effet de taille différente selon que l'analyse soit ajustée à l'IMC ou non. Pour 26 signaux, l'ajustement à l'IMC induit une atténuation de l'effet de ceux-ci, mettant en évidence une corrélation positive entre l'effet de ces variants et l'IMC et le DT2, suggérant que l'effet de ces variants sur le DT2 pourrait être principalement induit par l'adiposité. Cependant, les 15 signaux restants (correspondant à moins de 4% des signaux identifiés dans cette étude) présentent quant à eux des effets plus forts suite à un ajustement à l'IMC, mais variés : alors que certains de ces variants auraient un effet important sur la sécrétion d'insuline, d'autres pourraient influencer le risque de DT2 par une diminution de la capacité à stocker les graisses dans le tissu adipeux périphérique<sup>349</sup>. Dans notre cas, nous supposons que cet effet opposé du gène *OPRD1* sur le DT2 et l'adiposité pourrait être lié à des effets indépendants du récepteur DOP dans la cellule  $\beta$  pancréatique et dans le système nerveux central (SNC).

L'association entre les opioïdes et le métabolisme glucidique avait déjà été démontrée par plusieurs études assez anciennes, datant généralement des années 1980 et 1990, dans lesquelles la consommation d'opioïdes est associée à une augmentation de la glycémie et une

diminution de la sécrétion d'insuline<sup>298</sup>. La plupart des études ayant cherché à comprendre ce phénomène ont été réalisées chez des modèles murins, et se sont surtout intéressées au rôle du récepteur MOP dans la sécrétion d'insuline. Cependant, nous avons mis en évidence, par analyse de données de RNA-seq, que le gène *OPRD1* est le seul gène codant un récepteur aux opioïdes exprimé dans les îlots pancréatiques humains, et plus précisément dans les cellules  $\beta$  pancréatiques, au contraire d'*OPRM1* (codant le récepteur MOP) et *OPRK1* (codant le récepteur KOP) qui ne sont pas exprimés dans les îlots. D'autre part, ces gènes ne sont pas exprimés dans les îlots pancréatiques de souris, remettant en question les études réalisées dans les modèles murins. Nous avons donc supposé que chez l'humain, l'effet de la consommation d'opioïdes sur le métabolisme glucidique serait lié à une activation du récepteur DOP dans les cellules  $\beta$  pancréatiques. Grâce à la réalisation de tests de sécrétion d'insuline dans un modèle cellulaire  $\beta$  pancréatique surexprimant le gène *OPRD1*, nous avons démontré que l'activation du récepteur DOP est associée à une inhibition de la sécrétion d'insuline par les cellules  $\beta$ . Cet effet inhibiteur peut être mis en parallèle de l'augmentation du risque de DT2 chez les individus porteurs d'un variant GoF du gène *OPRD1* : chez ces individus, la plus forte activité du récepteur DOP en réponse aux opioïdes pourrait entraîner une plus forte inhibition de la sécrétion d'insuline. La réduction de la sécrétion d'insuline serait ainsi associée à une hyperglycémie et mènerait, à long terme, au phénotype diabétique observé chez ces patients. En outre, la diminution du risque d'hyperglycémie et de DT2 observée chez les patients porteurs de mutations LoF pourrait être liée à une inhibition moins forte de la sécrétion d'insuline en réponse aux opioïdes, menant de ce fait à une sécrétion plus importante qui permettrait une meilleure régulation de la glycémie.

Nous pensons que l'activation du récepteur DOP dans les cellules  $\beta$  pancréatiques pourrait avoir lieu grâce à l'action de peptides opioïdes endogènes. Pour rappel, le récepteur DOP peut être activé par plusieurs peptides endogènes : d'une part, les met- et leu-enképhalines, obtenues suite au clivage de la proenképhaline (codée par le gène *PENK*), et d'autre part la  $\beta$ -endorphine, qui est obtenue suite au clivage de la proopiomélanocortine (codée par le gène *POMC*). Par analyse de données de RNA-seq, nous avons observé que le gène *POMC*, ainsi que les gènes codant les enzymes PC1/3 et PC2 (codées par les gènes *PCSK1* et *PCSK2* respectivement) nécessaires au clivage de POMC en  $\beta$ -endorphine<sup>350</sup> sont exprimés dans les îlots pancréatiques humains. Nous envisageons donc que la régulation de la sécrétion

d'insuline via l'activation du récepteur DOP pourrait avoir lieu au sein des îlots de manière paracrine, via la sécrétion de  $\beta$ -endorphine par les cellules endocrines. Ceci est soutenu par d'anciennes études ayant mis en évidence la présence de  $\beta$ -endorphine dans les îlots humains, notamment dans les cellules  $\delta$  pancréatiques et dans une moindre mesure dans les cellules  $\alpha$  pancréatiques<sup>351</sup>. Ces deux types cellulaires sont capables de réguler la sécrétion d'insuline par l'action du glucagon et de la somatostatine sur les cellules  $\beta$  pancréatiques, respectivement. La sécrétion de  $\beta$ -endorphine par ces 2 types de cellules endocrines pourrait donc être impliquée dans la régulation paracrine de la sécrétion d'insuline au sein des îlots. Cependant, étant donné l'ancienneté de ces études et les outils utilisés, cette hypothèse resterait à confirmer. En outre, la présence de la  $\beta$ -endorphine dans les autres types cellulaires composant les îlots n'a pas été évaluée au cours de l'étude mentionnée précédemment : il n'est donc pas exclu que d'autres cellules du pancréas endocrine puissent également exprimer ce peptide, et ainsi participer à la régulation de l'insulinosécrétion. D'autre part, il est également possible que le récepteur DOP, présent à la surface des cellules  $\beta$  pancréatiques, ne soit pas stimulé uniquement par la  $\beta$ -endorphine sécrétée par les cellules des îlots, mais également par des opioïdes endogènes sécrétés par d'autres tissus de l'organisme, notamment par la glande surrénale ou la glande pituitaire, qui seraient ensuite transportés vers les îlots par la circulation sanguine<sup>352</sup>. Par exemple, plusieurs études menées chez des modèles mammifères ont mis en évidence l'expression des peptides opioïdes endogènes ( $\beta$ -endorphine, met-enképhaline et leu-enképhaline) dans la glande médullosurrénale, qui est impliquée dans la réponse au stress par sécrétion de l'adrénaline et la noradrénaline<sup>353</sup>. Ces deux hormones, à forte concentration (en cas de stress par exemple), sont impliquées dans l'inhibition de la sécrétion d'insuline par action sur les récepteurs  $\alpha_2$ -adrénergiques situés à la surface de la cellule  $\beta$ <sup>32</sup>. Ainsi, la sécrétion de peptides opioïdes par la médullosurrénale pourrait cibler le récepteur DOP présent à la surface de cellules  $\beta$ , et ainsi contribuer à l'inhibition de la sécrétion de l'insuline, notamment en cas de stress.

L'effet inhibiteur du récepteur DOP sur la sécrétion d'insuline pourrait être lié à une diminution de la concentration en AMPc dans la cellule  $\beta$  pancréatique. Pour rappel, Le récepteur DOP est couplé à une protéine  $G_i$ , capable d'inhiber l'adénylate cyclase dont le rôle est de générer de l'AMPc à partir de l'ATP. Comme expliqué dans l'introduction, l'AMPc est un second messager qui permet de stimuler la sécrétion de l'insuline en agissant à différents

niveaux dans la cellule  $\beta$ . Ainsi, l'inhibition de l'adénylate cyclase par le récepteur DOP pourrait engendrer une réduction de la concentration en AMPc dans la cellule, menant ainsi à une plus faible stimulation de la sécrétion d'insuline. Cette hypothèse est appuyée par une étude réalisée chez des individus sains : alors que l'injection de  $\beta$ -endorphine chez ces individus entraîne une inhibition de la sécrétion d'insuline stimulée par le glucose, l'administration de théophylline chez ces individus, permettant de maintenir la concentration intracellulaire en AMPc par inhibition de la phosphodiesterase impliquée dans sa dégradation, permet de restaurer la sécrétion d'insuline initialement inhibée par la  $\beta$ -endorphine<sup>354</sup>. Cependant, il n'est pas exclu que ce récepteur régule la sécrétion d'insuline via d'autres mécanismes.

Le modèle cellulaire EndoC $\beta$ H1, utilisé au cours de ce projet pour étudier l'effet du récepteur DOP sur la sécrétion d'insuline, a pour avantage d'être un modèle  $\beta$  pancréatique humain, par rapport aux autres lignées cellulaires  $\beta$  généralement d'origine murine. Ceci dit, il présente tout de même quelques limites. D'une part, il s'agit d'une lignée cellulaire obtenue à partir de cellules  $\beta$  pancréatiques humaines d'origine fœtale qui n'expriment pas *OPRD1*, alors que les cellules  $\beta$  pancréatiques matures l'expriment : nous avons donc dû surexprimer ce gène dans notre modèle cellulaire pour pouvoir réaliser notre étude. Cependant, le niveau d'expression « artificiel » de ce gène dans notre modèle est donc probablement plus important que son niveau d'expression dans des cellules  $\beta$  pancréatiques matures. L'effet observé lors de notre étude peut donc être un peu différent de l'effet réel en conditions physiologiques. D'autre part, comme mentionné précédemment, les cellules  $\beta$  pancréatiques n'agissent pas seules sur la régulation de la glycémie, mais sont en interaction avec les autres types cellulaires composant les îlots de Langerhans. L'utilisation de cellules  $\beta$  pancréatiques ne permet donc pas de tenir compte des régulations internes aux îlots de Langerhans. Ainsi, il serait judicieux de réitérer ces expériences sur des îlots pancréatiques humains, afin de confirmer nos observations dans un modèle se rapprochant de la physiologie humaine.

Comme mentionné précédemment, les variants du gène *OPRD1* sont également associés à l'adiposité : alors que les variants GoF sont associés à une diminution de celle-ci, les variants LoF sont associés à son augmentation. Ces données sont soutenues par quelques études observationnelles dans lesquelles une diminution de l'IMC et une amélioration des niveaux de lipides sériques ont été observées chez les consommateurs d'opium. Cependant, ces études observationnelles présentent souvent des effets contradictoires<sup>298</sup>. D'autre part, la majorité

des études réalisées chez des modèles murins ont mis en évidence que la consommation d'opioïdes stimule la prise alimentaire, alors que l'administration d'antagonistes des récepteurs aux opioïdes réduit la prise alimentaire, ce qui est complètement opposé à nos résultats<sup>355</sup>. Cependant, une grande partie de ces études se sont intéressées soit à l'effet global de l'administration d'opioïdes, sans se focaliser sur un type de récepteur en particulier, soit à l'effet du récepteur MOP et dans une moindre mesure du récepteur KOP<sup>356</sup>. Ainsi, les connaissances concernant l'effet du récepteur DOP sur l'adiposité sont très limitées. Une étude menée par Czyzyk et al., en 2012, a mis en évidence qu'un knockout du gène *OPRD1* chez des souris nourries avec un régime riche en calories entraîne une hyperphagie chez celles-ci<sup>357</sup>, allant dans le même sens que nos observations. Cependant, ces souris présentent une plus faible prise de poids corporel que les animaux WT, et ont une dépense énergétique plus importante, mais ce phénomène peut être expliqué par l'activation de la thermogenèse dans le tissu adipeux brun de ces animaux. Nous supposons donc que l'effet spécifique du récepteur DOP sur l'adiposité, a priori opposé à celui des récepteurs MOP et KOP, pourrait être médié par une régulation centrale de la prise alimentaire, liée au maintien de l'homéostasie énergétique via la voie leptine-mélanocortine, et/ou à la sensation de plaisir provoquée par la nourriture, par l'activation du système de récompense. Plusieurs éléments soutiennent notre hypothèse : d'une part, le récepteur DOP est exprimé (bien que faiblement) dans l'hypothalamus, qui est la région cérébrale clé impliquée dans la régulation de la balance énergétique par la voie leptine-mélanocortine, mais est également exprimé dans les régions neuronales impliquées dans le système de récompense<sup>356,358</sup>. D'autre part, la  $\beta$ -endorphine, obtenue par le clivage du précurseur POMC qui joue un rôle majeur dans la voie leptine-mélanocortine, ainsi que les met- et leu-enképhalines obtenues à partir de PENK sont exprimées dans l'hypothalamus<sup>359-361</sup>, et les enképhalines sont également fortement exprimées dans le système de récompense<sup>362</sup>.

En outre, nous ne pouvons pas exclure que l'effet du gène *OPRD1* sur l'adiposité soit associé à des effets périphériques, étant donné que ce gène est également exprimé dans certains tissus métaboliques périphériques, tel que le muscle<sup>288</sup>. D'ailleurs, chez des souris, l'administration de  $\beta$ -endorphine ou de l'agoniste spécifique du récepteur DOP DPDPE entraîne une augmentation de la capture de glucose par le muscle squelettique<sup>363</sup>. Le

récepteur DOP pourrait donc également être impliqué dans la régulation périphérique du métabolisme.

Afin de comprendre quels sont les mécanismes associant le gène *OPRD1* et l'adiposité, il serait intéressant d'investiguer l'effet central de ce gène par la génération d'animaux knockout pour celui-ci, et d'en étudier l'effet sur le métabolisme énergétique. Ainsi, la délétion de ce récepteur spécifiquement dans les neurones du système de récompense permettrait de déterminer si l'effet du récepteur sur l'adiposité est lié au plaisir ; le knockout de ce gène dans les neurones impliqués dans la voie leptine-mélanocortine permettrait quant à lui de savoir si cet effet est lié au maintien de la balance énergétique. De la même manière, on pourrait également envisager l'utilisation d'un modèle knockout pour le gène *PENK* (qui est le précurseur des met- et leu-enképhalines spécifiques au récepteur DOP) spécifiquement dans le système nerveux, qui permettrait également d'investiguer l'effet central du récepteur DOP. Cependant, ces modèles sont à prendre avec un certain recul, étant donné que les animaux peuvent avoir une distribution différente du récepteur DOP dans leurs tissus par rapport à l'humain, comme cela a été démontré pour l'expression du récepteur dans les îlots pancréatiques.

L'effet du gène *OPRD1* sur le métabolisme, et notamment sur le risque de DT2, nous laisse penser que ce gène pourrait être ciblé dans le cadre de la médecine de précision. Ainsi, dans le cas où l'effet du récepteur DOP sur l'adiposité serait bien dépendant d'un mécanisme central et non périphérique, nous supposons que le développement d'antagonistes du récepteur DOP, incapables de passer la barrière hémato-encéphalique, permettrait d'améliorer le phénotype des patients atteints de DT2 sans augmenter le risque d'obésité. Au contraire, l'utilisation d'un agoniste du récepteur DOP qui puisse agir dans le SNC afin de réduire le risque d'obésité nous semble compliqué à mettre en place : d'une part, l'utilisation d'agonistes du récepteur pourrait créer une dépendance chez les individus ; d'autre part, les récepteurs aux opioïdes sont exprimés dans de nombreuses régions du cerveau, et l'utilisation d'agonistes pourrait ainsi altérer d'autres fonctions neuronales.

## 2 Projet *GLIS3*

L'identification de variants rares du gène *GLIS3* par séquençage NGS, suivie de l'investigation de ces variants par génétique fonctionnelle et de l'application des critères

ACMG a permis d'identifier 18 variants P/LP de ce gène, qui sont associés au DT2 commun. Alors que ce gène était déjà connu pour son implication dans le développement du diabète néonatal, très peu d'études se sont intéressées à son association potentielle avec la forme polygénique du DT2. Quelques études GWAS ont pu identifier certains SNPs fréquents de ce gène associés à la maladie et à certains traits métaboliques associés<sup>364</sup>, mais présentant un faible effet sur le risque de développer ce trouble. Une seule étude datant de 2019, réalisée par Sun et al., s'est intéressée à l'effet de variants rares du gène *GLIS3* identifiés par séquençage NGS sur le risque de DT2 commun<sup>330</sup>. Cependant, cette étude n'a pas investigué l'effet fonctionnel des différents variants, et l'étude d'association entre l'agrégation de ces variants et le risque de DT2 n'a démontré qu'un faible effet de ces variants sur le risque de développer la maladie. Notre étude, combinant séquençage NGS et génétique fonctionnelle, a ainsi permis de démontrer une association forte entre les variants rares de *GLIS3* et le risque de DT2 commun.

Dans le cadre de notre étude, nous ne nous sommes intéressés qu'à l'effet des différents variants rares du gène *GLIS3* sur l'activation de la transcription du gène *INS*, ainsi qu'à leur capacité à stabiliser la liaison du facteur de transcription MAFA sur la région promotrice de ce gène. Bien que des altérations de ces mécanismes puissent expliquer l'association entre le gène *GLIS3* et le risque de DT2, d'autres éléments pourraient expliquer cette association. Ainsi, en plus d'être activée par les facteurs de transcription *GLIS3* et *MAFA*, la transcription du gène *INS* peut également être stimulée par les facteurs de transcription *NEUROD1* et *PDX1*. Comme pour *MAFA*, l'interaction de ces 2 facteurs de transcription avec la région promotrice du gène *INS* est stabilisée grâce à *GLIS3*<sup>324</sup>. Ainsi, les différents variants de *GLIS3* pourraient également réduire la transcription du gène de l'insuline en altérant le recrutement de *PDX1* et *NEUROD1* sur la région promotrice du gène *INS*.

D'autre part, *GLIS3* n'est pas seulement impliqué dans la transcription du gène de l'insuline, mais également dans la transcription de plusieurs autres gènes importants dans la cellule  $\beta$  pancréatique. Parmi ceux-ci, on peut notamment citer le gène *NGN3*, codant le facteur de transcription neurogénine 3 impliqué dans le développement des cellules pancréatiques endocrines<sup>364</sup>. Les souris knockout pour le gène *Glis3* présentent une forte réduction de l'expression de la neurogénine 3, qui est associée à une forte réduction de la taille des îlots et du contenu en cellules endocrines, couplée à une hyperglycémie<sup>321</sup>. Ce

phénotype ressemble fortement à celui observé chez les individus porteurs de mutations homozygotes rares du gène *GLIS3*, qui souffrent de diabète néonatal. Ainsi, on peut supposer que des mutations hétérozygotes de ce gène pourraient entraîner une légère altération du développement des cellules pancréatiques endocrines, et dont les conséquences se manifesteraient plus tard dans la vie, à cause d'une stimulation trop importante et d'un épuisement des cellules  $\beta$  pancréatiques. D'autre part, *GLIS3* est également impliqué dans la régulation de la prolifération et de l'expansion de la masse de cellules  $\beta$  via la transcription du gène *CCND2*, codant le facteur de transcription cycline D2<sup>364</sup>. Ce mécanisme est particulièrement important en cas de grossesse ou d'insulinorésistance, étant donné qu'il permet d'ajuster la quantité d'insuline sécrétée afin de maintenir une glycémie correcte. Ainsi, chez des souris déficientes hétérozygotes pour le gène *Glis3*, un diabète associé à une hypoinsulinémie ne se développe pas en cas de régime classique, mais uniquement en cas de régime riche en graisse<sup>325</sup>. Les variants hétérozygotes du gène *GLIS3* pourraient donc potentiellement être associés au développement d'un DT2 commun par altération du mécanisme d'expansion de la masse de cellules  $\beta$ . Cependant, cette hypothèse est à considérer avec un certain recul, étant donné que l'expansion de la masse de cellules  $\beta$  pancréatiques chez l'humain est controversée. Enfin, le gène *GLIS3* semblerait également impliqué dans l'apoptose des cellules  $\beta$ , en jouant un rôle protecteur dans les cellules<sup>364</sup>. Dans le modèle cellulaire  $\beta$  pancréatique de rat INS-1, le knockdown du gène *Glis3* est associé à une augmentation de l'apoptose des cellules  $\beta$  induite par des cytokines pro-inflammatoires ou par du palmitate, par activation des voies de l'apoptose<sup>365</sup>. Ce mécanisme pourrait donc également être impliqué dans l'association entre le risque de développer un DT2 commun et les variants du gène *GLIS3*.

Dans le cadre de notre étude, nous avons montré que les variants rares P/LP du gène *GLIS3* sont fortement enrichis dans les deux derniers exons du gène, correspondant à la région C-terminale de la protéine. Chez la souris, dont le gène *Glis3* présente une forte homologie de séquence avec le gène humain, le domaine de transactivation du facteur de transcription, qui contient notamment des sites de liaison pour des corégulateurs transcriptionnels, est situé dans la région C-terminale de la protéine<sup>348</sup>. Sans ce domaine, l'activité transcriptionnelle du facteur de transcription est altérée. Par homologie, on peut supposer que les mutations du gène *GLIS3* humain situées dans ce domaine sont vraisemblablement impliquées dans une

forte réduction, voire une abolition, de l'activité transactivatrice de ce facteur de transcription, expliquant ainsi leur forte pathogénicité. Ceci est soutenu par le fait que la délétion de la région C-terminale du gène *Glis3* de la souris, semblable à la délétion observée chez certains patients atteints d'un diabète néonatal, est associée à une perte de l'activité transcriptionnelle de la protéine<sup>348</sup>.

L'une des perspectives de ce projet serait, d'une part, de répliquer nos études d'association entre les variants P/LP du gène *GLIS3* et le DT2 dans une cohorte plus large afin de confirmer les observations qui ont été réalisées dans ce projet. D'autre part, il pourrait être intéressant d'investiguer l'effet fonctionnel de ces variants sur les autres mécanismes dans lesquels *GLIS3* est impliqué, à savoir le développement des cellules pancréatiques endocrines, l'expansion de la masse  $\beta$  pancréatique en réponse à une demande plus importante en insuline, ou la protection vis-à-vis de l'apoptose des cellules  $\beta$ , ainsi que sur son expression ou sa localisation nucléaire. Ceci permettrait potentiellement d'identifier les effets délétères de certains variants qui n'auraient pas pu être démontrés au cours de nos expériences de luciférase (par absence d'effet sur la transcription du gène de l'insuline), et qui pourraient renforcer l'association entre ce gène et le risque de DT2 commun.

# **CONCLUSION**

## Conclusion

La réalisation de séquençage NGS, couplée à la réalisation d'expériences de génétique fonctionnelle, s'est révélée être une combinaison très intéressante et novatrice dans le cadre de la découverte de variants rares associés au développement du DT2 commun. Jusqu'à il y a peu, ces outils génétiques étaient réservés à l'investigation de l'étiologie génétique des maladies monogéniques afin d'identifier des variants rares impliqués dans leur développement, les formes polygéniques étant plutôt investiguées par le génotypage de variants fréquents ou SNPs par GWAS. Cependant, ces SNPs sont généralement situés dans des régions non-codantes, et augmentent faiblement le risque de développer la maladie. Les expériences réalisées au cours de ce projet de thèse ont ainsi démontré l'intérêt de réaliser du séquençage NGS suivi de la réalisation d'expériences de génétique fonctionnelle afin d'identifier des variants rares, situés dans des régions codantes du génome, et présentant un effet fort sur le risque de développer un DT2 commun. Ces travaux ont ainsi permis l'identification du gène *OPRD1* comme étant associé au risque de DT2 commun et à l'adiposité de manière opposée, et a permis de confirmer l'association entre le gène *GLIS3*, dont l'implication dans le développement du diabète néonatal était déjà connue, avec le risque de DT2 commun. Ces études contribuent ainsi à la meilleure compréhension de l'héritabilité de la maladie, qui est encore mal connue. D'autre part, l'éclaircissement de l'étiologie génétique du DT2 commun pourrait contribuer au développement d'une médecine de précision, non seulement par séquençage du génome des patients afin d'identifier les gènes mutés, mais également par adaptation du traitement des patients en fonction des altérations génétiques sous-jacentes, comme c'est déjà le cas pour le diabète monogénique.

# **BIBLIOGRAPHIE**

## Bibliographie

1. Tattersall, R. B. The History of Diabetes Mellitus. in *Textbook of Diabetes* 1–23 (John Wiley & Sons, Ltd, 2010). doi:10.1002/9781444324808.ch1.
2. Lakhtakia, R. The History of Diabetes Mellitus. *Sultan Qaboos Univ. Med. J.* **13**, 368–370 (2013).
3. National Diabetes Data Group. Classification and Diagnosis of Diabetes Mellitus and Other Categories of Glucose Intolerance. *Diabetes* **28**, 1039–1057 (1979).
4. Organization, W. H. Definition, diagnosis and classification of diabetes mellitus and its complications : report of a WHO consultation. Part 1, Diagnosis and classification of diabetes mellitus. (1999).
5. American Diabetes Association. 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes—2019. *Diabetes Care* **42**, S13–S28 (2019).
6. Alberti, K. G. M. M. The Classification and Diagnosis of Diabetes Mellitus. in *Textbook of Diabetes* 24–30 (John Wiley & Sons, Ltd, 2010). doi:10.1002/9781444324808.ch2.
7. Organisation Mondiale de la Santé. *Rapport mondial sur le diabète*. (Organisation Mondiale de la Santé, 2016).
8. WHO | Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia. WHO [https://www.who.int/diabetes/publications/diagnosis\\_diabetes2006/en/](https://www.who.int/diabetes/publications/diagnosis_diabetes2006/en/).
9. Fayfman, M., Pasquel, F. J. & Umpierrez, G. E. Management of Hyperglycemic Crises: Diabetic ketoacidosis and hyperglycemic hyperosmolar state. *Med. Clin. North Am.* **101**, 587–606 (2017).
10. Seaquist, E. R. Microvascular complications of diabetes. *Postgrad. Med.* **103**, 61–68 (1998).
11. Hirai, F. E., Tielsch, J. M., Klein, B. E. K. & Klein, R. Ten-year change in vision-related quality of life in type 1 diabetes: Wisconsin epidemiologic study of diabetic retinopathy. *Ophthalmology* **118**, 353–358 (2011).
12. Nathan, D. M. Long-term complications of diabetes mellitus. *N. Engl. J. Med.* **328**, 1676–1685 (1993).
13. Forbes, J. M. & Cooper, M. E. Mechanisms of Diabetic Complications. *Physiol. Rev.* **93**, 137–188 (2013).
14. Gilbertson, D. T. *et al.* Projecting the number of patients with end-stage renal disease in the United States to the year 2015. *J. Am. Soc. Nephrol. JASN* **16**, 3736–3741 (2005).
15. Abbott, C. A., Malik, R. A., van Ross, E. R. E., Kulkarni, J. & Boulton, A. J. M. Prevalence and characteristics of painful diabetic neuropathy in a large community-based diabetic population in the U.K. *Diabetes Care* **34**, 2220–2224 (2011).
16. Ross, M. A. Neuropathies associated with diabetes. *Med. Clin. North Am.* **77**, 111–124 (1993).
17. Obrosova, I. G. Diabetic painful and insensate neuropathy: Pathogenesis and potential treatments. *Neurotherapeutics* **6**, 638–647 (2009).

18. Fowler, M. J. Microvascular and Macrovascular Complications of Diabetes. *Clin. Diabetes* **26**, 77–82 (2008).
19. Morrish, N. J., Wang, S. L., Stevens, L. K., Fuller, J. H. & Keen, H. Mortality and causes of death in the WHO Multinational Study of Vascular Disease in Diabetes. *Diabetologia* **44 Suppl 2**, S14-21 (2001).
20. Tripathi, B. K. & Srivastava, A. K. Diabetes mellitus: complications and therapeutics. *Med. Sci. Monit. Int. Med. J. Exp. Clin. Res.* **12**, RA130-147 (2006).
21. Saeedi, P. *et al.* Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Res. Clin. Pract.* **157**, (2019).
22. Fédération Internationale du Diabète. Atlas du diabète de la FID, 9ème édition. Bruxelles, Belgique: Fédération Internationale du Diabète, 2019.
23. Bommer, C. *et al.* The global economic burden of diabetes in adults aged 20-79 years: a cost-of-illness study. *Lancet Diabetes Endocrinol.* **5**, 423–430 (2017).
24. Association, A. D. Economic Costs of Diabetes in the U.S. in 2017. *Diabetes Care* **41**, 917–928 (2018).
25. Fosse-Edorh S., Mandereau-Bruno L., Piffaretti C. Le poids du diabète en France en 2016. Synthèse épidémiologique. Saint-Maurice : Santé publique France ; 2018. 8 p.
26. Prévalence et incidence du diabète. /maladies-et-traumatismes/diabete/prevalence-et-incidence-du-diabete.
27. Alsahli, M. & Gerich, J. E. Abnormalities of Insulin Secretion and  $\beta$ -Cell Defects in Type 2 Diabetes. in *Textbook of Diabetes* 160–173 (John Wiley & Sons, Ltd, 2010). doi:10.1002/9781444324808.ch10.
28. Yki-Järvinen, H. Insulin Resistance in Type 2 Diabetes. in *Textbook of Diabetes* 174–190 (John Wiley & Sons, Ltd, 2010). doi:10.1002/9781444324808.ch11.
29. Leung, P. S. Overview of the Pancreas. in *The Renin-Angiotensin System: Current Research Progress in The Pancreas: The RAS in the Pancreas* (ed. Leung, P. S.) 3–12 (Springer Netherlands, 2010). doi:10.1007/978-90-481-9060-7\_1.
30. Leung, P. S. Physiology of the Pancreas. in *The Renin-Angiotensin System: Current Research Progress in The Pancreas: The RAS in the Pancreas* (ed. Leung, P. S.) 13–27 (Springer Netherlands, 2010). doi:10.1007/978-90-481-9060-7\_2.
31. Githens, S. The pancreatic duct cell: proliferative capabilities, specific characteristics, metaplasia, isolation, and culture. *J. Pediatr. Gastroenterol. Nutr.* **7**, 486–506 (1988).
32. Jones, P. M. & Persaud, S. J. Islet Function and Insulin Secretion. in *Textbook of Diabetes* 85–103 (John Wiley & Sons, Ltd, 2010). doi:10.1002/9781444324808.ch6.
33. Cabrera, O. *et al.* The unique cytoarchitecture of human pancreatic islets has implications for islet cell function. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 2334–2339 (2006).
34. Fu, Z., Gilbert, E. R. & Liu, D. Regulation of Insulin Synthesis and Secretion and Pancreatic Beta-Cell Dysfunction in Diabetes. *Curr. Diabetes Rev.* **9**, 25–53 (2013).
35. Vakilian, M., Tahamtani, Y. & Ghaedi, K. A review on insulin trafficking and exocytosis. *Gene* **706**, 52–61 (2019).

36. Barg, S., Eliasson, L., Renström, E. & Rorsman, P. A Subset of 50 Secretory Granules in Close Contact With I-Type Ca<sup>2+</sup> Channels Accounts for First-Phase Insulin Secretion in Mouse  $\beta$ -Cells. *Diabetes* **51**, S74–S82 (2002).
37. Hutton, J. C. The insulin secretory granule. *Diabetologia* **32**, 271–281 (1989).
38. Weiss, M., Steiner, D. F. & Philipson, L. H. Insulin Biosynthesis, Secretion, Structure, and Structure-Activity Relationships. in *Endotext* (eds. Feingold, K. R. et al.) (MDText.com, Inc., 2000).
39. Boland, B. B., Rhodes, C. J. & Grimsby, J. S. The dynamic plasticity of insulin production in  $\beta$ -cells. *Mol. Metab.* **6**, 958–973 (2017).
40. Matschinsky, F. M. & Wilson, D. F. The Central Role of Glucokinase in Glucose Homeostasis: A Perspective 50 Years After Demonstrating the Presence of the Enzyme in Islets of Langerhans. *Front. Physiol.* **10**, (2019).
41. Goody, R. S. & Mannherz, H. G. The Role of the Cytoskeleton in Transport and Release of Insulin-Containing Granules by Pancreatic  $\beta$ -Cells. in *BetaSys: Systems Biology of Regulated Exocytosis in Pancreatic  $\beta$ -Cells* (eds. Booß-Bavnbek, B., Klösgen, B., Larsen, J., Pociot, F. & Renström, E.) 83–95 (Springer New York, 2011). doi:10.1007/978-1-4419-6956-9\_5.
42. Rizo, J. & Xu, J. The Synaptic Vesicle Release Machinery. *Annu. Rev. Biophys.* **44**, 339–367 (2015).
43. Tang, J. *et al.* A complexin/synaptotagmin 1 switch controls fast synaptic vesicle exocytosis. *Cell* **126**, 1175–1187 (2006).
44. Gaisano, H. Y. Here come the newcomer granules, better late than never. *Trends Endocrinol. Metab.* **25**, 381–388 (2014).
45. Söllner, T., Bennett, M. K., Whiteheart, S. W., Scheller, R. H. & Rothman, J. E. A protein assembly-disassembly pathway in vitro that may correspond to sequential steps of synaptic vesicle docking, activation, and fusion. *Cell* **75**, 409–418 (1993).
46. Henquin, J. C. Triggering and amplifying pathways of regulation of insulin secretion by glucose. *Diabetes* **49**, 1751–1760 (2000).
47. Jones, P. M., Persaud, S. J. & Howell, S. L. Time-course of Ca<sup>2+</sup>-induced insulin secretion from perfused, electrically permeabilised islets of Langerhans: effects of cAMP and a phorbol ester. *Biochem. Biophys. Res. Commun.* **162**, 998–1003 (1989).
48. Gembal, M., Gilon, P. & Henquin, J. C. Evidence that glucose can control insulin release independently from its action on ATP-sensitive K<sup>+</sup> channels in mouse B cells. *J. Clin. Invest.* **89**, 1288–1295 (1992).
49. Kalwat, M. A. & Cobb, M. H. Mechanisms of the amplifying pathway of insulin secretion in the  $\beta$  cell. *Pharmacol. Ther.* **179**, 17–30 (2017).
50. Kolczynska, K., Loza-Valdes, A., Hawro, I. & Sumara, G. Diacylglycerol-evoked activation of PKC and PKD isoforms in regulation of glucose and lipid metabolism: a review. *Lipids Health Dis.* **19**, 113 (2020).

51. Dixit, S. S. *et al.* Effects of CaMKII-Mediated Phosphorylation of Ryanodine Receptor Type 2 on Islet Calcium Handling, Insulin Secretion, and Glucose Tolerance. *PLoS ONE* **8**, e58655 (2013).
52. Easom, R. A. CaM kinase II: a protein kinase with extraordinary talents germane to insulin exocytosis. *Diabetes* **48**, 675–684 (1999).
53. Tengholm, A. & Gylfe, E. cAMP signalling in insulin and glucagon secretion. *Diabetes Obes. Metab.* **19 Suppl 1**, 42–53 (2017).
54. Light, P. E., Manning Fox, J. E., Riedel, M. J. & Wheeler, M. B. Glucagon-like peptide-1 inhibits pancreatic ATP-sensitive potassium channels via a protein kinase A- and ADP-dependent mechanism. *Mol. Endocrinol. Baltim. Md* **16**, 2135–2144 (2002).
55. Weiss Sharon, Oz Shimrit, Benmocha Adva, & Dascal Nathan. Regulation of Cardiac L-Type Ca<sup>2+</sup> Channel CaV1.2 Via the  $\beta$ -Adrenergic-cAMP-Protein Kinase A Pathway. *Circ. Res.* **113**, 617–631 (2013).
56. Kang, G. *et al.* cAMP sensor Epac as a determinant of ATP-sensitive potassium channel activity in human pancreatic beta cells and rat INS-1 cells. *J. Physiol.* **573**, 595–609 (2006).
57. Dyachok, O. & Gylfe, E. Ca(2+)-induced Ca(2+) release via inositol 1,4,5-trisphosphate receptors is amplified by protein kinase A and triggers exocytosis in pancreatic beta-cells. *J. Biol. Chem.* **279**, 45455–45461 (2004).
58. Dzhura, I. *et al.* Epac2-dependent mobilization of intracellular Ca<sup>2+</sup> by glucagon-like peptide-1 receptor agonist exendin-4 is disrupted in  $\beta$ -cells of phospholipase C- $\epsilon$  knockout mice. *J. Physiol.* **588**, 4871–4889 (2010).
59. Wan, Q.-F. *et al.* Protein Kinase Activation Increases Insulin Secretion by Sensitizing the Secretory Machinery to Ca<sup>2+</sup>. *J. Gen. Physiol.* **124**, 653–662 (2004).
60. Vikman, J. *et al.* Truncation of SNAP-25 reduces the stimulatory action of cAMP on rapid exocytosis in insulin-secreting cells. *Am. J. Physiol. Endocrinol. Metab.* **297**, E452-461 (2009).
61. Persaud, S. J., Muller, D. & Jones, P. M. Insulin signalling in islets. *Biochem. Soc. Trans.* **36**, 290–293 (2008).
62. Leibiger, B. *et al.* Selective Insulin Signaling through A and B Insulin Receptors Regulates Transcription of Insulin and Glucokinase Genes in Pancreatic  $\beta$  Cells. *Mol. Cell* **7**, 559–570 (2001).
63. Xu, G. G. & Rothenberg, P. L. Insulin receptor signaling in the beta-cell influences insulin gene expression and insulin content: evidence for autocrine beta-cell regulation. *Diabetes* **47**, 1243–1252 (1998).
64. Banarar, S., McGregor, V. P. & Cryer, P. E. Intra-islet hyperinsulinemia prevents the glucagon response to hypoglycemia despite an intact autonomic response. *Diabetes* **51**, 958–965 (2002).
65. Kawamori, D. *et al.* Insulin signaling in alpha cells modulates glucagon secretion in vivo. *Cell Metab.* **9**, 350–361 (2009).

66. Franklin, I., Gromada, J., Gjinovci, A., Theander, S. & Wollheim, C. B. Beta-cell secretory products activate alpha-cell ATP-dependent potassium channels to inhibit glucagon release. *Diabetes* **54**, 1808–1815 (2005).
67. Xu, E. *et al.* Intra-islet insulin suppresses glucagon release via GABA-GABAA receptor system. *Cell Metab.* **3**, 47–58 (2006).
68. Elliott, A. D., Ustione, A. & Piston, D. W. Somatostatin and insulin mediate glucose-inhibited glucagon secretion in the pancreatic  $\alpha$ -cell by lowering cAMP. *Am. J. Physiol. Endocrinol. Metab.* **308**, E130-143 (2015).
69. Honey, R. N., Fallon, M. B. & Weir, G. C. Effects of exogenous insulin, glucagon, and somatostatin on islet hormone secretion in the perfused chicken pancreas. *Metabolism.* **29**, 1242–1246 (1980).
70. Gerber, P. P., Trimble, E. R., Wollheim, C. B. & Renold, A. E. Effect of insulin on glucose- and arginine-stimulated somatostatin secretion from the isolated perfused rat pancreas. *Endocrinology* **109**, 279–283 (1981).
71. Hauge-Evans, A. C., Anderson, R. L., Persaud, S. J. & Jones, P. M. Delta cell secretory responses to insulin secretagogues are not mediated indirectly by insulin. *Diabetologia* **55**, 1995–2004 (2012).
72. Braun, M. *et al.*  $\gamma$ -Aminobutyric Acid (GABA) Is an Autocrine Excitatory Transmitter in Human Pancreatic  $\beta$ -Cells. *Diabetes* **59**, 1694–1701 (2010).
73. van der Meulen, T. *et al.* Urocortin3 mediates somatostatin-dependent negative feedback control of insulin secretion. *Nat. Med.* **21**, 769–776 (2015).
74. Briant, L. J. B. *et al.*  $\delta$ -cells and  $\beta$ -cells are electrically coupled and regulate  $\alpha$ -cell activity via somatostatin. *J. Physiol.* **596**, 197–215 (2018).
75. Hartig, S. M. & Cox, A. R. Paracrine signaling in islet function and survival. *J. Mol. Med. Berl. Ger.* (2020) doi:10.1007/s00109-020-01887-x.
76. Moens, K. *et al.* Dual glucagon recognition by pancreatic beta-cells via glucagon and glucagon-like peptide 1 receptors. *Diabetes* **47**, 66–72 (1998).
77. Marie, J. C., Boissard, C., Skoglund, G., Rosselin, G. & Breant, B. Glucagon acts through its own receptors in the presence of functional glucagon-like peptide-1 receptors on hamster insulinoma. *Endocrinology* **137**, 4108–4114 (1996).
78. Svendsen, B. *et al.* Insulin Secretion Depends on Intra-islet Glucagon Signaling. *Cell Rep.* **25**, 1127-1134.e2 (2018).
79. Svendsen, B. & Holst, J. J. Paracrine regulation of somatostatin secretion by insulin and glucagon in mouse pancreatic islets. *Diabetologia* **64**, 142–151 (2021).
80. Berts, A., Ball, A., Dryselius, G., Gylfe, E. & Hellman, B. Glucose stimulation of somatostatin-producing islet cells involves oscillatory Ca<sup>2+</sup> signaling. *Endocrinology* **137**, 693–697 (1996).
81. Patton, G. S. *et al.* Pancreatic immunoreactive somatostatin release. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 2140–2143 (1977).

82. Ipp, E. *et al.* Release of immunoreactive somatostatin from the pancreas in response to glucose, amino acids, pancreozymin-cholecystokinin, and tolbutamide. *J. Clin. Invest.* **60**, 760–765 (1977).
83. Hauge-Evans, A. C. *et al.* Somatostatin secreted by islet delta-cells fulfills multiple roles as a paracrine regulator of islet function. *Diabetes* **58**, 403–411 (2009).
84. Kailey, B. *et al.* SSTR2 is the functionally dominant somatostatin receptor in human pancreatic  $\beta$ - and  $\alpha$ -cells. *Am. J. Physiol. Endocrinol. Metab.* **303**, E1107–1116 (2012).
85. Khandekar, N., Berning, B. A., Sainsbury, A. & Lin, S. The role of pancreatic polypeptide in the regulation of energy homeostasis. *Mol. Cell. Endocrinol.* **418**, 33–41 (2015).
86. Aragón, F. *et al.* Pancreatic polypeptide regulates glucagon release through PPYR1 receptors expressed in mouse and human alpha-cells. *Biochim. Biophys. Acta* **1850**, 343–351 (2015).
87. Khan, D., Vasu, S., Moffett, R. C., Irwin, N. & Flatt, P. R. Influence of neuropeptide Y and pancreatic polypeptide on islet function and beta-cell survival. *Biochim. Biophys. Acta Gen. Subj.* **1861**, 749–758 (2017).
88. Adriaenssens, A. E. *et al.* Transcriptomic profiling of pancreatic alpha, beta and delta cell populations identifies delta cells as a principal target for ghrelin in mouse islets. *Diabetologia* **59**, 2156–2165 (2016).
89. Drucker, D. J. The biology of incretin hormones. *Cell Metab.* **3**, 153–165 (2006).
90. Müller, T. D. *et al.* Glucagon-like peptide 1 (GLP-1). *Mol. Metab.* **30**, 72–130 (2019).
91. De Marinis, Y. Z. *et al.* GLP-1 inhibits and adrenaline stimulates glucagon release by differential modulation of N- and L-type Ca<sup>2+</sup> channel-dependent exocytosis. *Cell Metab.* **11**, 543–553 (2010).
92. Guida, C. *et al.* Promiscuous receptor activation mediates glucagonostatic effects of GLP-1(9-36) and GLP-1(7-36). *bioRxiv* 785667 (2019) doi:10.1101/785667.
93. Moens, K. *et al.* Expression and functional activity of glucagon, glucagon-like peptide I, and glucose-dependent insulinotropic peptide receptors in rat pancreatic islet cells. *Diabetes* **45**, 257–261 (1996).
94. Winzell, M. S. & Ahrén, B. Role of VIP and PACAP in islet function. *Peptides* **28**, 1805–1813 (2007).
95. Gromada, J. *et al.* Glucagon-like peptide 1 (7-36) amide stimulates exocytosis in human pancreatic beta-cells by both proximal and distal regulatory steps in stimulus-secretion coupling. *Diabetes* **47**, 57–65 (1998).
96. Ehses, J. A., Lee, S. S., Pederson, R. A. & McIntosh, C. H. A new pathway for glucose-dependent insulinotropic polypeptide (GIP) receptor signaling: evidence for the involvement of phospholipase A2 in GIP-stimulated insulin secretion. *J. Biol. Chem.* **276**, 23667–23673 (2001).
97. El, K. & Campbell, J. E. The role of GIP in  $\alpha$ -cells and glucagon secretion. *Peptides* **125**, 170213 (2020).
98. Chia, C. W. *et al.* Exogenous glucose-dependent insulinotropic polypeptide worsens post prandial hyperglycemia in type 2 diabetes. *Diabetes* **58**, 1342–1349 (2009).

99. Petersen, M. C. & Shulman, G. I. Mechanisms of Insulin Action and Insulin Resistance. *Physiol. Rev.* **98**, 2133–2223 (2018).
100. Hubbard, S. R. The Insulin Receptor: Both a Prototypical and Atypical Receptor Tyrosine Kinase. *Cold Spring Harb. Perspect. Biol.* **5**, a008946 (2013).
101. Taniguchi, C. M., Emanuelli, B. & Kahn, C. R. Critical nodes in signalling pathways: insights into insulin action. *Nat. Rev. Mol. Cell Biol.* **7**, 85–96 (2006).
102. Wu, X. & Garvey, W. T. Insulin Action. in *Textbook of Diabetes* 104–125 (John Wiley & Sons, Ltd, 2010). doi:10.1002/9781444324808.ch7.
103. Myers, M. G. *et al.* IRS-1 activates phosphatidylinositol 3'-kinase by associating with src homology 2 domains of p85. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 10350–10354 (1992).
104. Huang, X., Liu, G., Guo, J. & Su, Z. The PI3K/AKT pathway in obesity and type 2 diabetes. *Int. J. Biol. Sci.* **14**, 1483–1496 (2018).
105. Leto, D. & Saltiel, A. R. Regulation of glucose transport by insulin: traffic control of GLUT4. *Nat. Rev. Mol. Cell Biol.* **13**, 383–396 (2012).
106. Santos, J. M., Ribeiro, S. B., Gaya, A. R., Appell, H.-J. & Duarte, J. A. Skeletal muscle pathways of contraction-enhanced glucose uptake. *Int. J. Sports Med.* **29**, 785–794 (2008).
107. Mu, J., Brozinick, J. T., Valladares, O., Bucan, M. & Birnbaum, M. J. A role for AMP-activated protein kinase in contraction- and hypoxia-regulated glucose transport in skeletal muscle. *Mol. Cell* **7**, 1085–1094 (2001).
108. Printz, R. L. *et al.* Hexokinase II mRNA and gene structure, regulation by insulin, and evolution. *J. Biol. Chem.* **268**, 5209–5219 (1993).
109. Jensen, J., Rustad, P., Kolnes, A. & Lai, Y.-C. The Role of Skeletal Muscle Glycogen Breakdown for Regulation of Insulin Sensitivity by Exercise. *Front. Physiol.* **2**, 112 (2011).
110. Newgard, C. B., Brady, M. J., O'Doherty, R. M. & Saltiel, A. R. Organizing glucose disposal: emerging roles of the glycogen targeting subunits of protein phosphatase-1. *Diabetes* **49**, 1967–1977 (2000).
111. Cohen, P. & Frame, S. The renaissance of GSK3. *Nat. Rev. Mol. Cell Biol.* **2**, 769–776 (2001).
112. Zhang, J. N., Hiken, J., Davis, A. E. & Lawrence, J. C. Insulin stimulates dephosphorylation of phosphorylase in rat epitrochlearis muscles. *J. Biol. Chem.* **264**, 17513–17523 (1989).
113. Zhang, T. *et al.* Acetylation negatively regulates glycogen phosphorylase by recruiting protein phosphatase 1. *Cell Metab.* **15**, 75–87 (2012).
114. Duncan, R. E., Ahmadian, M., Jaworski, K., Sarkadi-Nagy, E. & Sul, H. S. Regulation of Lipolysis in Adipocytes. *Annu. Rev. Nutr.* **27**, 79–101 (2007).
115. Brunet, A. *et al.* Akt promotes cell survival by phosphorylating and inhibiting a Forkhead transcription factor. *Cell* **96**, 857–868 (1999).
116. Matsumoto, M. *et al.* Role of the Insulin Receptor Substrate 1 and Phosphatidylinositol 3-Kinase Signaling Pathway in Insulin-Induced Expression of Sterol Regulatory Element Binding Protein 1c and Glucokinase Genes in Rat Hepatocytes. *Diabetes* **51**, 1672–1680 (2002).

117. Witters, L. A., Watts, T. D., Daniels, D. L. & Evans, J. L. Insulin stimulates the dephosphorylation and activation of acetyl-CoA carboxylase. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 5473–5477 (1988).
118. Witters, L. A. & Kemp, B. E. Insulin activation of acetyl-CoA carboxylase accompanied by inhibition of the 5'-AMP-activated protein kinase. *J. Biol. Chem.* **267**, 2864–2867 (1992).
119. Hughes, K., Ramakrishna, S., Benjamin, W. B. & Woodgett, J. R. Identification of multifunctional ATP-citrate lyase kinase as the alpha-isoform of glycogen synthase kinase-3. *Biochem. J.* **288**, 309–314 (1992).
120. Berwick, D. C., Hers, I., Heesom, K. J., Moule, S. K. & Tavaré, J. M. The identification of ATP-citrate lyase as a protein kinase B (Akt) substrate in primary adipocytes. *J. Biol. Chem.* **277**, 33895–33900 (2002).
121. Marín-Peñalver, J. J., Martín-Timón, I., Sevillano-Collantes, C. & del Cañizo-Gómez, F. J. Update on the treatment of type 2 diabetes mellitus. *World J. Diabetes* **7**, 354–395 (2016).
122. Reed, J., Bain, S. & Kanamarlapudi, V. A Review of Current Trends with Type 2 Diabetes Epidemiology, Aetiology, Pathogenesis, Treatments and Future Perspectives. *Diabetes Metab. Syndr. Obes. Targets Ther.* **14**, 3567–3602 (2021).
123. Inzucchi, S. E. *et al.* Management of Hyperglycemia in Type 2 Diabetes, 2015: A Patient-Centered Approach: Update to a Position Statement of the American Diabetes Association and the European Association for the Study of Diabetes. *Diabetes Care* **38**, 140–149 (2015).
124. Look AHEAD Research Group *et al.* Reduction in weight and cardiovascular disease risk factors in individuals with type 2 diabetes: one-year results of the look AHEAD trial. *Diabetes Care* **30**, 1374–1383 (2007).
125. Andrews, R. C. *et al.* Diet or diet plus physical activity versus usual care in patients with newly diagnosed type 2 diabetes: the Early ACTID randomised controlled trial. *Lancet Lond. Engl.* **378**, 129–139 (2011).
126. Nguyen, N. T., Nguyen, X.-M. T., Lane, J. & Wang, P. Relationship Between Obesity and Diabetes in a US Adult Population: Findings from the National Health and Nutrition Examination Survey, 1999–2006. *Obes. Surg.* **21**, 351–355 (2011).
127. Escalante-Pulido, M., Escalante-Herrera, A., Milke-Najar, M. E. & Alpizar-Salazar, M. Effects of weight loss on insulin secretion and in vivo insulin sensitivity in obese diabetic and non-diabetic subjects. *Diabetes Nutr. Metab.* **16**, 277–283 (2003).
128. Guiding Principles for the Care of People With or at Risk for Diabetes | NIDDK. *National Institute of Diabetes and Digestive and Kidney Diseases* <https://www.niddk.nih.gov/health-information/professionals/clinical-tools-patient-management/diabetes/guiding-principles-care-people-risk-diabetes>.
129. Phielix, E., Meex, R., Moonen-Kornips, E., Hesselink, M. K. C. & Schrauwen, P. Exercise training increases mitochondrial content and ex vivo mitochondrial function similarly in patients with type 2 diabetes and in control individuals. *Diabetologia* **53**, 1714–1721 (2010).

130. Colberg, S. R. *et al.* Exercise and type 2 diabetes: the American College of Sports Medicine and the American Diabetes Association: joint position statement. *Diabetes Care* **33**, e147-167 (2010).
131. Duclos, M. *et al.* Physical activity and type 2 diabetes. Recommendations of the SFD (Francophone Diabetes Society) diabetes and physical activity working group. *Diabetes Metab.* **39**, 205–216 (2013).
132. Rena, G., Hardie, D. G. & Pearson, E. R. The mechanisms of action of metformin. *Diabetologia* **60**, 1577–1585 (2017).
133. Eldor, R. & Raz, I. Diabetes therapy--focus on Asia: second-line therapy debate: insulin/secretagogues. *Diabetes Metab. Res. Rev.* **28 Suppl 2**, 85–89 (2012).
134. Inzucchi, S. E. *et al.* Management of Hyperglycemia in Type 2 Diabetes: A Patient-Centered Approach. *Diabetes Care* **35**, 1364–1379 (2012).
135. Ferrannini, E. & DeFronzo, R. A. Impact of glucose-lowering drugs on cardiovascular disease in type 2 diabetes. *Eur. Heart J.* **36**, 2288–2296 (2015).
136. Lau, D. C. W. & Teoh, H. Impact of Current and Emerging Glucose-Lowering Drugs on Body Weight in Type 2 Diabetes. *Can. J. Diabetes* **39 Suppl 5**, S148-154 (2015).
137. Garber, A. J. *et al.* CONSENSUS STATEMENT BY THE AMERICAN ASSOCIATION OF CLINICAL ENDOCRINOLOGISTS AND AMERICAN COLLEGE OF ENDOCRINOLOGY ON THE COMPREHENSIVE TYPE 2 DIABETES MANAGEMENT ALGORITHM--2016 EXECUTIVE SUMMARY. *Endocr. Pract. Off. J. Am. Coll. Endocrinol. Am. Assoc. Clin. Endocrinol.* **22**, 84–113 (2016).
138. Takahashi, A. *et al.* Sulfonylurea and glinide reduce insulin content, functional expression of K(ATP) channels, and accelerate apoptotic beta-cell death in the chronic phase. *Diabetes Res. Clin. Pract.* **77**, 343–350 (2007).
139. McIntosh, B. *et al.* Second-line therapy in patients with type 2 diabetes inadequately controlled with metformin monotherapy: a systematic review and mixed-treatment comparison meta-analysis. *Open Med. Peer-Rev. Indep. Open-Access J.* **5**, e35-48 (2011).
140. Abe, M., Okada, K. & Soma, M. Antidiabetic agents in patients with chronic kidney disease and end-stage renal disease on dialysis: metabolism and clinical practice. *Curr. Drug Metab.* **12**, 57–69 (2011).
141. Rosak, C. & Mertes, G. Critical evaluation of the role of acarbose in the treatment of diabetes: patient considerations. *Diabetes Metab. Syndr. Obes. Targets Ther.* **5**, 357–367 (2012).
142. Amori, R. E., Lau, J. & Pittas, A. G. Efficacy and safety of incretin therapy in type 2 diabetes: systematic review and meta-analysis. *JAMA* **298**, 194–206 (2007).
143. Nauck, M. A. Update on developments with SGLT2 inhibitors in the management of type 2 diabetes. *Drug Des. Devel. Ther.* **8**, 1335–1380 (2014).
144. Heerspink, H. J. L., de Zeeuw, D., Wie, L., Leslie, B. & List, J. Dapagliflozin a glucose-regulating drug with diuretic properties in subjects with type 2 diabetes. *Diabetes Obes. Metab.* **15**, 853–862 (2013).

145. Ruggenti, P. *et al.* Glomerular hyperfiltration and renal disease progression in type 2 diabetes. *Diabetes Care* **35**, 2061–2068 (2012).
146. Rosenwasser, R. F., Sultan, S., Sutton, D., Choksi, R. & Epstein, B. J. SGLT-2 inhibitors and their potential in the treatment of diabetes. *Diabetes Metab. Syndr. Obes. Targets Ther.* **6**, 453–467 (2013).
147. Cahn, A., Miccoli, R., Dardano, A. & Del Prato, S. New forms of insulin and insulin therapies for the treatment of type 2 diabetes. *Lancet Diabetes Endocrinol.* **3**, 638–652 (2015).
148. Silver, B. *et al.* EADSG Guidelines: Insulin Therapy in Diabetes. *Diabetes Ther.* **9**, 449–492 (2018).
149. Horvath, K. *et al.* Long-acting insulin analogues versus NPH insulin (human isophane insulin) for type 2 diabetes mellitus. *Cochrane Database Syst. Rev.* CD005613 (2007) doi:10.1002/14651858.CD005613.pub3.
150. Bonnefond, A. & Froguel, P. Rare and Common Genetic Events in Type 2 Diabetes: What Should Biologists Know? *Cell Metab.* **21**, 357–368 (2015).
151. Bonnefond, A. & Froguel, P. Rare and common genetic events in type 2 diabetes: what should biologists know? *Cell Metab.* **21**, 357–368 (2015).
152. Ali, O. Genetics of type 2 diabetes. *World J. Diabetes* **4**, 114–123 (2013).
153. Génétique. [http://www.edu.upmc.fr/sdv/masselot\\_05001/polymorphisme/rflp.html](http://www.edu.upmc.fr/sdv/masselot_05001/polymorphisme/rflp.html).
154. Les marqueurs moléculaires. *Semence Pédagogie* <https://www.gnis-pedagogie.org/sujet/marqueurs-moleculaires/>.
155. Sanger Sequencing Steps & Method. <https://www.sigmaaldrich.com/BE/en/technical-documents/protocol/genomics/sequencing/sanger-sequencing>.
156. Vaxillaire, M. & Froguel, P. Facteurs génétiques et gènes candidats dans la physiopathologie du diabète de type 2. *Corresp. En MHDN Métabolismes Horm. Diabètes Nutr.* **VII**, 11 (2003).
157. LaFramboise, T. Single nucleotide polymorphism arrays: a decade of biological, computational and technological advances. *Nucleic Acids Res.* **37**, 4181–4193 (2009).
158. EMBL-EBI. What are genome wide association studies (GWAS)? | GWAS Catalog. <https://www.ebi.ac.uk/training/online/courses/gwas-catalogue-exploring-snp-trait-associations/what-is-gwas-catalog/what-are-genome-wide-association-studies-gwas/>.
159. Morris, A. P. *et al.* Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat. Genet.* **44**, 981–990 (2012).
160. Smemo, S. *et al.* Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature* **507**, 371–375 (2014).
161. The Human Genome Project. *Genome.gov* <https://www.genome.gov/human-genome-project>.
162. WES vs WGS vs Custom Panels. <https://sequencing.roche.com/content/rochesequence/en-us/science-education/education/articles/wes-wgs-custom.html>.

163. Le séquençage de nouvelle génération // Sacha Schutz // bioinformatique génétique médecine. <https://dridk.me/ngs.html>.
164. Zuk, O. *et al.* Searching for missing heritability: designing rare variant association studies. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E455-464 (2014).
165. Murphy, R., Ellard, S. & Hattersley, A. T. Clinical implications of a molecular genetic classification of monogenic beta-cell diabetes. *Nat. Clin. Pract. Endocrinol. Metab.* **4**, 200–213 (2008).
166. Misra, S. & Owen, K. R. Genetics of Monogenic Diabetes: Present Clinical Challenges. *Curr. Diab. Rep.* **18**, 141 (2018).
167. Vaxillaire, M. & Froguel, P. Monogenic diabetes: Implementation of translational genomic research towards precision medicine. *J. Diabetes* **8**, 782–795 (2016).
168. Vaxillaire, M. & Froguel, P. Genetic basis of maturity-onset diabetes of the young. *Endocrinol. Metab. Clin. North Am.* **35**, 371–384, x (2006).
169. Vaxillaire, M. & Froguel, P. Monogenic diabetes in the young, pharmacogenetics and relevance to multifactorial forms of type 2 diabetes. *Endocr. Rev.* **29**, 254–264 (2008).
170. Shields, B. M. *et al.* Maturity-onset diabetes of the young (MODY): how many cases are we missing? *Diabetologia* **53**, 2504–2508 (2010).
171. Hattersley, A. T. & Patel, K. A. Precision diabetes: learning from monogenic diabetes. *Diabetologia* **60**, 769–777 (2017).
172. Naylor, R., Knight Johnson, A. & del Gaudio, D. Maturity-Onset Diabetes of the Young Overview. in *GeneReviews*<sup>®</sup> (eds. Adam, M. P. *et al.*) (University of Washington, Seattle, 2018).
173. Sladek, F. M., Zhong, W. M., Lai, E. & Darnell, J. E. Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily. *Genes Dev.* **4**, 2353–2365 (1990).
174. Xanthopoulos, K. G. *et al.* The different tissue transcription patterns of genes for HNF-1, C/EBP, HNF-3, and HNF-4, protein factors that govern liver-specific transcription. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 3807–3811 (1991).
175. Miquerol, L. *et al.* Expression of the L-type pyruvate kinase gene and the hepatocyte nuclear factor 4 transcription factor in exocrine and endocrine pancreas. *J. Biol. Chem.* **269**, 8944–8951 (1994).
176. Stoffel, M. & Duncan, S. A. The maturity-onset diabetes of the young (MODY1) transcription factor HNF4 $\alpha$  regulates expression of genes required for glucose transport and metabolism. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 13209–13214 (1997).
177. Gupta, R. K. *et al.* The MODY1 gene HNF-4 $\alpha$  regulates selected genes involved in insulin secretion. *J. Clin. Invest.* **115**, 1006–1015 (2005).
178. Bartoov-Shifman, R. *et al.* Activation of the insulin gene promoter through a direct effect of hepatocyte nuclear factor 4 alpha. *J. Biol. Chem.* **277**, 25914–25919 (2002).
179. Boj, S. F., Párrizas, M., Maestro, M. A. & Ferrer, J. A transcription factor regulatory circuit in differentiated pancreatic cells. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 14481–14486 (2001).

180. Thomas, H. *et al.* A distant upstream promoter of the HNF-4 $\alpha$  gene connects the transcription factors involved in maturity-onset diabetes of the young. *Hum. Mol. Genet.* **10**, 2089–2097 (2001).
181. Bell, G. I. *et al.* Gene for non-insulin-dependent diabetes mellitus (maturity-onset diabetes of the young subtype) is linked to DNA polymorphism on human chromosome 20q. *Proc. Natl. Acad. Sci.* **88**, 1484–1488 (1991).
182. Irwin, M., Cox, N. & Kong, A. Sequential imputation for multilocus linkage analysis. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 11684–11688 (1994).
183. Stoffel, M. *et al.* A yeast artificial chromosome-based map of the region of chromosome 20 containing the diabetes-susceptibility gene, MODY1, and a myeloid leukemia related gene. *Proc. Natl. Acad. Sci.* **93**, 3937–3941 (1996).
184. Yamagata, K. *et al.* Mutations in the hepatocyte nuclear factor-4 $\alpha$  gene in maturity-onset diabetes of the young (MODY1). *Nature* **384**, 458–460 (1996).
185. Pearson, E. R. *et al.* Molecular genetics and phenotypic characteristics of MODY caused by hepatocyte nuclear factor 4 $\alpha$  mutations in a large European collection. *Diabetologia* **48**, 878–885 (2005).
186. Delvecchio, M., Pastore, C. & Giordano, P. Treatment Options for MODY Patients: A Systematic Review of Literature. *Diabetes Ther.* **11**, 1667–1685 (2020).
187. Pearson, E. R. *et al.* Macrosomia and hyperinsulinaemic hypoglycaemia in patients with heterozygous mutations in the HNF4A gene. *PLoS Med.* **4**, e118 (2007).
188. Pearson, E. R. *et al.* Genetic cause of hyperglycaemia and response to treatment in diabetes. *Lancet Lond. Engl.* **362**, 1275–1281 (2003).
189. Urakami, T. Maturity-onset diabetes of the young (MODY): current perspectives on diagnosis and treatment. *Diabetes Metab. Syndr. Obes. Targets Ther.* **12**, 1047–1056 (2019).
190. Matschinsky, F. M. Glucokinase, glucose homeostasis, and diabetes mellitus. *Curr. Diab. Rep.* **5**, 171–176 (2005).
191. Froguel, P. *et al.* Close linkage of glucokinase locus on chromosome 7p to early-onset non-insulin-dependent diabetes mellitus. *Nature* **356**, 162–164 (1992).
192. Thomson, K. L. *et al.* Identification of 21 novel glucokinase (GCK) mutations in UK and European Caucasians with maturity-onset diabetes of the young (MODY). *Hum. Mutat.* **22**, 417 (2003).
193. Velho, G. *et al.* Primary pancreatic beta-cell secretory defect caused by mutations in glucokinase gene in kindreds of maturity onset diabetes of the young. *Lancet Lond. Engl.* **340**, 444–448 (1992).
194. Velho, G. *et al.* Impaired hepatic glycogen synthesis in glucokinase-deficient (MODY-2) subjects. *J. Clin. Invest.* **98**, 1755–1761 (1996).
195. Stride, A. *et al.* Cross-sectional and longitudinal studies suggest pharmacological treatment used in patients with glucokinase mutations does not alter glycaemia. *Diabetologia* **57**, 54–56 (2014).

196. Odom, D. T. *et al.* Control of Pancreas and Liver Gene Expression by HNF Transcription Factors. *Science* **303**, 1378–1381 (2004).
197. D'Angelo, A. *et al.* Hepatocyte nuclear factor 1alpha and beta control terminal differentiation and cell fate commitment in the gut epithelium. *Dev. Camb. Engl.* **137**, 1573–1582 (2010).
198. Low, B. S. J. *et al.* Decreased GLUT2 and glucose uptake contribute to insulin secretion defects in MODY3/HNF1A hiPSC-derived mutant  $\beta$  cells. *Nat. Commun.* **12**, 3133 (2021).
199. Vaxillaire, M. *et al.* A gene for maturity onset diabetes of the young (MODY) maps to chromosome 12q. *Nat. Genet.* **9**, 418–423 (1995).
200. Menzel, S. *et al.* Localization of MODY3 to a 5-cM region of human chromosome 12. *Diabetes* **44**, 1408–1413 (1995).
201. Yamagata, K. *et al.* Mutations in the hepatocyte nuclear factor-1alpha gene in maturity-onset diabetes of the young (MODY3). *Nature* **384**, 455–458 (1996).
202. Radha, V. *et al.* Identification of novel variants in the hepatocyte nuclear factor-1alpha gene in South Indian patients with maturity onset diabetes of young. *J. Clin. Endocrinol. Metab.* **94**, 1959–1965 (2009).
203. Stride, A. *et al.* Beta-cell dysfunction, insulin sensitivity, and glycosuria precede diabetes in hepatocyte nuclear factor-1alpha mutation carriers. *Diabetes Care* **28**, 1751–1756 (2005).
204. Isomaa, B. *et al.* Chronic diabetic complications in patients with MODY3 diabetes. *Diabetologia* **41**, 467–473 (1998).
205. Bingham, C. *et al.* The Generalized Aminoaciduria Seen in Patients With Hepatocyte Nuclear Factor-1 $\alpha$  Mutations Is a Feature of All Patients With Diabetes and Is Associated With Glucosuria. *Diabetes* **50**, 2047–2052 (2001).
206. Stoffers, D. A., Thomas, M. K. & Habener, J. F. Homeodomain protein IDX-1: a master regulator of pancreas development and insulin gene expression. *Trends Endocrinol. Metab. TEM* **8**, 145–151 (1997).
207. Schwitzgebel, V. M. *et al.* Agenesis of human pancreas due to decreased half-life of insulin promoter factor 1. *J. Clin. Endocrinol. Metab.* **88**, 4398–4406 (2003).
208. Stoffers, D. A., Ferrer, J., Clarke, W. L. & Habener, J. F. Early-onset type-II diabetes mellitus (MODY4) linked to IPF1. *Nat. Genet.* **17**, 138–139 (1997).
209. Deng, M., Xiao, X., Zhou, L. & Wang, T. First Case Report of Maturity-Onset Diabetes of the Young Type 4 Pedigree in a Chinese Family. *Front. Endocrinol.* **10**, 406 (2019).
210. Maestro, M. A. *et al.* Distinct roles of HNF1beta, HNF1alpha, and HNF4alpha in regulating pancreas development, beta-cell function and growth. *Endocr. Dev.* **12**, 33–45 (2007).
211. Haumaitre, C. *et al.* Lack of TCF2/vHNF1 in mice leads to pancreas agenesis. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 1490–1495 (2005).
212. Bellanné-Chantelot, C. *et al.* Clinical spectrum associated with hepatocyte nuclear factor-1beta mutations. *Ann. Intern. Med.* **140**, 510–517 (2004).
213. Horikawa, Y. *et al.* Mutation in hepatocyte nuclear factor-1 beta gene (TCF2) associated with MODY. *Nat. Genet.* **17**, 384–385 (1997).

214. Kim, E. K. *et al.* Identification and Functional Characterization of P159L Mutation in HNF1B in a Family with Maturity-Onset Diabetes of the Young 5 (MODY5). *Genomics Inform.* **12**, 240–246 (2014).
215. Malecki, M. T. *et al.* Mutations in NEUROD1 are associated with the development of type 2 diabetes mellitus. *Nat. Genet.* **23**, 323–328 (1999).
216. Kristinsson, S. Y. *et al.* MODY in Iceland is associated with mutations in HNF-1alpha and a novel mutation in NeuroD1. *Diabetologia* **44**, 2098–2103 (2001).
217. Horikawa, Y. *et al.* NEUROD1-deficient diabetes (MODY6): Identification of the first cases in Japanese and the clinical features. *Pediatr. Diabetes* **19**, 236–242 (2018).
218. Rubio-Cabezas, O. *et al.* Homozygous mutations in NEUROD1 are responsible for a novel syndrome of permanent neonatal diabetes and neurological abnormalities. *Diabetes* **59**, 2326–2331 (2010).
219. Horikawa, Y. & Enya, M. Genetic Dissection and Clinical Features of MODY6 (NEUROD1-MODY). *Curr. Diab. Rep.* **19**, 12 (2019).
220. Rubio-Cabezas, O., Flanagan, S. E., Damhuis, A., Hattersley, A. T. & Ellard, S. KATP channel mutations in infants with permanent diabetes diagnosed after 6 months of life. *Pediatr. Diabetes* **13**, 322–325 (2012).
221. Rubio-Cabezas, O. & Ellard, S. Diabetes Mellitus in Neonates and Infants: Genetic Heterogeneity, Clinical Approach to Diagnosis, and Therapeutic Options. *Horm. Res. Paediatrics* **80**, 137–146 (2013).
222. Lemelman, M. B., Letourneau, L. & Greeley, S. A. W. Neonatal Diabetes Mellitus: An Update on Diagnosis and Management. *Clin. Perinatol.* **45**, 41–59 (2018).
223. Wiedemann, B. *et al.* Incidence of neonatal diabetes in Austria-calculation based on the Austrian Diabetes Register. *Pediatr. Diabetes* **11**, 18–23 (2010).
224. INSERM US14-- ALL RIGHTS. Orphanet: Transient neonatal diabetes mellitus. [https://www.orpha.net/consor/cgi-bin/OC\\_Exp.php?Ing=EN&Expert=99886](https://www.orpha.net/consor/cgi-bin/OC_Exp.php?Ing=EN&Expert=99886).
225. Steck, A. K. & Winter, W. E. Review on monogenic diabetes. *Curr. Opin. Endocrinol. Diabetes Obes.* **18**, 252–258 (2011).
226. Greeley, S. A. W. *et al.* Update in neonatal diabetes. *Curr. Opin. Endocrinol. Diabetes Obes.* **17**, 13–19 (2010).
227. Cryns, K. *et al.* Mutational spectrum of the WFS1 gene in Wolfram syndrome, nonsyndromic hearing impairment, diabetes mellitus, and psychiatric disease. *Hum. Mutat.* **22**, 275–287 (2003).
228. Rubio-Cabezas, O. *et al.* Wolcott-Rallison Syndrome Is the Most Common Genetic Cause of Permanent Neonatal Diabetes in Consanguineous Families. *J. Clin. Endocrinol. Metab.* **94**, 4162–4170 (2009).
229. Temple, I. K. & Shield, J. P. H. Transient neonatal diabetes, a disorder of imprinting. *J. Med. Genet.* **39**, 872–875 (2002).
230. Aguilar-Bryan, L. & Bryan, J. Neonatal Diabetes Mellitus. *Endocr. Rev.* **29**, 265–291 (2008).
231. Polak, M. & Cavé, H. Neonatal diabetes mellitus: a disease linked to multiple mechanisms. *Orphanet J. Rare Dis.* **2**, 12 (2007).

232. Polak, M. *et al.* Heterozygous missense mutations in the insulin gene are linked to permanent diabetes appearing in the neonatal period or in early infancy: a report from the French ND (Neonatal Diabetes) Study Group. *Diabetes* **57**, 1115–1119 (2008).
233. Arima, T., Drewell, R. A., Oshimura, M., Wake, N. & Surani, M. A. A Novel Imprinted Gene, HYMAI, Is Located within an Imprinted Domain on Human Chromosome 6 Containing ZAC. *Genomics* **67**, 248–255 (2000).
234. Das, S. *et al.* Partial Paternal Uniparental Disomy of Chromosome 6 in an Infant with Neonatal Diabetes, Macroglossia, and Craniofacial Abnormalities. *Am. J. Hum. Genet.* **67**, 1586–1591 (2000).
235. Cavé, H., Polak, M., Drunat, S., Denamur, E. & Czernichow, P. Refinement of the 6q chromosomal region implicated in transient neonatal diabetes. *Diabetes* **49**, 108–113 (2000).
236. Gardner, R. J. *et al.* An imprinted locus associated with transient neonatal diabetes mellitus. *Hum. Mol. Genet.* **9**, 589–596 (2000).
237. Docherty, L. E. *et al.* Clinical presentation of 6q24 transient neonatal diabetes mellitus (6q24 TNDM) and genotype–phenotype correlation in an international cohort of patients. *Diabetologia* **56**, 758–762 (2013).
238. Gloyn, A. L. *et al.* Relapsing diabetes can result from moderately activating mutations in KCNJ11. *Hum. Mol. Genet.* **14**, 925–934 (2005).
239. Babenko, A. P. *et al.* Activating mutations in the ABCC8 gene in neonatal diabetes mellitus. *N. Engl. J. Med.* **355**, 456–466 (2006).
240. Gloyn, A. L. *et al.* KCNJ11 activating mutations are associated with developmental delay, epilepsy and neonatal diabetes syndrome and other neurological features. *Eur. J. Hum. Genet. EJHG* **14**, 824–830 (2006).
241. Koster, J. C. *et al.* The G53D mutation in Kir6.2 (KCNJ11) is associated with neonatal diabetes and motor dysfunction in adulthood that is improved with sulfonylurea therapy. *J. Clin. Endocrinol. Metab.* **93**, 1054–1061 (2008).
242. Edghill, E. L. *et al.* Insulin mutation screening in 1,044 patients with diabetes: mutations in the INS gene are a common cause of neonatal diabetes but a rare cause of diabetes diagnosed in childhood or adulthood. *Diabetes* **57**, 1034–1042 (2008).
243. Støy, J. *et al.* Insulin gene mutations as a cause of permanent neonatal diabetes. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 15040–15044 (2007).
244. Garin, I. *et al.* Recessive mutations in the INS gene result in neonatal diabetes through reduced insulin biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 3105–3110 (2010).
245. Donath, X. *et al.* Next-generation sequencing identifies monogenic diabetes in 16% of patients with late adolescence/adult-onset diabetes selected on a clinical basis: a cross-sectional analysis. *BMC Med.* **17**, 132 (2019).
246. 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. Off. J. Am. Coll. Med. Genet.* **17**, 405–424 (2015).

247. Almgren, P. *et al.* Heritability and familiarity of type 2 diabetes and related quantitative traits in the Botnia Study. *Diabetologia* **54**, 2811–2819 (2011).
248. Poulsen, P., Kyvik, K. O., Vaag, A. & Beck-Nielsen, H. Heritability of type II (non-insulin-dependent) diabetes mellitus and abnormal glucose tolerance--a population-based twin study. *Diabetologia* **42**, 139–145 (1999).
249. Willemsen, G. *et al.* The Concordance and Heritability of Type 2 Diabetes in 34,166 Twin Pairs From International Twin Registers: The Discordant Twin (DISCOTWIN) Consortium. *Twin Res. Hum. Genet. Off. J. Int. Soc. Twin Stud.* **18**, 762–771 (2015).
250. Flannick, J. The Contribution of Low-Frequency and Rare Coding Variation to Susceptibility to Type 2 Diabetes. *Curr. Diab. Rep.* **19**, 25 (2019).
251. Hanis, C. L. *et al.* A genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome 2. *Nat. Genet.* **13**, 161–166 (1996).
252. Vasseur, F. *et al.* Single-nucleotide polymorphism haplotypes in the both proximal promoter and exon 3 of the APM1 gene modulate adipocyte-secreted adiponectin hormone levels and contribute to the genetic risk for type 2 diabetes in French Caucasians. *Hum. Mol. Genet.* **11**, 2607–2614 (2002).
253. Permutt, M. A. *et al.* A Genome Scan for Type 2 Diabetes Susceptibility Loci in a Genetically Isolated Population. *Diabetes* **50**, 681–685 (2001).
254. Silander, K. *et al.* Genetic Variation Near the Hepatocyte Nuclear Factor-4 $\alpha$  Gene Predicts Susceptibility to Type 2 Diabetes. *Diabetes* **53**, 1141–1149 (2004).
255. Meyre, D. *et al.* Variants of ENPP1 are associated with childhood and adult obesity and increase the risk of glucose intolerance and type 2 diabetes. *Nat. Genet.* **37**, 863–867 (2005).
256. Grant, S. F. A. *et al.* Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat. Genet.* **38**, 320–323 (2006).
257. Sladek, R. *et al.* A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* **445**, 881–885 (2007).
258. Kooner, J. S. *et al.* Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. *Nat. Genet.* **43**, 984–989 (2011).
259. März, W. *et al.* G(-30)A polymorphism in the pancreatic promoter of the glucokinase gene associated with angiographic coronary artery disease and type 2 diabetes mellitus. *Circulation* **109**, 2844–2849 (2004).
260. Sandhu, M. S. *et al.* Common variants in WFS1 confer risk of type 2 diabetes. *Nat. Genet.* **39**, 951–953 (2007).
261. Winckler, W. *et al.* Evaluation of common variants in the six known maturity-onset diabetes of the young (MODY) genes for association with type 2 diabetes. *Diabetes* **56**, 685–693 (2007).
262. Hani, E. H. *et al.* Missense mutations in the pancreatic islet beta cell inwardly rectifying K<sup>+</sup> channel gene (KIR6.2/BIR): a meta-analysis suggests a role in the polygenic basis of Type II diabetes mellitus in Caucasians. *Diabetologia* **41**, 1511–1515 (1998).

263. Bonnefond, A. *et al.* Whole-exome sequencing and high throughput genotyping identified KCNJ11 as the thirteenth MODY gene. *PLoS One* **7**, e37423 (2012).
264. Barroso, I. *et al.* Dominant negative mutations in human PPARgamma associated with severe insulin resistance, diabetes mellitus and hypertension. *Nature* **402**, 880–883 (1999).
265. Deeb, S. S. *et al.* A Pro12Ala substitution in PPARgamma2 associated with decreased receptor activity, lower body mass index and improved insulin sensitivity. *Nat. Genet.* **20**, 284–287 (1998).
266. Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, and Novartis Institutes of BioMedical Research *et al.* Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* **316**, 1331–1336 (2007).
267. Qi, L. *et al.* Genetic variants at 2q24 are associated with susceptibility to type 2 diabetes. *Hum. Mol. Genet.* **19**, 2706–2715 (2010).
268. Scott, L. J. *et al.* A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* **316**, 1341–1345 (2007).
269. Zeggini, E. *et al.* Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* **316**, 1336–1341 (2007).
270. Zeggini, E. *et al.* Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat. Genet.* **40**, 638–645 (2008).
271. Dina, C. *et al.* Variation in FTO contributes to childhood obesity and severe adult obesity. *Nat. Genet.* **39**, 724–726 (2007).
272. 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).
273. 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).
274. Bouatia-Naji, N. *et al.* A variant near MTNR1B is associated with increased fasting plasma glucose levels and type 2 diabetes risk. *Nat. Genet.* **41**, 89–94 (2009).
275. Dupuis, J. *et al.* New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat. Genet.* **42**, 105–116 (2010).
276. Prokopenko, I. *et al.* Variants in MTNR1B influence fasting glucose levels. *Nat. Genet.* **41**, 77–81 (2009).
277. Fuchsberger, C. *et al.* The genetic architecture of type 2 diabetes. *Nature* **536**, 41–47 (2016).
278. Flannick, J. *et al.* Exome sequencing of 20,791 cases of type 2 diabetes and 24,440 controls. *Nature* **570**, 71–76 (2019).
279. Bonnefond, A. *et al.* Pathogenic variants in actionable MODY genes are associated with type 2 diabetes. *Nat. Metab.* **2**, 1126–1134 (2020).
280. Pasquali, L. *et al.* Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. *Nat. Genet.* **46**, 136–143 (2014).

281. Jockers, R., Maurice, P., Boutin, J. A. & Delagrangé, P. Melatonin receptors, heterodimerization, signal transduction and binding sites: what's new? *Br. J. Pharmacol.* **154**, 1182–1195 (2008).
282. Dubocovich, M. L. *et al.* International Union of Basic and Clinical Pharmacology. LXXV. Nomenclature, Classification, and Pharmacology of G Protein-Coupled Melatonin Receptors. *Pharmacol. Rev.* **62**, 343–380 (2010).
283. Bonnefond, A. *et al.* Rare MTNR1B variants impairing melatonin receptor 1B function contribute to type 2 diabetes. *Nat. Genet.* **44**, 297–301 (2012).
284. Majithia, A. R. *et al.* Rare variants in PPAR $\gamma$  with decreased activity in adipocyte differentiation are associated with increased risk of type 2 diabetes. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 13127–13132 (2014).
285. Stumvoll, M. & Häring, H. The peroxisome proliferator-activated receptor- $\gamma$ 2 Pro12Ala polymorphism. *Diabetes* **51**, 2341–2347 (2002).
286. SIGMA Type 2 Diabetes Consortium *et al.* Association of a low-frequency variant in HNF1A with type 2 diabetes in a Latino population. *JAMA* **311**, 2305–2314 (2014).
287. Dhaliwal, A. & Gupta, M. Physiology, Opioid Receptor. in *StatPearls* (StatPearls Publishing, 2021).
288. Peng, J., Sarkar, S. & Chang, S. L. Opioid receptor expression in human brain and peripheral tissues using absolute quantitative real-time RT-PCR. *Drug Alcohol Depend.* **124**, 223–228 (2012).
289. Al-Hasani, R. & Bruchas, M. R. Molecular Mechanisms of Opioid Receptor-Dependent Signaling and Behavior. *Anesthesiology* **115**, 1363–1381 (2011).
290. Gendron, L., Cahill, C. M., von Zastrow, M., Schiller, P. W. & Pineyro, G. Molecular Pharmacology of  $\delta$ -Opioid Receptors. *Pharmacol. Rev.* **68**, 631–700 (2016).
291. Pathan, H. & Williams, J. Basic opioid pharmacology: an update. *Br. J. Pain* **6**, 11–16 (2012).
292. Borsodi, A. *et al.* Opioid receptors (version 2019.4) in the IUPHAR/BPS Guide to Pharmacology Database. *IUPHARBPS Guide Pharmacol. CITE* **2019**, (2019).
293. Abdallah, K. & Gendron, L. The Delta Opioid Receptor in Pain Control. in *Delta Opioid Receptor Pharmacology and Therapeutic Applications* (ed. Jutkiewicz, E. M.) 147–177 (Springer International Publishing, 2018). doi:10.1007/164\_2017\_32.
294. Lefkowitz, R. J. & Shenoy, S. K. Transduction of receptor signals by beta-arrestins. *Science* **308**, 512–517 (2005).
295. Kramer, H. K. *et al.* Mutation of tyrosine 318 (Y318F) in the delta-opioid receptor attenuates tyrosine phosphorylation, agonist-dependent receptor internalization, and mitogen-activated protein kinase activation. *Mol. Brain Res.* **79**, 55–66 (2000).
296. Belcheva, M. M. *et al.* Opioid Modulation of Extracellular Signal-Regulated Protein Kinase Activity Is Ras-Dependent and Involves G $\beta\gamma$  Subunits. *J. Neurochem.* **70**, 635–645 (1998).
297. Sriram, K. & Insel, P. A. G Protein-Coupled Receptors as Targets for Approved Drugs: How Many Targets and How Many Drugs? *Mol. Pharmacol.* **93**, 251–258 (2018).

298. Najafipour, H. & Beik, A. The Impact of Opium Consumption on Blood Glucose, Serum Lipids and Blood Pressure, and Related Mechanisms. *Front. Physiol.* **7**, (2016).
299. Bone, R. N. *et al.* A Computational Approach for Defining a Signature of  $\beta$ -Cell Golgi Stress in Diabetes. *Diabetes* **69**, 2364–2376 (2020).
300. Kolic, J. *et al.* Differential Effects of Voclosporin and Tacrolimus on Insulin Secretion From Human Islets. *Endocrinology* **161**, bqaa162 (2020).
301. Mousavy Gharavy, S. N. *et al.* Sexually dimorphic roles for the type 2 diabetes-associated C2cd4b gene in murine glucose homeostasis. *Diabetologia* **64**, 850–864 (2021).
302. Berthault, C., Staels, W. & Scharfmann, R. Purification of pancreatic endocrine subsets reveals increased iron metabolism in beta cells. *Mol. Metab.* **42**, 101060 (2020).
303. Nica, A. C. *et al.* Cell-type, allelic, and genetic signatures in the human pancreatic beta cell transcriptome. *Genome Res.* **23**, 1554–1562 (2013).
304. Jetten, A. M. GLIS1-3 transcription factors: critical roles in the regulation of multiple physiological processes and diseases. *Cell. Mol. Life Sci. CMLS* **75**, 3473–3494 (2018).
305. Lichti-Kaiser, K., ZeRuth, G., Kang, H. S., Vasanth, S. & Jetten, A. M. Gli-similar proteins: their mechanisms of action, physiological functions, and roles in disease. *Vitam. Horm.* **88**, 141–171 (2012).
306. Kim, Y.-S. *et al.* Identification of Glis1, a novel Gli-related, Kruppel-like zinc finger protein containing transactivation and repressor functions. *J. Biol. Chem.* **277**, 30901–30913 (2002).
307. Zhang, F. & Jetten, A. M. Genomic structure of the gene encoding the human Gli-related, Krüppel-like zinc finger protein GLIS2. *Gene* **280**, 49–57 (2001).
308. Zhang, F. *et al.* Characterization of Glis2, a novel gene encoding a Gli-related, Krüppel-like transcription factor with transactivation and repressor functions. Roles in kidney development and neurogenesis. *J. Biol. Chem.* **277**, 10139–10149 (2002).
309. Kim, Y.-S., Nakanishi, G., Lewandoski, M. & Jetten, A. M. GLIS3, a novel member of the GLIS subfamily of Krüppel-like zinc finger proteins with repressor and activation functions. *Nucleic Acids Res.* **31**, 5513–5525 (2003).
310. Senée, V. *et al.* Mutations in GLIS3 are responsible for a rare syndrome with neonatal diabetes mellitus and congenital hypothyroidism. *Nat. Genet.* **38**, 682–687 (2006).
311. Yang, Y., Chang, B. H.-J., Samson, S. L., Li, M. V. & Chan, L. The Krüppel-like zinc finger protein Glis3 directly and indirectly activates insulin gene transcription. *Nucleic Acids Res.* **37**, 2529–2538 (2009).
312. Wen, X. & Yang, Y. Emerging roles of GLIS3 in neonatal diabetes, type 1 and type 2 diabetes. *J. Mol. Endocrinol.* **58**, R73–R85 (2017).
313. Dimitri, P. *et al.* Novel GLIS3 mutations demonstrate an extended multisystem phenotype. *Eur. J. Endocrinol.* **164**, 437–443 (2011).
314. Dimitri, P. *et al.* Expanding the Clinical Spectrum Associated With GLIS3 Mutations. *J. Clin. Endocrinol. Metab.* **100**, E1362–1369 (2015).
315. Dimitri, P. *et al.* An emerging, recognizable facial phenotype in association with mutations in Gli-similar 3 (GLIS3). *Am. J. Med. Genet. A.* **170**, 1918–1923 (2016).

316. Barrett, J. C. *et al.* Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat. Genet.* **41**, 703–707 (2009).
317. Kiani, A. K. *et al.* Association of 32 type 1 diabetes risk loci in Pakistani patients. *Diabetes Res. Clin. Pract.* **108**, 137–142 (2015).
318. Winkler, C. *et al.* Feature ranking of type 1 diabetes susceptibility genes improves prediction of type 1 diabetes. *Diabetologia* **57**, 2521–2529 (2014).
319. Awata, T. *et al.* A low-frequency GLIS3 variant associated with resistance to Japanese type 1 diabetes. *Biochem. Biophys. Res. Commun.* **437**, 521–525 (2013).
320. Shih, H. P., Wang, A. & Sander, M. Pancreas organogenesis: from lineage determination to morphogenesis. *Annu. Rev. Cell Dev. Biol.* **29**, 81–105 (2013).
321. Yang, Y. *et al.* The Krüppel-like zinc finger protein GLIS3 transactivates neurogenin 3 for proper fetal pancreatic islet differentiation in mice. *Diabetologia* **54**, 2595–2605 (2011).
322. Hay, C. W. & Docherty, K. Comparative analysis of insulin gene promoters: implications for diabetes research. *Diabetes* **55**, 3201–3213 (2006).
323. Kang, H. S. *et al.* Transcription factor Glis3, a novel critical player in the regulation of pancreatic beta-cell development and insulin gene expression. *Mol. Cell. Biol.* **29**, 6366–6379 (2009).
324. ZeRuth, G. T., Takeda, Y. & Jetten, A. M. The Krüppel-Like Protein Gli-Similar 3 (Glis3) Functions as a Key Regulator of Insulin Transcription. *Mol. Endocrinol.* **27**, 1692–1705 (2013).
325. Yang, Y., Chang, B. H.-J. & Chan, L. Sustained expression of the transcription factor GLIS3 is required for normal beta cell function in adults. *EMBO Mol. Med.* **5**, 92–104 (2013).
326. Ackermann, A. M. & Gannon, M. Molecular regulation of pancreatic beta-cell mass development, maintenance, and expansion. *J. Mol. Endocrinol.* **38**, 193–206 (2007).
327. Georgia, S. & Bhushan, A. Beta cell replication is the primary mechanism for maintaining postnatal beta cell mass. *J. Clin. Invest.* **114**, 963–968 (2004).
328. Kushner, J. A. *et al.* Cyclins D2 and D1 are essential for postnatal pancreatic beta-cell growth. *Mol. Cell. Biol.* **25**, 3752–3762 (2005).
329. Georgia, S. *et al.* Cyclin D2 is essential for the compensatory beta-cell hyperplastic response to insulin resistance in rodents. *Diabetes* **59**, 987–996 (2010).
330. Sun, J. *et al.* Sequencing reveals protective and pathogenic effects on development of diabetes of rare GLIS3 variants. *PLoS ONE* **14**, e0220805 (2019).
331. Balkau, B. [An epidemiologic survey from a network of French Health Examination Centres, (D.E.S.I.R.): epidemiologic data on the insulin resistance syndrome]. *Rev. Epidemiol. Sante Publique* **44**, 373–375 (1996).
332. Sladek, R. *et al.* A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* **445**, 881–885 (2007).
333. 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).

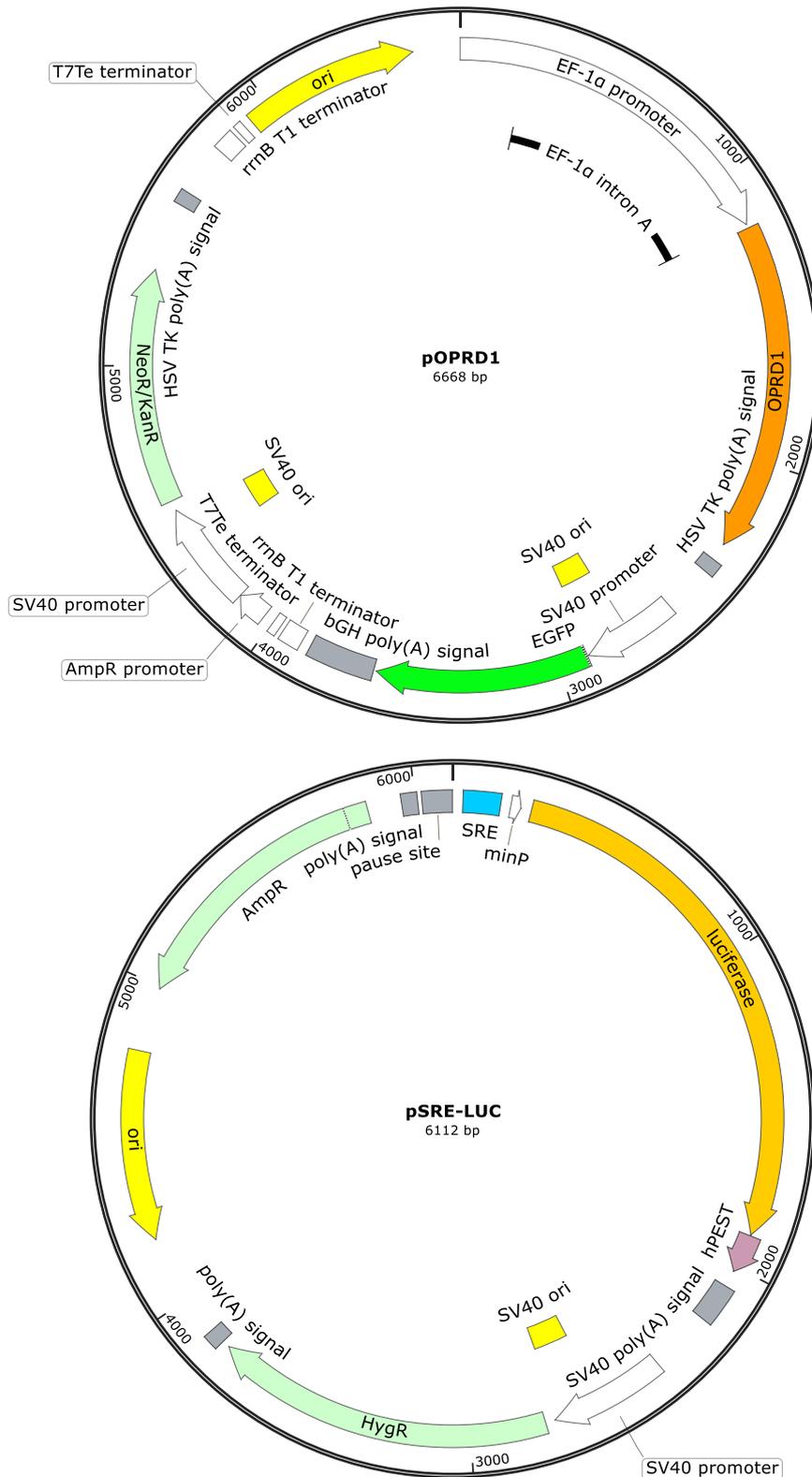
334. Romon, M. *et al.* Relationships between physical activity and plasma leptin levels in healthy children: the Fleurbaix-Laventie Ville Santé II Study. *Int. J. Obes. Relat. Metab. Disord. J. Int. Assoc. Study Obes.* **28**, 1227–1232 (2004).
335. Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–311 (2001).
336. Liu, X., Wu, C., Li, C. & Boerwinkle, E. dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Non-synonymous and Splice Site SNVs. *Hum. Mutat.* **37**, 235–241 (2016).
337. Sun, J., Zheng, Y. & Hsu, L. A unified mixed-effects model for rare-variant association in sequencing studies. *Genet. Epidemiol.* **37**, 334–344 (2013).
338. MacGregor, G. R. & Caskey, C. T. Construction of plasmids that express E. coli beta-galactosidase in mammalian cells. *Nucleic Acids Res.* **17**, 2365 (1989).
339. Marshak, S., Totary, H., Cerasi, E. & Melloul, D. Purification of the  $\beta$ -cell glucose-sensitive factor that transactivates the insulin gene differentially in normal and transformed islet cells. *Proc. Natl. Acad. Sci.* **93**, 15057–15062 (1996).
340. Stepanenko, A. A. & Dmitrenko, V. V. HEK293 in cell biology and cancer research: phenotype, karyotype, tumorigenicity, and stress-induced genome-phenotype evolution. *Gene* **569**, 182–190 (2015).
341. HEK293 Cell Line Origins, Cytogenetics, and Expression. <https://www.hek293.com/> (2014).
342. Graham, F. L., Smiley, J., Russell, W. C. & Nairn, R. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.* **36**, 59–74 (1977).
343. Thomas, P. & Smart, T. G. HEK293 cell line: a vehicle for the expression of recombinant proteins. *J. Pharmacol. Toxicol. Methods* **51**, 187–200 (2005).
344. Ravassard, P. *et al.* A genetically engineered human pancreatic  $\beta$  cell line exhibiting glucose-inducible insulin secretion. *J. Clin. Invest.* **121**, 3589–3597 (2011).
345. 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).
346. Najafipour, H. & Beik, A. The Impact of Opium Consumption on Blood Glucose, Serum Lipids and Blood Pressure, and Related Mechanisms. *Front. Physiol.* **7**, 436 (2016).
347. Blodgett, D. M. *et al.* Novel Observations From Next-Generation RNA Sequencing of Highly Purified Human Adult and Fetal Islet Cell Subsets. *Diabetes* **64**, 3172–3181 (2015).
348. Beak, J. Y., Kang, H. S., Kim, Y.-S. & Jetten, A. M. Functional analysis of the zinc finger and activation domains of Glis3 and mutant Glis3(NDH1). *Nucleic Acids Res.* **36**, 1690–1702 (2008).
349. Mahajan, A. *et al.* Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat. Genet.* **46**, 234–244 (2014).
350. Harno, E., Gali Ramamoorthy, T., Coll, A. P. & White, A. POMC: The Physiological Power of Hormone Processing. *Physiol. Rev.* **98**, 2381–2430 (2018).
351. von Dorsche, H. H., Fält, K. & Zühlke, H. Immunohistochemical investigations of beta-endorphin in human pancreatic islets. *Acta Histochem.* **85**, 131–134 (1989).

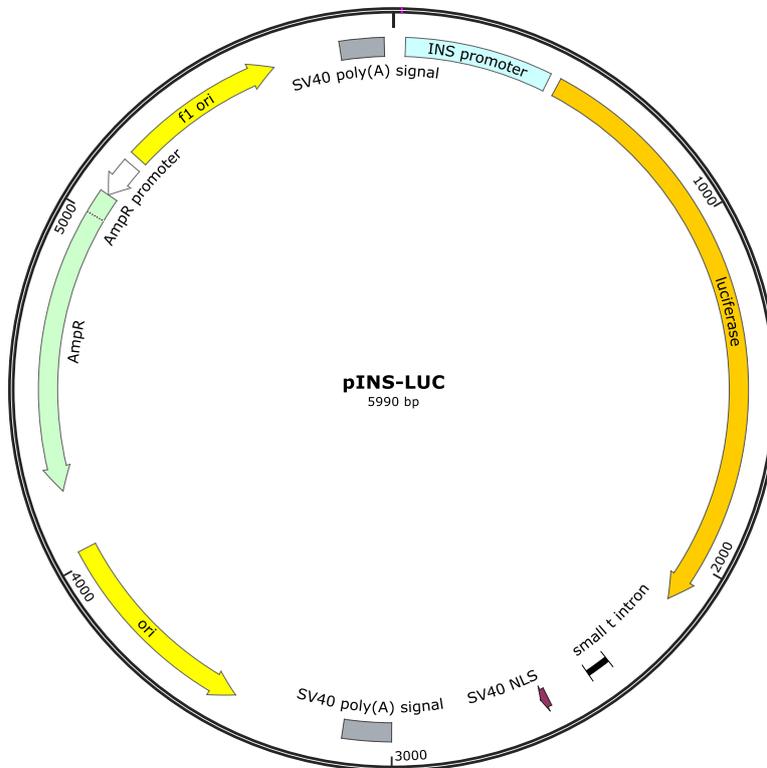
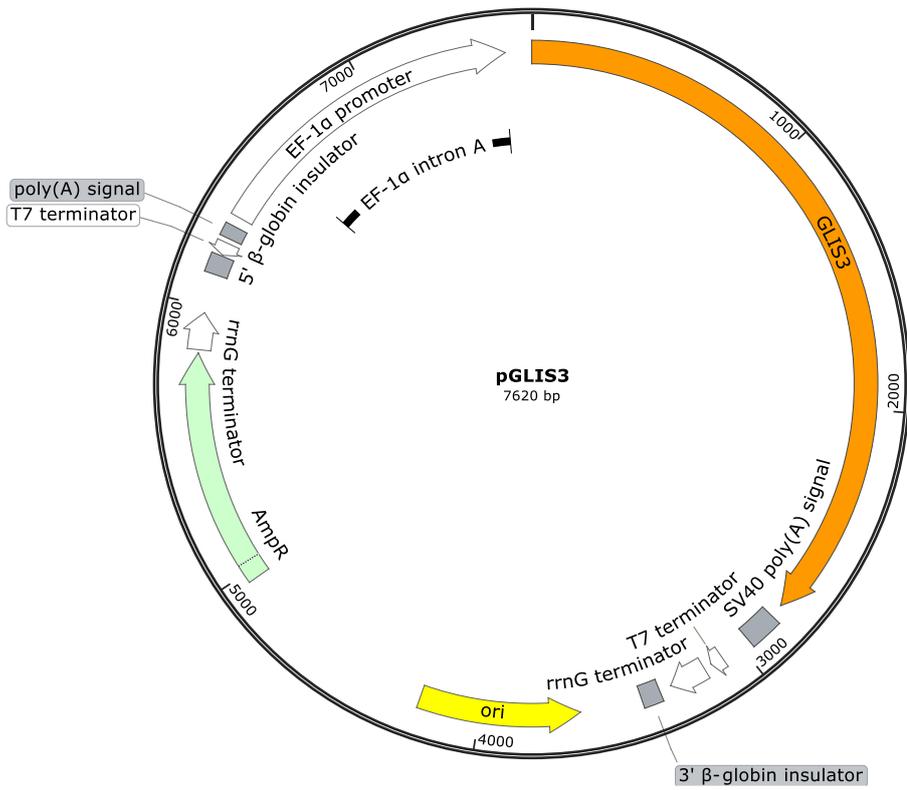
352. Froehlich, J. C. Opioid Peptides. *Alcohol Health Res. World* **21**, 132–136 (1997).
353. Evans, C. J., Erdelyi, E., Weber, E. & Barchas, J. D. Identification of Pro-Opiomelanocortin-Derived Peptides in the Human Adrenal Medulla. *Science* **221**, 957–960 (1983).
354. Giugliano, D. *et al.* Beta-endorphin and islet hormone release in humans: evidence for interference with cAMP. *Am. J. Physiol.-Endocrinol. Metab.* **257**, E361–E366 (1989).
355. Gosnell, B. A. & Levine, A. S. Reward systems and food intake: role of opioids. *Int. J. Obes.* **2005** **33 Suppl 2**, S54–58 (2009).
356. Nogueiras, R. *et al.* The opioid system and food intake: homeostatic and hedonic mechanisms. *Obes. Facts* **5**, 196–207 (2012).
357. Czyzyk, T. A. *et al.* Mice lacking  $\delta$ -opioid receptors resist the development of diet-induced obesity. *FASEB J.* **26**, 3483–3492 (2012).
358. Simonin, F. *et al.* The human delta-opioid receptor: genomic organization, cDNA cloning, functional expression, and distribution in human brain. *Mol. Pharmacol.* **46**, 1015–1021 (1994).
359. Millington, W. R. & Smith, D. L. The Posttranslational Processing of  $\beta$ -Endorphin in Human Hypothalamus. *J. Neurochem.* **57**, 775–781 (1991).
360. Sukhov, R. R., Walker, L. C., Rance, N. E., Price, D. L. & Young, W. S. Opioid precursor gene expression in the human hypothalamus. *J. Comp. Neurol.* **353**, 604–622 (1995).
361. Kubek, M. J. & Wilber, J. F. Regional distribution of leucine-enkephalin in hypothalamic and extrahypothalamic loci of the human nervous system. *Neurosci. Lett.* **18**, 155–161 (1980).
362. Hurd, Y. L. Differential messenger RNA expression of prodynorphin and proenkephalin in the human brain. *Neuroscience* **72**, 767–783 (1996).
363. Evans, A. A. L., Tunnicliffe, G., Knights, P., Bailey, C. J. & Smith, M. E. Delta opioid receptors mediate glucose uptake in skeletal muscles of lean and obese-diabetic (ob/ob) mice. *Metab. - Clin. Exp.* **50**, 1402–1408 (2001).
364. Wen, X. & Yang, Y. Emerging roles of GLIS3 in neonatal diabetes, type 1 and type 2 diabetes. *J. Mol. Endocrinol.* **58**, R73–R85 (2017).
365. Nogueira, T. C. *et al.* GLIS3, a Susceptibility Gene for Type 1 and Type 2 Diabetes, Modulates Pancreatic Beta Cell Apoptosis via Regulation of a Splice Variant of the BH3-Only Protein Bim. *PLoS Genet.* **9**, e1003532 (2013).

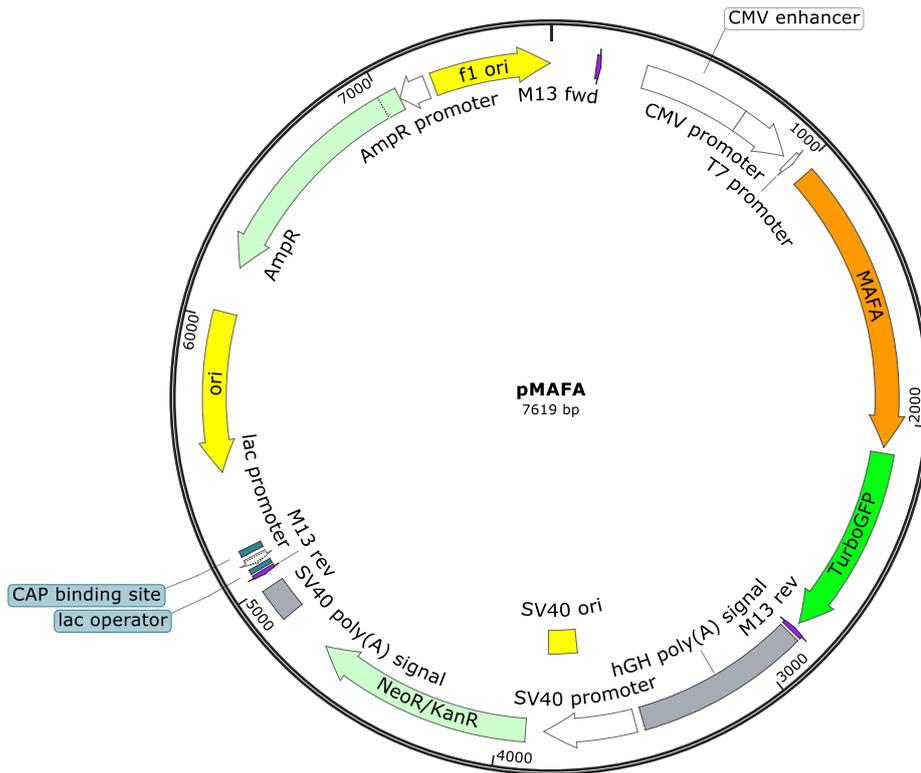
# **ANNEXES**

# Annexes

**Annexe 1.** Cartes représentant la construction des différents plasmides utilisés dans le cadre des projets *OPRD1* et *GLIS3*. Ces cartes ont été réalisées à l'aide du logiciel SnapGene (version 5.3.2).







**Annexe 2.** Liste des amorces modifiées utilisées pour la réalisation des PCR de mutagenèse du projet *OPRD1*.

Variants		Séquences des amorces forward (5'→3')	Séquence des amorces reverse (5'→3')
c.41C>G	p.P14R	GCTGCAGCCCCGGCTCTTCGCCA	TGGCGAAGAGCCGGGGCTGCAGC
c.41C>T	p.P14L	GAGCTGCAGCCCCTGCTCTTCGCCAAC	GTTGGCGAAGAGCAGGGGCTGCAGCTC
c.70C>T	p.P24S	TCGGACGCCTACTCTAGCGCCTGCC	GGCAGGCGCTAGAGTAGGCGTCCGA
c.80G>T	p.C27F	CCTACCCTAGCGCCTTCCCCAGCG	CGCTGGGGAAGGCGCTAGGGTAGG
c.95C>G	p.A32G	CAGCGCTGGCGGCAATGCGTCGG	CCGACGCATTGCCGCCAGCGCTG
c.107G>A	p.G36E	CAATGCGTCGGAGCCGCCAGGCG	CGCCTGGCGGCTCCGACGCATTG
c.126C>G	p.S42R	GCGCGCGGAGGGCCTCGTCCC	GGGACGAGGCCCTCCGCGCGC
c.154A>G	p.I52V	CCTGGCAATCGCCGTCACCGCGCTCTA	TAGAGCGCGGTGACGGCGATTGCCAGG
c.184G>C	p.V62L	CCGTGTGCGCCCTGGGGCTGCTG	CAGCAGCCCCAGGGCGCACACGG
c.243G>T	p.K81N	CGTCCGGTACACTAAGATGAATACGGCCACCAA	TTGGTGGCCGTATTCATCTTAGTGTACCGGACG
c.283G>A	p.D95N	CCTGGCCTTAGCCAATGCGCTGGCCAC	GTGGCCAGCGCATTGGCTAAGGCCAGG
c.287C>T	p.A96V	GCCTTAGCCGATGTGCTGGCCACCAGC	GCTGGTGGCCAGCACATCGGCTAAGGC
c.314A>T	p.Q105L	CACGCTGCCTTCTCTGAGTGCCAAGTACC	GGTACTTGGCACTCAGGAAAGGCAGCGTG
c.391A>G	p.N131D	CTCTCCATCGACTACTACGATATGTTACCAGCATCT	AGATGCTGGTGAACATATCGTAGTAGTCGATGGAGAG
c.396G>C	p.M132I	GCTCTCCATCGACTACTACAATATCTTACCAGCATCT	AGATGCTGGTGAAGATATTGTAGTAGTCGATGGAGAGC
c.424A>T	p.M142L	CTTCACGCTCACCATGTTGAGTGTTGACCGCTA	TAGCGGTCAACACTCAACATGGTGAGCGTGAAG
c.478C>G	p.R160G	CCCTGGACTTCGGCACGCCTGCC	GGCAGGCGTGCCGAAGTCCAGGG
c.547A>G	p.I183V	CGTTGGCGTGCCCGTCATGGTCATGGC	GCCATGACCATGACGGGCACGCCAACG
c.575G>A	p.R192Q	GACCCGTCCCCAGGACGGGGCAG	CTGCCCCGTCTGGGGACGGGTC
c.658T>C	p.F220L	Plasmide commandé chez e-Zyvec	
c.661G>T	p.A221S	GCGTGTTCTCTTCTCCTTCGTGGTGCCC	GGGCACCACGAAGGAGAAGAGGAACACGC
c.709C>A	p.L237M	CTATGGCCTCATGATGCTGCGCCTGCG	CGCAGGCGCAGCATCATGAGGCCATAG
c.725G>A	p.S242N	CTGCGCCTGCGCAATGTGCGCCTGC	GCAGGCGCACATTGCGCAGGCGCAG
c.743G>T	p.G248V	CGCCTGCTGTCGGTCTCCAAGGAGAAG	CTTCTCCTTGAGACCGACAGCAGGCG
c.749A>G	p.K250R	GCTGTGCGGGCTCCAGGGAGAAGGACCG	CGGTCCTTCTCCTGGAGCCCCGACAGC
c.781C>T	p.R261C	CGGCGCATCACGTGCATGGTGCTGG	CCAGCACCATGCACGTGATGCGCCG

c.840C>G	p.F280L	Plasmide commandé chez e-Zyvec	
c.925G>A	p.A309T	CGCGCTGGGCTACACCAATAGCAGCCT	AGGCTGCTATTGGTGTAGCCCAGCGCG
c.966C>G	p.D322E	CTACGCTTTCCTCGAGGAGAACTTCAAGCGC	GCGCTTGAAGTTCTCCTCGAGGAAAGCGTAG
c.989G>T	p.R330L	AGCGCTGCTTCCTCCAGCTCTGCCG	CGGCAGAGCTGGAGGAAGCAGCGCT
c.1063G>T	p.E355*	CCACGGCCCGCTAGCGTGTACC	GGTGACACGCTAGCGGGCCGTGG
c.1082C>A	p.T361N	GTGTCACCGCCTGCAACCCGTCGATG	CATCGGACGGGTTGCAGGCGGTGACAC
c.1094G>A	p.G365D	Plasmide commandé chez e-Zyvec	
c.1097C>T	p.P366L	GTCCGATGGTCTCGGCGGTGGCG	CGCCACCGCCGAGACCATCGGAC

**Annexe 3.** Liste des amorces modifiées utilisées pour la réalisation des PCR de mutagenèse du projet *GLIS3*.

Variants		Séquence des amorces forward (5'-3')	Séquence des amorces reverse (5'->3')
c.34C>T	p.R12W	GCATGAGTCTCCACTGGACATCGGGAACC	GGTTCCTCGATGTCCAGTGGAGACTCATGC
c.64A>G	p.M22V	CACAGGGGCCTAGGGTGGTCAGTGGTCAT	ATGACCACTGACCACCCTAGGCCCTGTG
c.82A>G	p.I28V	GTCAGTGGTCATCACGTTCTGCCATCCGAG	CTCGGATGGCAGGAACGTGATGACCACTGAC
c.86C>T	p.P29L	GTGGTCATCACATTCTTGCCATCCGAGCCC	GGGCTCGGATGGCAAGAATGTGATGACCAC
c.106G>A	p.G36R	CGAGCCCACTCCAGGACTCCTGGCC	GGCCAGGAGTCCTGGAGTGGGCTCG
c.113C>G	p.P38R	CTCCGGGACTCGTGGCCCCCTCGC	GCGAGGGGGCCACGAGTCCCGGAG
c.202A>G	p.M68V	CCTCAGGAGGAGGGGTGGCTCCTCAGAAC	GTTCTGAGGAGCCACCCCTCCTCTGAGG
c.232C>G	p.R78G	CGTGGCTGAGAGCGGCATCCATCTGCC	GGCAGATGGATGCCGCTCTCAGCCACG
c.233G>A	p.R78H	CGTGGCTGAGAGCCACATCCATCTGCCTG	CAGGCAGATGGATGTGGCTCTCAGCCACG
c.252A>T	p.L84F	CATCCATCTGCCTGCCTTTAGCCCCAGGA	TCCTGGGGCTAAAGGCAGGCAGATGGATG
c.256C>T	p.P86S	CTGCCTGCCTTAAGCTCCAGGAGACAAATGC	GCATTTGTCTCCTGGAGCTTAAGGCAGGCAG
c.275C>T	p.T92I	GGAGACAAATGCTCATCAATGGGAAGCCGCG	CGCGGCTTCCCATTGATGAGCATTTGTCTCC
c.287C>T	p.P96L	TCACCAATGGGAAGCTGCGATTCCAGGTCAC	GTGACCTGGAATCGCAGCTTCCCATTGGTGA
c.358G>C	p.G120R	AGCCAAAGCAGCAGGAGTTTGAAGCCCTTTTC	GAAAAGGGCTTCGAAACTCCTGCTGCTTTGGCT
c.370C>T	p.P124S	AGGAGTTTGAAGCCCTTTTCTCCAAATCCTGGG	CCCAGGATTTGGAGAAAAGGGCTTCCAAACTCCT
c.413G>A	p.C138Y	TGGCTTTGGGCTCAGTACAAGTCCATTGGAAAAG	CTTTTCCAATGGACTTGTACTGAGGCCAAAGCCA
c.422T>C	p.I141T	GGCCTCAGTGCAAGTCCACTGGAAAAGGCAG	CTGCCTTTCCAGTGGACTTGCCTGAGGCC
c.425G>A	p.G142E	TCAGTGCAAGTCCATTGAAAAAGGCAGCTGCAACA	TGTTGCAGCTGCCTTTTTCAATGGACTTGCCTGA
c.467T>G	p.M156R	CACCAGCAGTCCCAGGATGGTTCAGCGAC	GTCGCTGAACCATCCTGGGACTGCTGGTG
c.472G>T	p.V158F	AGCAGTCCCATGATGTTTCAGCGACTGGGAC	GTCCCAGTCGCTGAAACATCATGGGACTGCT
c.496C>G	p.P166A	CTGGGACTCATTTCAGCTCCAGCAAGCCAG	CTGGCTTGCTGGAGCTGAAATGAGTCCCAG
c.503C>G	p.A168G	CTCATTTACCTCCAGGAAGCCAGGTCTCTACA	TGTAGAGACCTGGCTTCTGGAGGTGAAATGA
c.515C>T	p.S172F	CCAGCAAGCCAGGTCTTTACAGCATGCAACCAG	CTGGTTGCATGCTGTAAAGACCTGGCTTGTCTGG
c.713C>G	p.S238C	CTCCCTTCGCTCTGCAACCACGGCTCT	AGAGCCGTGGTTGCAGAGCGAAGGGAG
c.758G>C	p.S253T	CTAGGGGATCTCCTTACCCTTCTCCCGG	CCGGGAGGAAGGGTAAGGAGATCCCCTAG
c.793G>T	p.V265F	GACATCCATGTCCAGCAATAGTTTCTCTAACTCATTACC	GGTAATGAGTTAGAGAAACTATTGCTGGACATGGATGTC
c.814T>A	p.Y272N	CTCTAACTCATTACCATCCAACCTTTTGGCACGGAAAAG	CTTTCCGTGCCAAAAGGTTGGATGGTAATGAGTTAGAG
c.827C>T	p.T276M	CCATCCTACCTTTTTGGCATGGAAAGTAGCCACTC	GAGTGGCTACTTTCCATGCCAAAAGGTAGGATGG
c.833G>A	p.S278N	CTACCTTTTGGCACGGAAAATAGCCACTCTCCTT	AAGGAGAGTGGCTATTTCCGTGCCAAAAGGTAG
c.842C>T	p.S281F	TTGGCACGGAAAAGTAGCCACTTTCCTTACCCTAG	CTAGGGTAAGGAAAAGTGGCTACTTTCCGTGCCAA

c.844C>G	p.P282A	GGAAAGTAGCCACTCTGCTTACCCTAGTCCTCG	CGAGGACTAGGGTAAGCAGAGTGGCTACTTTCC
c.890G>A	p.R297H	TCCCACTCGGCCACTCCAAGAAGAGAG	CTCTCTTCTGGAGTGGGCCGAGTGGGA
c.893C>A	p.S298H	CACTCGGCCCGCTACAAGAAGAGAGCG	CGCTCTCTTCTGTAGCGGGCCGAGTG
c.937G>A	p.G313R	CGCTGTCCGATGGCATCAGGATAGATTTCAATACC	GGTATTGAAATCTATCCTGATGCCATCGGACAGCG
c.968C>T	p.S323L	TACCATCATCCGCACGTTGCCACGTCC	GGACGTGGGCAACGTGCGGATGATGGTA
c.974C>T	p.T325M	GCACGTCGCCCATGTCCTTGGTGGC	GCCACCAAGGACATGGGCGACGTGC
c.1012C>T	p.P338S	GTCGAGGGCTTCGTCGGCCAACCTGTC	GACAGGTTGGCCGACGAAGCCCTCGAC
c.1063C>T	p.R355C	TCCTGGGCGTGTGCGGCAGCTGC	GCAGCTGCCGCACACGCCAGGA
c.1073G>A	p.C358Y	TGCGCGGCAGCTACATCCCCAGCC	GGCTGGGGAATGTAGCTGCCGCGCA
c.1085C>A	p.P362Q	CATTCCCAGCAGCGCCCGGTGC	GCACCGGGCGCTGCTGGGGAATG
c.1090C>T	p.P364S	CCAGCCGCGCTCGGTGCCCGG	CCGGGCACCGAGCGCGGCTGG
c.1111G>T	p.G371C	GGCAGCCAGAAGTGCCTGCTGGTGG	CCACCAGCACGCACTTCTGGCTGCC
c.1117C>G	p.L373V	CAGAAGGGCGTGGTGGTGCCCTG	CAGGGGCCACCACCACGCCCTTCTG
c.1126C>T	p.P376S		Plasmide commandé chez Twist
c.1145C>A	p.P382Q	CCTGGCGCTGCAGGCCTACGGCG	CGCCGTAGGCCTGCAGCGCCAGG
c.1154G>T	p.G385V	GCCGGCCTACGTCGAGGACGGGG	CCCCGTCTCGACGTAGGCCGGC
c.1185G>A	p.M395I	GAGCACGAGCGCATACAACAGCTGGAGCA	TGCTCCAGCTGTTGTATGCGCTCGTGCTC
c.1191G>C	p.Q397H	GCGCATGCAACACCTGGAGCACGGC	GCCGTGCTCCAGGTGTTGCATGCGC
c.1199A>G	p.H400R	ACAGCTGGAGCGCGGGCCTGC	GCAGGCCGCCGCGCTCCAGCTGT
c.1200C>G	p.H400Q	GCAACAGCTGGAGCAGGGCGGCCT	AGGCCGCCCTGCTCCAGCTGTTGC
c.1210C>A	p.Q404K	ACGGCGGCCTGAAGCCAGGCCTG	CAGGCCTGGCTTCAGGCCCGCT
c.1214C>T	p.P405L	GCGGCCTGAAGCCAGGCCTGGTCAA	TTGACCAGGCCTAGCTGCAGGCCGC
c.1216G>T	p.G406C	GCGGCCTGCAGCCATGCCTGGTC	GACCAGGCATGGCTGCAGGCCGC
c.1227C>A	p.N409K	GCCAGGCCTGGTCAAACACATGGTGGTG	CACCACCATGTGTTTGACCAGGCCTGGC
c.1232T>A	p.M411K	CCAGGCCTGGTCAACCACAAGGTGGTGCA	CTGCACCACCTTGTGGTTGACCAGGCCTGG
c.1271C>G	p.S424W	GCCCCGACAGCCAGTGGGCCGGCC	GGCCGGCCCACTGGCTGTCGGGGC
c.1295G>A	p.R432H	TGTTCAAGACCGAACACCTGGAGGAGTTCCC	GGGAACTCCTCCAGGTGTTGCGTCTTGAACA
c.1318A>G	p.T440A	GTTCCCGGGCAGCGCCGTAGACCTACC	GGTAGGTCTACGGCGCTGCCCGGGAAC
c.1325A>C	p.D442A	GCAGCACCGTAGCCCTACCCCCGC	GCGGGGGGTAGGGCTACGGTGCTGC
c.1336G>A	p.A446T	ACCTACCCCCACGCCTCCGCTC	GAGCGGAGGCGTGGGGGTAGGT
c.1388A>C	p.H463P	CCACCCCTTACCCTGCCATGCGCAC	GTGCGCATGGGCAGGGTAAGGGGGTGG
c.1417C>T	p.L473F	ACCACCCGGAGTTCCGGGCCCCAC	GTGGGGCCCAACTCCGGGTGGT
c.1430C>T	p.A477V	CGGGCCCCACGTCCAGCAGCTGG	CCAGCTGCTGGACGTGGGGCCCG

c.1480G>A	p.D494N	GACGGGGAGATGAACGGCATCGGGG	CCCCGATGCCGTTTCATCTCCCCGTC
c.1483G>A	p.G495S	GGGGAGATGGACAGCATCGGGGGCA	TGCCCCGATGCTGTCCATCTCCCC
c.1518C>G	p.D506E	CTGCCGCTGGATCGAGTGCAGCGC	GCGCTGCACTCGATCCAGCGGCAG
c.1545G>C	p.E515D	CGACCAGCAGGACGAGCTCGTGCGG	CCGCACGAGCTCGTCTGTGGTTCG
c.1709C>T	p.T570M	GAAGCCCAACAAGTGTATGTTTGAAGGTTGCGAGA	TCTCGCAACCTTCAAACATACACTTGTGGGCTTC
c.1715A>T	p.E572V	CAACAAGTGTACGTTTGTAGGTTGCGAGAAGGCCT	AGGCCTTCTCGCAACCTACAAACGTACACTTGTG
c.1775C>G	p.T592R	CCACTTGCGGAGCCACAGAGGCGAGAA	TTCTCGCTCTGTGGCTCCGCAAGTGG
c.1843G>A	p.A615T	CTCCAGTGACCGCACCAAACACCAGCG	CGCTGGTGTGGTGGTGGTCACTGGAG
c.1881T>A	p.Y627*	GACGCATCTGGACACCAAACCTTAAGCTTGTCAAATTC	GAATTTGACAAGCTTAAGTTTTGGTGTCCAGATGCGTC
c.1994G>A	p.S665N	GAAAAAGTTGCGGTCCAACACAGAGCTCCATCCAG	CTGGATGGAGCTCTGTGTTGGACCGCAACTTTTTTC
c.2051C>T	p.P684L	Plasmide commandé chez Twist	
c.2060C>T	p.S687F	CTGCAGCCGGCCACTTTCCCTAGAGATGC	GCATCTCTAGGGAAAGTGGCCGGCTGCAG
c.2067A>T	p.R689S	GCCACTTCCCCTAGTGATGCTGCTGCTGA	TCAGCAGCAGCATCACTAGGGGAAGTGGC
c.2089G>A	p.V69M	GCTGAAGGGACCATGGGACGCTCCC	GGGAGCGTCCCATGGTCCCTTCAGC
c.2096G>A	p.R699H	GGACCGTGGGACACTCCCCTGGACC	GGTCCAGGGGAGTGTCCACGGTCC
c.2107C>T	p.P703S	GGACGCTCCCCTGGATCCGGGCT	AGGCCCGGATCCAGGGGAGCGTCC
c.2141C>G	p.S714C	CTATTCAGTCCCATTTTCTGCAGCAATTATTCAAGCCG	CGGCTTGAATAATTGCTGCAGAAAATGGGAGCTGAATAG
c.2281G>A	p.D761N	CACTCACAGCTGTGAACGCAGGAGCTGAG	CTCAGCTCCTGCGTTCACAGCTGTGAGTG
c.2305C>A	p.P769T	AGCTGAGAGGTTTGCAACTTCTGCTCCATCTCC	GGAGATGGAGCAGAAGTTGCAAACCTCTCAGCT
c.2312C>A	p.A771D	AGGTTTGCACCTTCTGATCCATCTCCTCACCAC	GTGGTGAGGAGATGGATCAGAAGGTGCAAACCT
c.2404G>A	p.G802S	Plasmide commandé chez Twist	
c.2431G>C	p.E811Q	CCTGAAGTCTATCAGCCACAAACAACTCTTCTTTTCAA	TTGAAAAGAAGAGTTTGTGGTGGCTGATAGGACTTCAGG
c.2439C>G	p.N813K	CCTGAAGTCTATCAGCCAGAAACAAAGTCTTCTTTTCAACC	GGTTGAAAAGAAGACTTTGTTTCTGGCTGATAGGACTTCAGG
c.2461A>G	p.I821V	ACAAACTCTTCTTTTCAACCAAATGGTGTCCATGTCCATGGA	TCCATGGACATGGACACCATTTGGTTGAAAAGAAGAGTTTGT
c.2494A>G	p.K832E	TGGGCAGCTGCAGGAGTTCTGTCCCCC	GGGGGACAGAACTCCTGCAGCTGCCCA
c.2518G>A	p.D840N	CCCCCACACTACCCCAATTCCAGAGAATTG	CAATTCTCTGGGAATTGGGGTAGTGTGGGGG
c.2531T>C	p.I844T	CCGATTCCCAGAGAACTGTGCCGCTGTC	GACAGGCGGCACAGTTCTCTGGGAATCGG
c.2549C>G	p.S850C	CCGCCTGTCAGCTGCTGCAGTGTGGTG	CACCACACTGCAGCAGCTGACAGGCGG
c.2609G>A	p.S870N	TCCATGGGCCAGGCCAATTTTGTATGTTTTCCACA	TGTGGAAAACATCAAATTTGGCCTGGCCCATGGA
c.2614G>C	p.D872H	CATGGGCCAGGCCAGTTTTTCTATGTTTTCCACAGA	TCTGTGGAAAACATGAAAATTTGGCCTGGCCCATG
c.2623C>T	p.H875Y	GGCCAGTTTTGATGTTTTCTACAGAGCCTTCTCGA	TCGAGAAGGCTCTGTAGAAAACATCAAACCTGGCC
c.2636C>G	p.S879W	CACAGAGCCTTCTTGACTCACTCGGGC	GCCCCGAGTGAGTCAAGAAGGCTCTGTG
c.2636C>T	p.S879L	Plasmide commandé chez Twist	

c.2648G>C	p.G883A	CTTCTCGACTCACTCGGCCATTACAGTGTATGATTT	AAATCATACACTGTAATGGCCGAGTGAGTCGAGAAG
c.2686C>T	p.L896F	CTTCAAGTTCCTCGAGCTTCTTTGGGGAGTCTCTC	GAGAGACTCCCCAAAGAAGCTCGAGGAACTTGAAG
c.2705G>A	p.R902H	GGGAGTCTCTCCACAGCGGGGCTGA	TCAGCCCCGCTGTGGAGAGACTCCC
c.2710G>A	p.G904R	GTCTCTCCGCAGCAGGGCTGAAGATGC	GCATCTTCAGCCCTGCTGCGGAGAGAC
c.2741G>A	p.S914N	CCTTCTTG CAGATCAACACCGTGGACCGCTG	CAGCGGTCCACGGTGTGATCTGCAAGAAGG
c.2746G>A	p.V91M	GCAGATCAGCACCATGGACCGCTGTCC	GGACAGCGGTCCATGGTGCTGATCTGC
c.2747_2749del	p.V916del	GCAGATCAGCACCGACCGCTGTCCTAGC	GCTAGGACAGCGGTTCGGTGCTGATCTGC
c.2749G>A	p.D917N	AGATCAGCACCGTGAACCGCTGTCCTAGC	GCTAGGACAGCGGTTACGGTGCTGATCT
c.2770T>G	p.S924A	GTCCTAGCCAGCTCGCCTCTGTCTACACC	GGTGTAGACAGAGGCGAGCTGGCTAGGAC
c.2785G>A	p.E929K	Plasmide commandé chez Twist	

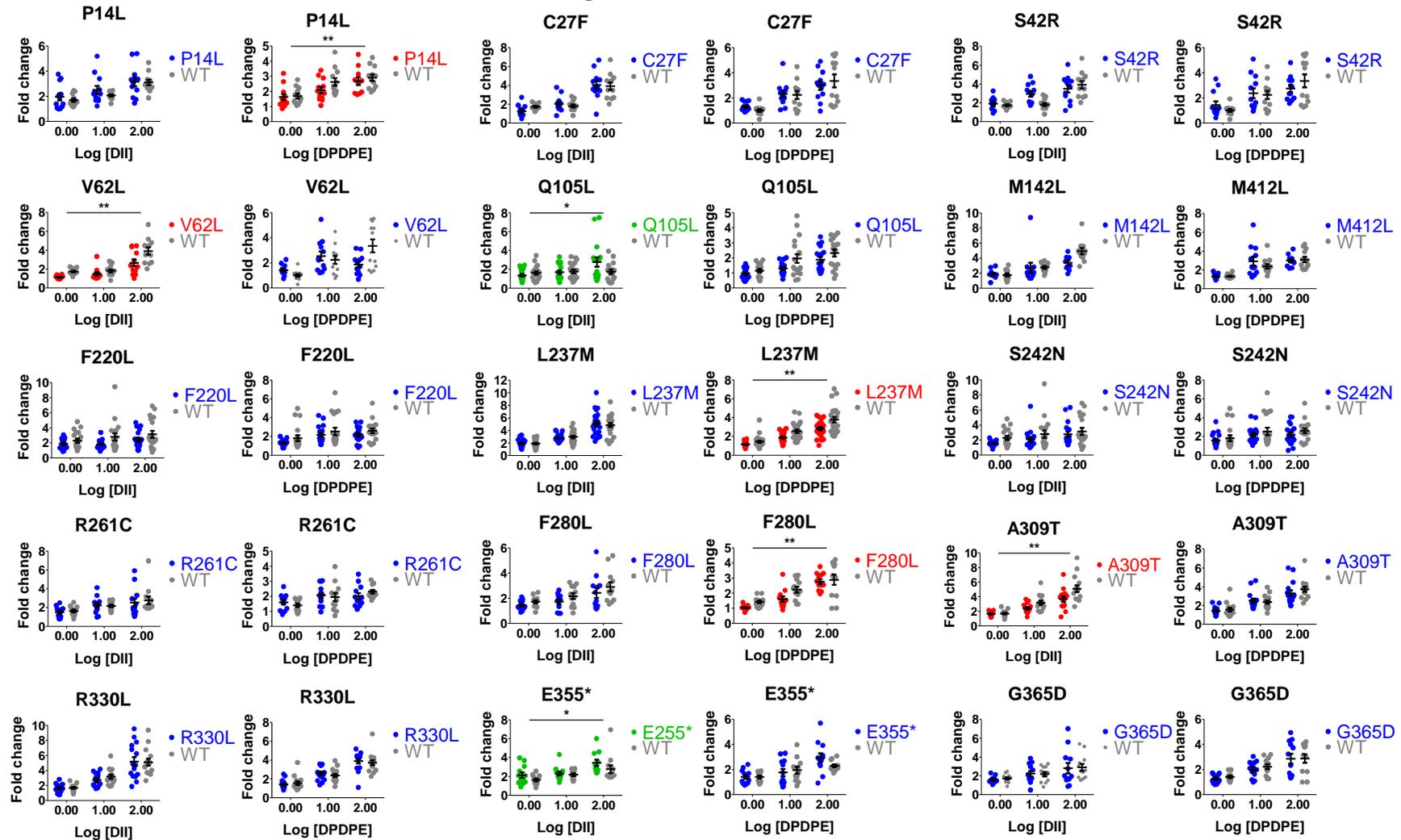
Annexe 4. Liste des amorces utilisées pour le séquençage des plasmides pOPRD1.

Région du plasmide ciblée	Sens de l'amorce	Séquence de l'amorce (5'→3')
pEF1α (en amont du gène <i>OPRD1</i> )	Forward	TCAAGCCTCAGACAGTGGTTC
	Reverse	CCTCGGACGCCTACCCTA
Gène <i>OPRD1</i>	Forward	GATGCCGAACATGACAAGC
	Reverse	CTAAGATGAAGACGGCCACCA
	Forward	CTGAGGCCAGGACCCAGATA
	Reverse	GGTGTGCATGCTCCAGTTC
	Forward	TCCAGACGATGACGAAGATG
	Reverse	GTCTGGACGCTGGTGGAC
	Forward	CTGAAGCTGCTGGGGTCT
	Reverse	CTGAAGCTGCTGGGGTCT
tHSV (en aval du gène <i>OPRD1</i> )	Reverse	GGTGGGGTATCGACAGAGTG

Annexe 5. Liste des amorces utilisées pour le séquençage des plasmides pGLIS3.

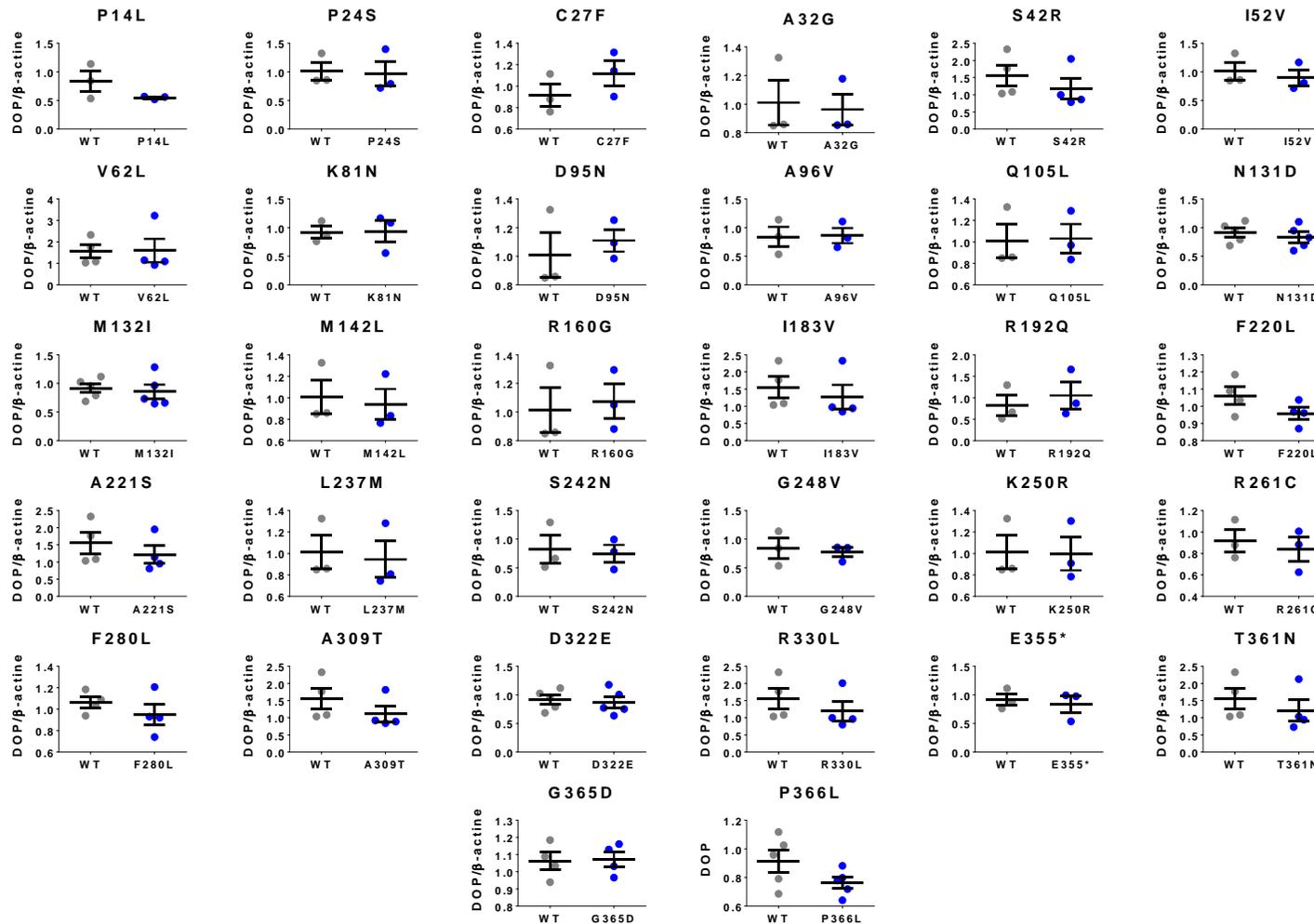
Région du plasmide ciblée	Sens de l'amorce	Séquence de l'amorce (5'→3')
Gène <i>GLIS3</i>	Forward	TAAGCCCCAGGAGACAAATG
	Reverse	CATTCTGAGAGCCGTGGTTG
	Forward	GAAAGTCAGTCGGCCTCAAG
	Reverse	CATGCTGCACCACCATGT
	Forward	GAGGTCTACGGGCATTTCT
	Reverse	GCCCAGAAGCAAGTGAAGTC
	Forward	GCATCTGGACACCAAACCTT
	Reverse	GAAGGTGCAAACCTCTCAGC
	Forward	ATTCAAGCCGAAGTGAACA
	Reverse	TGTTCCACTTCGGCTTGAAT
	Reverse	CGATTCCCAGAGAATTGTGC

Annexe 6. Activité luciférase des variants neutres du gène *OPRD1*.



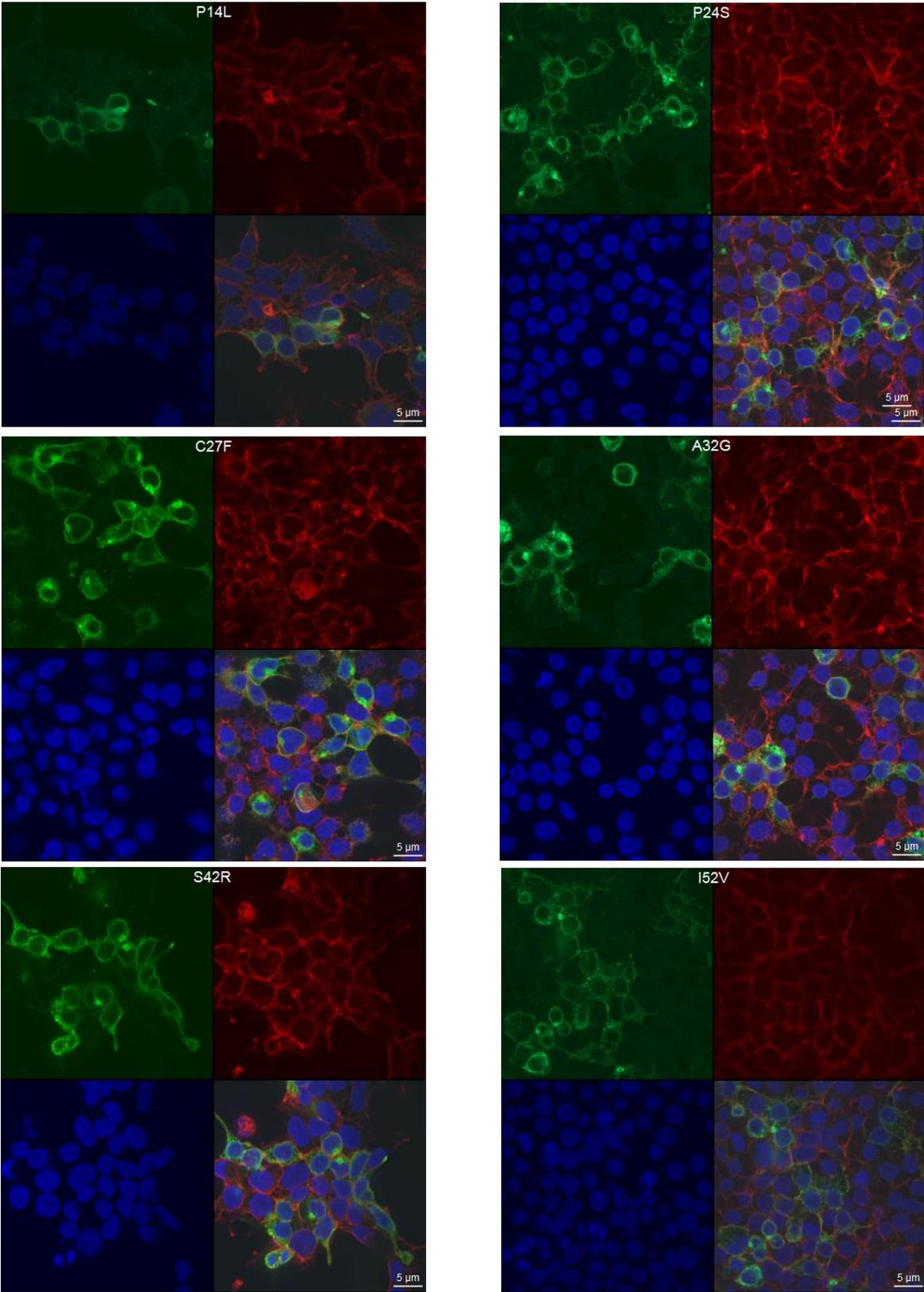
Activité luciférase des variants neutres du gène *OPRD1*, en réponse aux agonistes de DOP DII et DPDPE. Ces graphiques représentent, pour chaque variant, le fold change de l'activité de la SRE-luciférase (normalisée par les valeurs de  $\beta$ -galactosidase) en réponse à des concentrations en DII et en DPDPE de 1, 10 et 100 nM, par rapport à l'activité basale du récepteur DOP WT. L'effet de chaque variant a été analysé selon le modèle ANOVA suivant :  $FC = \beta_0 + \beta_1C + \delta_1V + \delta_2C \times V + \beta_2E + \epsilon$ , avec C correspondant à la concentration de l'agoniste, V correspondant au variant étudié, et E correspondant au numéro d'identification de l'expérience. Les graphiques représentent la moyenne  $\pm$  SEM des fold change d'au moins 4 expériences indépendantes. \* $P \leq 0,05$ , \*\* $P \leq 0,01$ , \*\*\* $P \leq 0,001$ . DII : [D-Ala<sup>2</sup>]-Deltorphin II, DPDPE : [D-Pen<sup>2</sup>,D-Pen<sup>5</sup>]-Enkephalin, WT : wild-type.

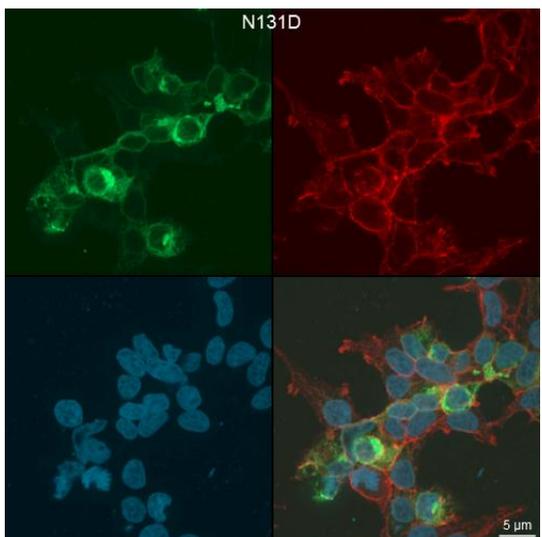
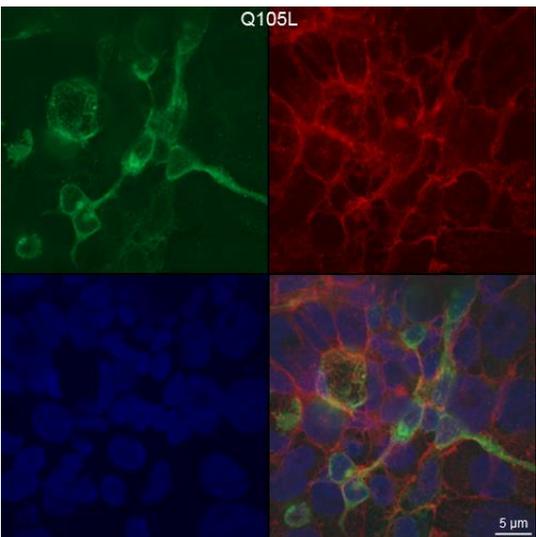
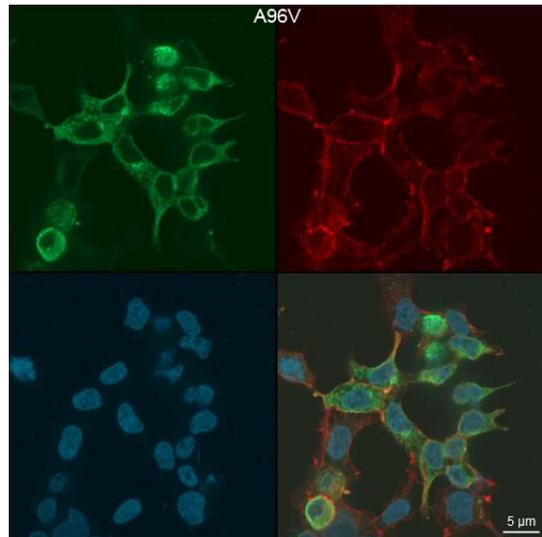
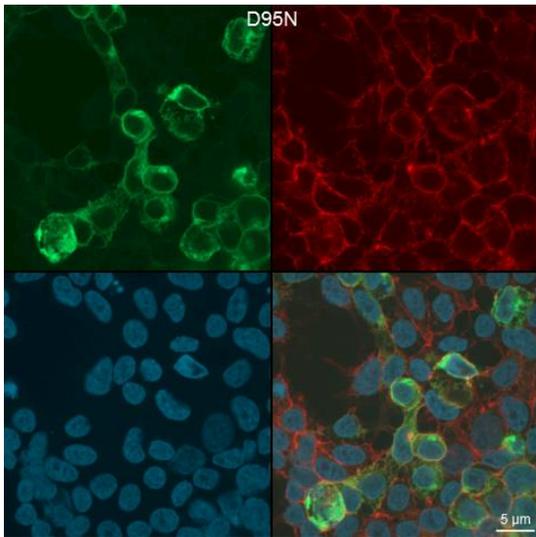
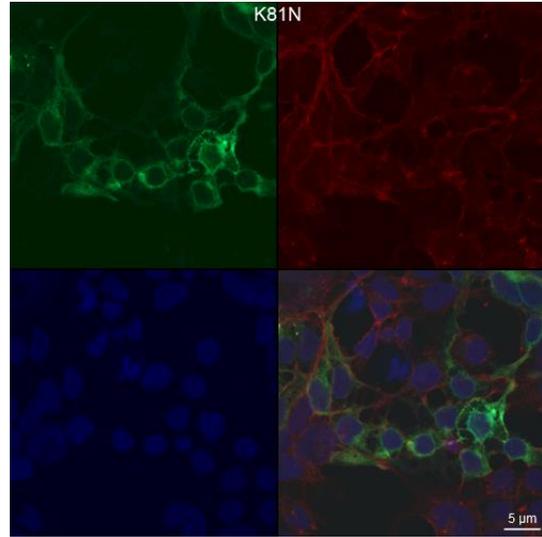
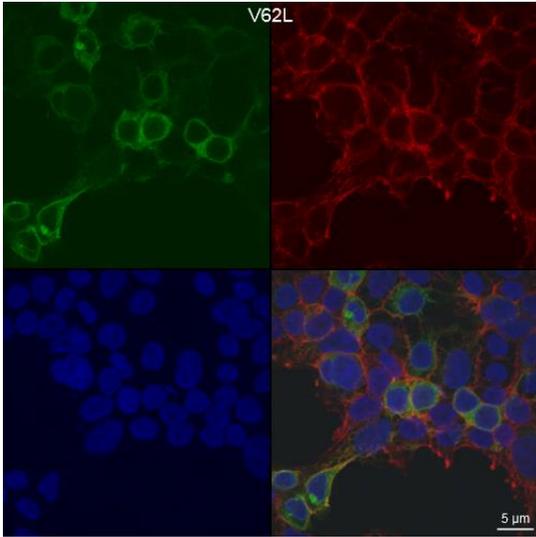
## Annexe 7. Expression des différents mutants du récepteur DOP.

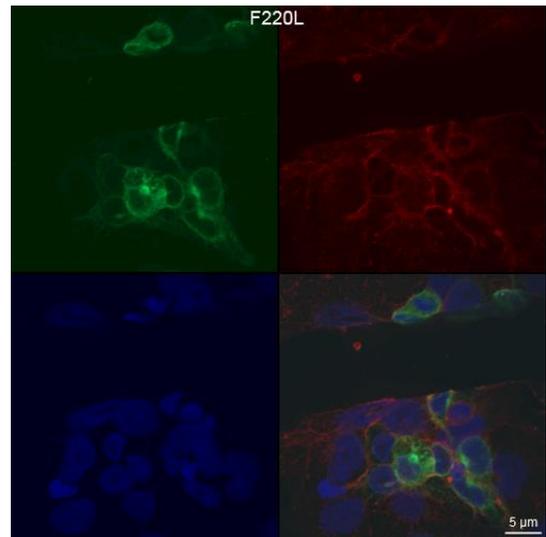
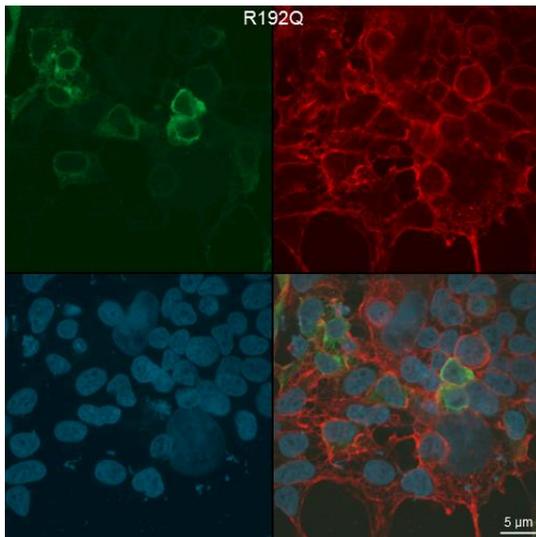
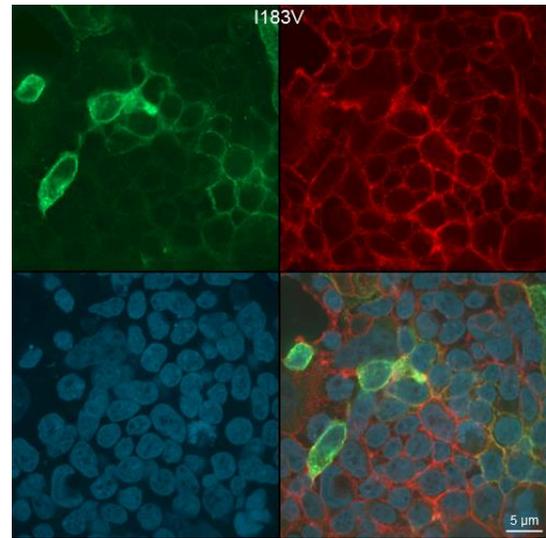
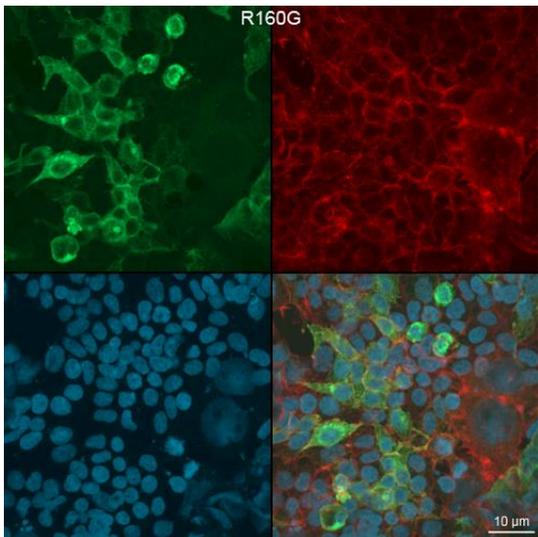
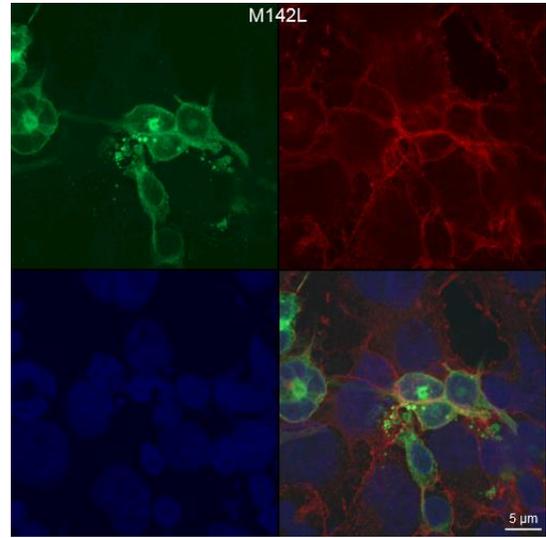
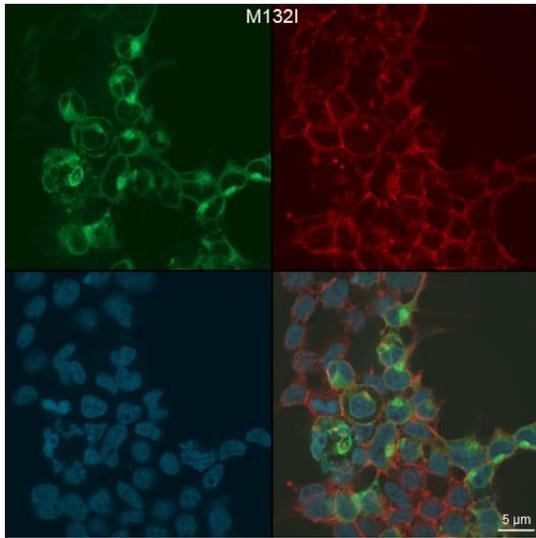


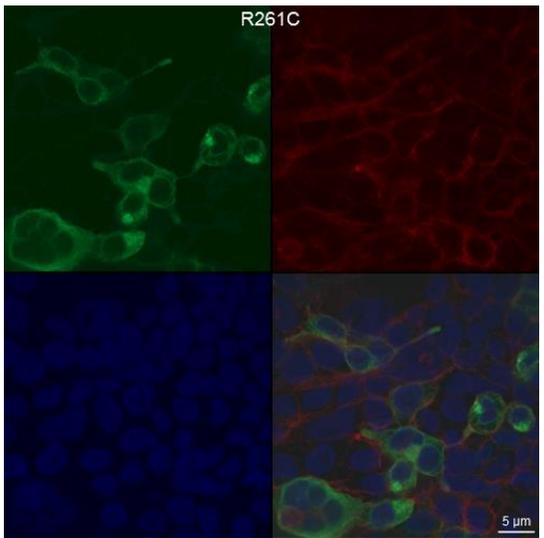
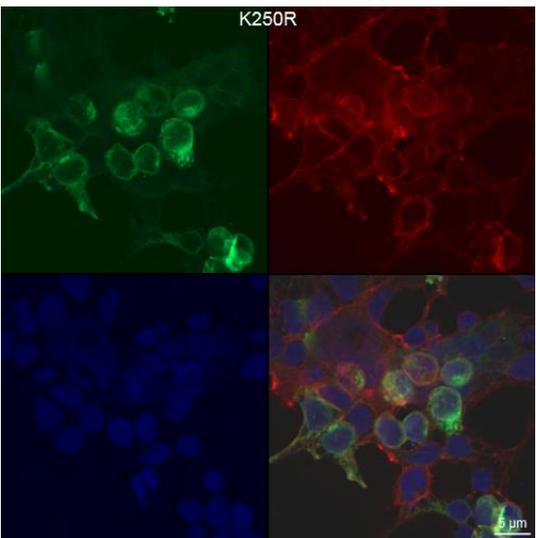
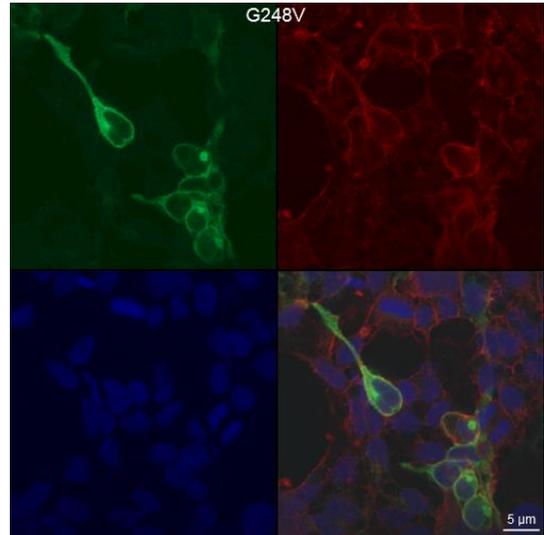
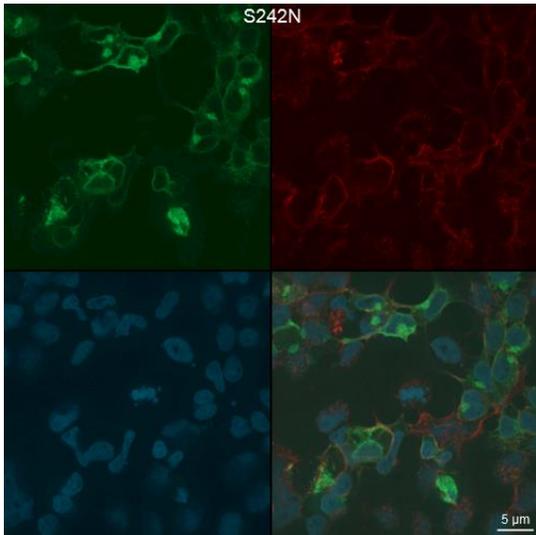
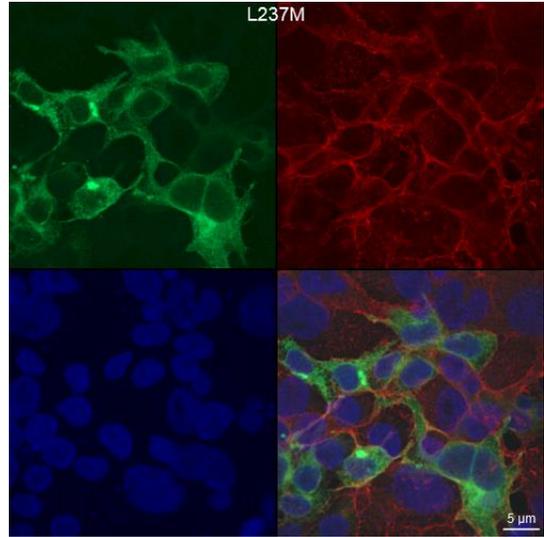
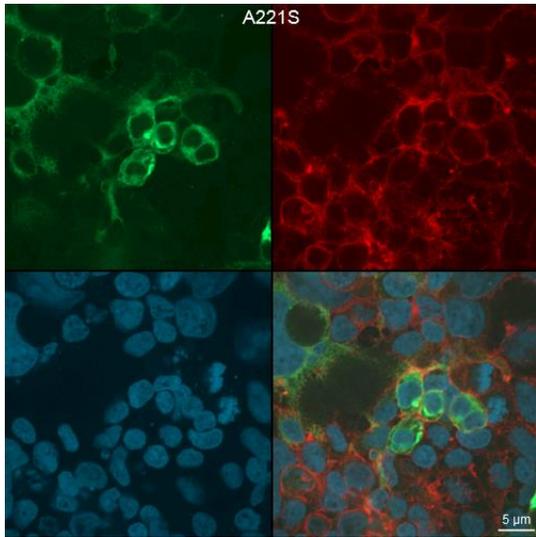
Expression des différents mutants du récepteur DOP mesurée par western blot. Pour chaque condition, l'intensité de la bande correspondant au récepteur DOP a été normalisée par celle de la  $\beta$ -actine associée, puis comparée à l'expression du récepteur DOP en condition WT par un test de la somme des rangs de Wilcoxon. Ces graphiques représentent la moyenne de l'expression normalisée du récepteur DOP  $\pm$  SEM d'au moins 3 expériences indépendantes. **WT** : wild-type.

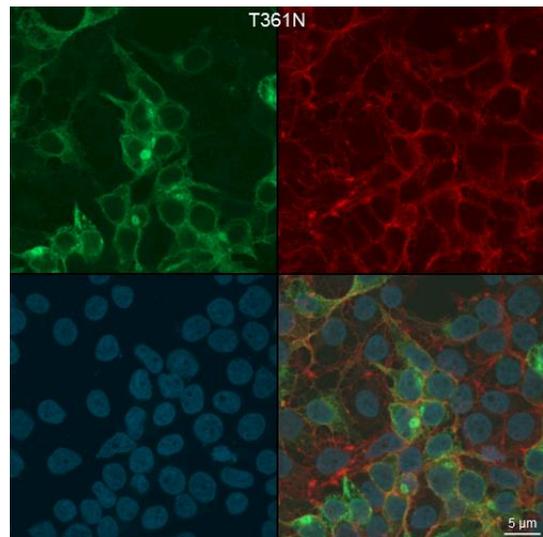
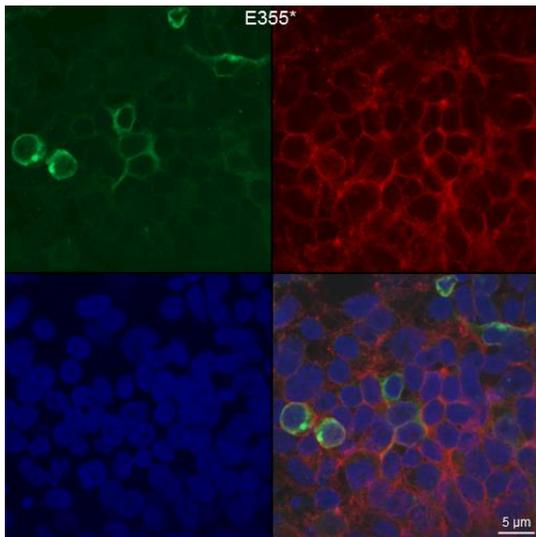
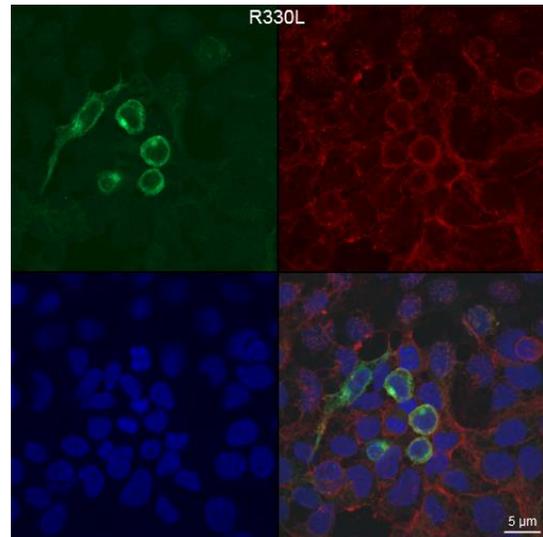
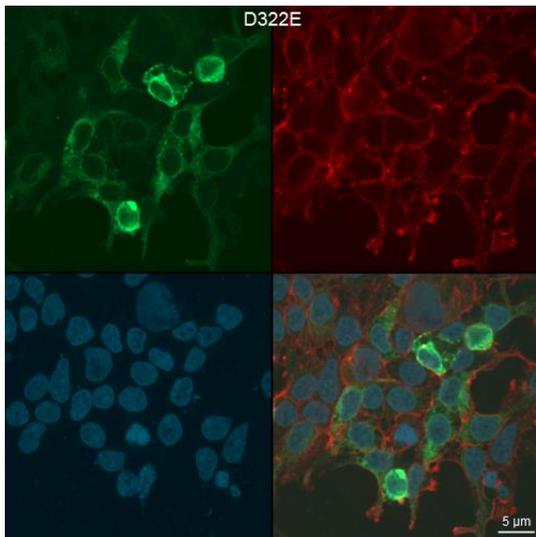
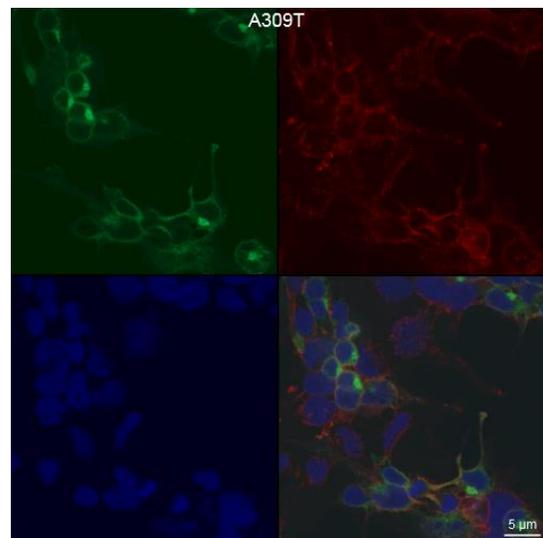
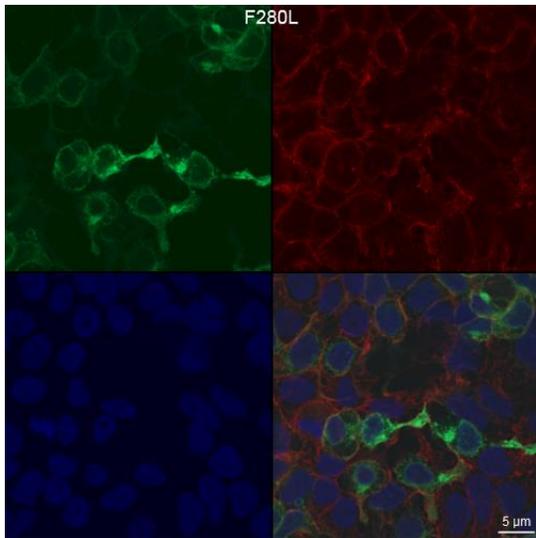
Annexe 8. Localisation cellulaire des différents mutants du récepteur DOP.

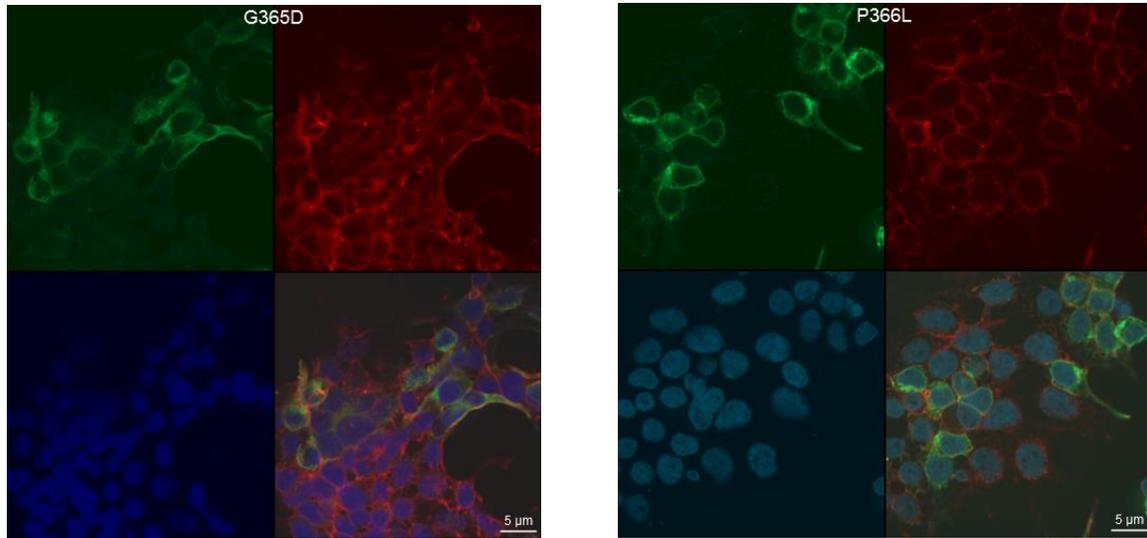










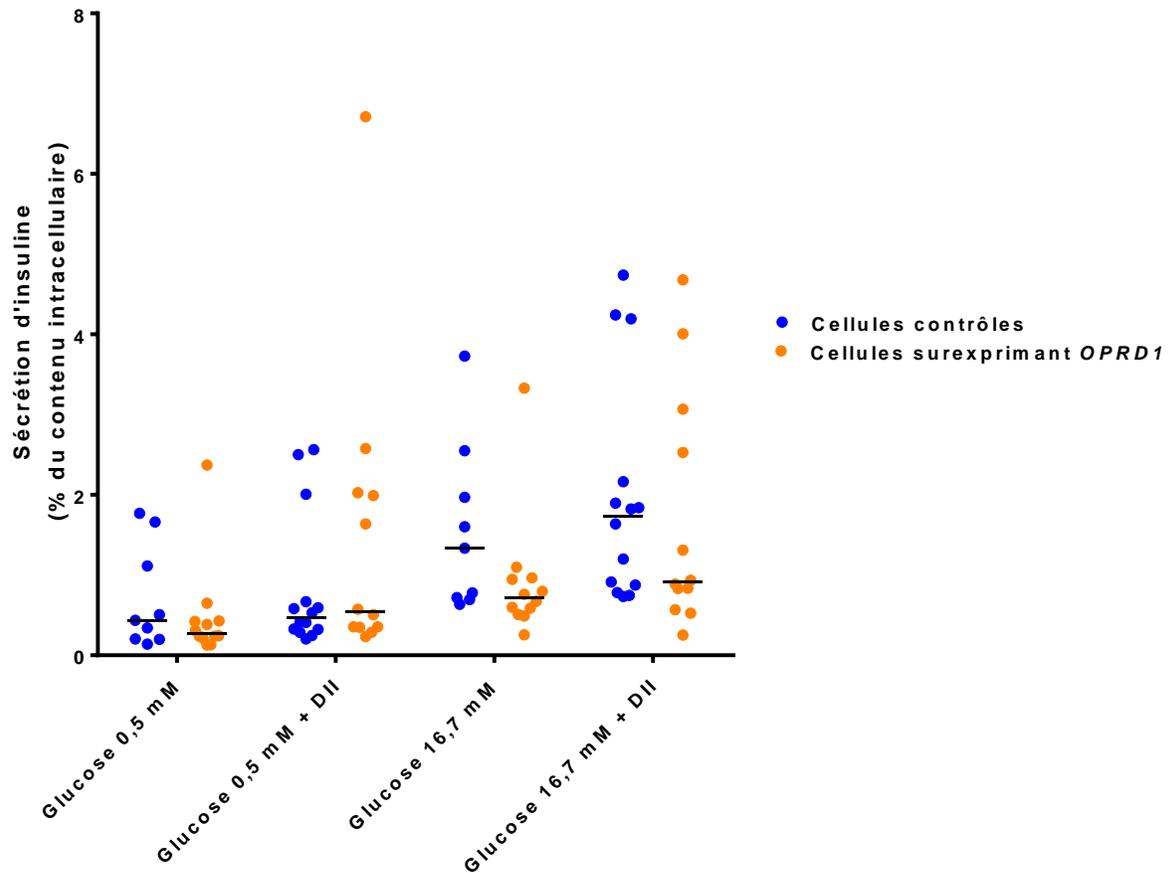


Immunofluorescence du récepteur DOP dans des cellules HEK293 surexprimant les différents variants du gène OPRD1. Le récepteur DOP est marqué en vert, l'actine F en rouge, et les noyaux cellulaires en bleu. Pour ces variants, une seule expérience d'immunofluorescence a été réalisée.

**Annexe 9.** Données cliniques des individus porteurs ou non d'un variant rare GoF ou LoF dans l'étude RaDiO. Ces données représentent le nombre, la moyenne  $\pm$  SD, ou la médiane [écart interquartile] des différents paramètres. **N** : nombre d'individus, **M** : mâle, **F** : femelle, **IMC** : indice de masse corporelle, **HDL** : HDL-cholestérol, **TG** : triglycérides, **PAS** : pression artérielle systolique, **PAD** : pression artérielle diastolique, **GoF** : gain de fonction, **LoF** : perte de fonction.

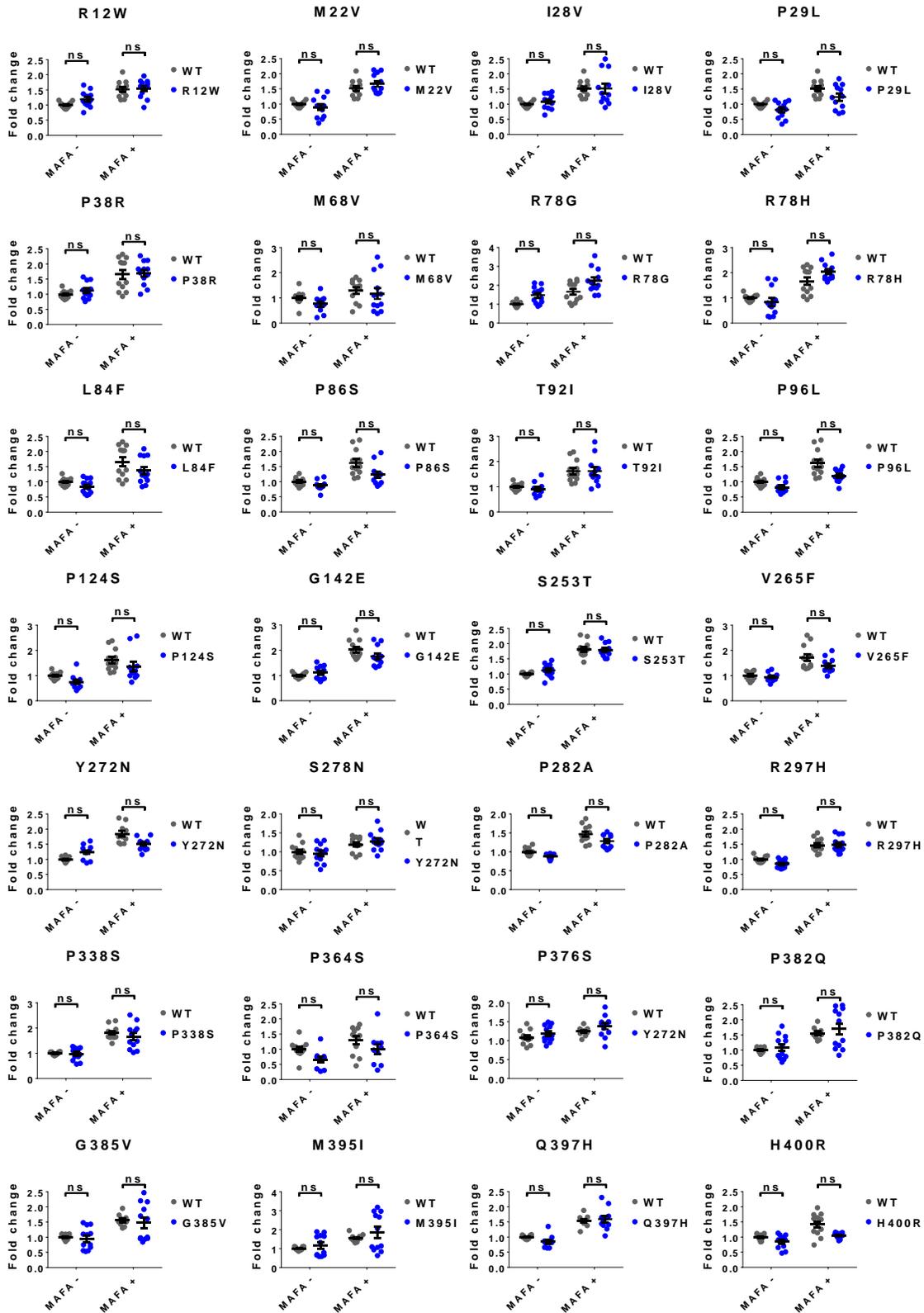
Données cliniques	Porteurs d'un variant rare GoF	Porteurs d'un variant rare LoF	Non-porteurs de variants rares LoF ou GoF
<b>N</b>	8	13	6,950
<b>Sexe</b>	M: 3/ F: 5	M: 6 / F: 7	M: 3,429/ F: 3,521
<b>Age lors de l'investigation</b>	51 $\pm$ 12	47 $\pm$ 13	51 $\pm$ 13
<b>IMC (kg/m<sup>2</sup>)</b>	25 [24–27]	27 [25–30]	26 [23–29]
<b>HDL (mmol/L)</b>	1.8 $\pm$ 0.39	1.3 $\pm$ 0.23	1.5 $\pm$ 0.45
<b>TG (mmol/L)</b>	0.55 [0.54–0.92]	0.90 [0.81–1.1]	1.1 [0.76–1.6]
<b>PAS (mmHg)</b>	120 $\pm$ 17	129 $\pm$ 20	130 $\pm$ 15
<b>PAD (mmHg)</b>	70 $\pm$ 11	77 $\pm$ 11	78 $\pm$ 9.4
<b>N surpoids/obésité (%)</b>	3 (38%)	12 (92%)	3,592 (58%)
<b>N hyperglycémie (%)</b>	4 (50%)	2 (15%)	3,188 (47%)

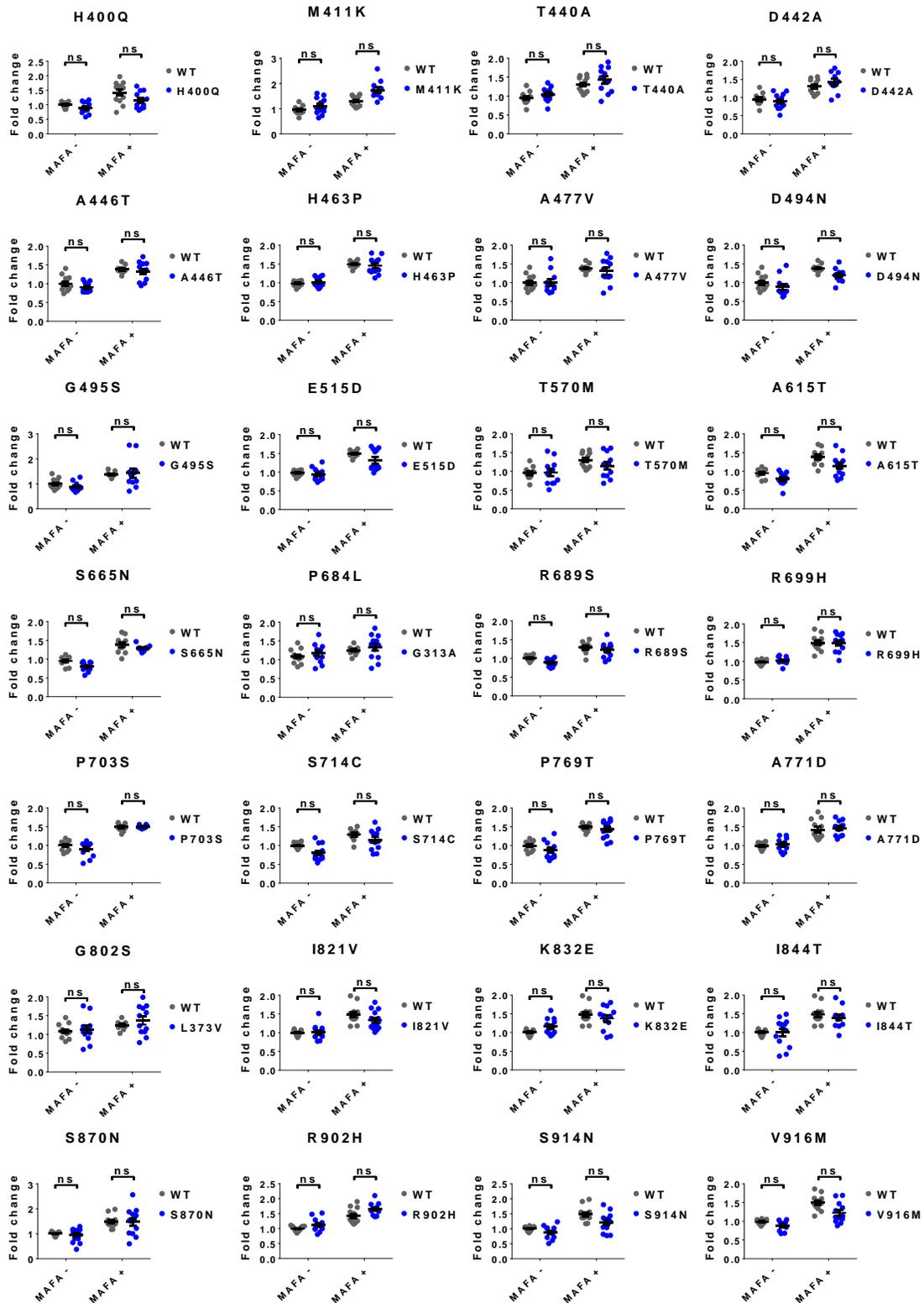
**Annexe 10.** Sécrétion d'insuline par les cellules EndoC $\beta$ H1 surexprimant le gène *OPRD1* et stimulées à bas et haut glucose, en combinaison avec l'agoniste de DOP DII.



Niveaux d'insuline sécrétée par les cellules EndoC $\beta$ H1 surexprimant ou non le gène *OPRD1* WT, à bas glucose (0,5 mM) et haut glucose (16,7 mM), couplé ou non à l'agoniste de DOP DII à 100 nM. Ces données représentent, pour chaque condition, la médiane de la sécrétion d'insuline de 5 expériences indépendantes. *DII* : [D-Ala<sup>2</sup>]-Deltorphan II.

Annexe 11. Activité luciférase des variants neutres du gène *GLIS3*.





Activité luciférase des variants neutres du gène *GLIS3*. Ces graphiques représentent, pour chaque variant, le fold change de l'activité de la SRE-luciférase (normalisée par les valeurs de  $\beta$ -galactosidase), en présence ou en absence de MAFA, par rapport à l'activité du facteur de transcription *GLIS3* WT. L'effet de chaque variant a été analysé selon un modèle ANOVA. Les graphiques représentent la moyenne  $\pm$  SEM des fold change d'au moins 4 expériences indépendantes. **WT** : wild-type.

**Annexe 12.** Evaluation des critères ACMG pour chaque variant rare du gène *GLIS3* identifié dans l'étude RaDiO. **PVS1** : très forte preuve de pathogénicité 1 (variant nul dans un gène pour lequel la perte de fonction est un mécanisme connu de la maladie) ; **PS3** : forte preuve de pathogénicité 3 (effet délétère sur le gène ou sur son produit démontré par des études fonctionnelles réalisées *in vitro* ou *in vivo*) ; **PM2** : preuve modérée de pathogénicité 2 (variant absent des individus contrôles (ou à très basse fréquence si récessif) dans des bases de données telles que Exome Sequencing Project, 1000 Genomes, ou ExAC) ; **PM4** : preuve modérée de pathogénicité 4 (changement de la longueur de la protéine dû à des délétions/insertions dans le cadre de lecture dans une région non répétée ou à des variants stop-loss (perte du codon STOP) ; **PP3** : preuve secondaire de pathogénicité (multiples preuves informatiques soutenant un effet délétère sur le gène ou son produit). **P/LP** : variant défini comme pathogène ou probablement pathogène suite à la combinaison des critères ACMG.

Mutation	PVS1	PS3	PM2	PM4	PP3	P/LP
c.34C>T, p.R12W						
c.64A>G, p.M22V						
c.82A>G, p.I28V						
c.86C>T, p.P29L			1			
c.106G>A, p.G36R		1				
c.113C>G, p.P38R						
c.202A>G, p.M68V						
c.232C>G, p.R78G						
c.233G>A, p.R78H						
c.252A>T, p.L84F						
c.256C>T, p.P86S			1			
c.275C>T, p.T92I						
c.287C>T, p.P96L						
c.358G>C, p.G120R		1				
c.370C>T, p.P124S						
c.413G>A, p.C138Y		1	1			1
c.422T>C, p.I141T		1				
c.425G>A, p.G142E			1			
c.467T>G, p.M156R		1	1		1	1
c.472G>T, p.V158F		1				
c.496C>G, p.P166A		1				
c.503C>G, p.A168G		1	1			1
c.515C>T, p.S172F		1	1			1
c.713C>G, p.S238C		1				
c.758G>C, p.S253T			1			
c.793G>T, p.V265F						
c.814T>A, p.Y272N						
c.827C>T, p.T276M		1				
c.833G>A, p.S278N			1			
c.842C>T, p.S281F		1	1			1

c.844C>G, p.P282A					
c.890G>A, p.R297H					
c.893C>A, p.S298Y		1			
c.937G>A, p.G313R		1			
c.968C>T, p.S323L		1		1	
c.974C>T, p.T325M		1			
c.1012C>T, p.P338S			1		
c.1063C>T, p.R355C		1			
c.1073G>A, p.C358Y		1	1		1
c.1085C>A, p.P362Q		1			
c.1090C>T, p.P364S					
c.1111G>T, p.G371C		1	1		1
c.1117C>G, p.L373V		1			
c.1126C>T, p.P376S					
c.1145C>A, p.P382Q					
c.1154G>T, p.G385V					
c.1185G>A, p.M395I			1		
c.1191G>C, p.Q397H					
c.1199A>G, p.H400R					
c.1200C>G, p.H400Q					
c.1210C>A, p.Q404K		1			
c.1214C>T, p.P405L		1			
c.1216G>T, p.G406C		1			
c.1227C>A, p.N409K		1			
c.1232T>A, p.M411K			1		
c.1271C>G, p.S424W		1	1		1
c.1295G>A, p.R432H		1	1		1
c.1318A>G, p.T440A					
c.1325A>C, p.D442A			1		
c.1336G>A, p.A446T			1		
c.1388A>C, p.H463P					
c.1417C>T, p.L473F		1			
c.1430C>T, p.A477V					
c.1480G>A, p.D494N					
c.1483G>A, p.G495S					
c.1518C>G, p.D506E		1			
c.1545G>C, p.E515D					
c.1709C>T, p.T570M					
c.1715A>T, p.E572V		1		1	
c.1775C>G, p.T592R		1		1	
c.1843G>A, p.A615T					
c.1881T>A, p.Y627*	1	1	1		1
c.1994G>A, p.S665N					
c.2051C>T, p.P684L					
c.2060C>T, p.S687F		1			
c.2067A>T, p.R689S			1		
c.2089G>A, p.V697M		1			
c.2096G>A, p.R699H					
c.2107C>T, p.P703S					
c.2141C>G, p.S714C					
c.2281G>A, p.D761N		1			
c.2305C>A, p.P769T					
c.2312C>A, p.A771D			1		

c.2404G>A, p.G802S					
c.2431G>C, p.E811Q		1			
c.2439C>G, p.N813K		1	1		1
c.2461A>G, p.I821V			1		
c.2494A>G, p.K832E					
c.2518G>A, p.D840N		1			
c.2531T>C, p.I844T					
c.2549C>G, p.S850C		1	1		1
c.2609G>A, p.S870N					
c.2614G>C, p.D872H		1			
c.2623C>T, p.H875Y		1	1		1
c.2636C>T, p.S879L		1			
c.2636C>G, p.S879W		1	1		1
c.2648G>C, p.G883A		1	1		1
c.2705G>A, p.R902H					
c.2710G>A, p.G904R		1			
c.2741G>A, p.S914N			1		
c.2746G>A, p.V916M					
c.2747_2749del, p.V916del		1	1	1	1
c.2749G>A, p.D917N		1	1		1
c.2770T>G, p.S924A		1	1		1
c.2785G>A, p.E929K		1			

**Annexe 13.** Données cliniques des individus porteurs ou non d'un variant rare P/LP dans l'étude RaDiO. Ces données représentent le nombre, la moyenne  $\pm$  SD, ou la médiane [écart interquartile] des différents paramètres. **N** : nombre d'individus, **M** : mâle, **F** : femelle, **IMC** : indice de masse corporelle, **P/LP** : pathogène/probablement pathogène.

Données cliniques	Porteurs d'un variant rare P/LP	Non-porteurs de variants rares P/LP
<b>N</b>	12	2160
<b>Sexe</b>	M: 7 / F: 11	M: 2620/ F: 2833
<b>Age lors de l'investigation</b>	55 $\pm$ 12	53 $\pm$ 12
<b>IMC (kg/m<sup>2</sup>)</b>	27,7 [24,0–29,6]	26,8 [23,2–29,0]
<b>N surpoids/obésité (%)</b>	8 (44%)	2109 (47%)
<b>N hyperglycémie (%)</b>	2 (11%)	249 (5,6%)

**Annexe 14.** Article « Mirror effect of delta opioid receptor mutations on type 2 diabetes and obesity” soumis pour publication

## Mirror effect of delta opioid receptor mutations on type 2 diabetes and obesity

Sarah Meulebrouck<sup>1,2</sup>, Gurvan Queniat<sup>1,2</sup>, Mehdi Derhourhi<sup>1,2</sup>, Alaa Badreddine<sup>1,2</sup>, Raphaël Boutry<sup>1,2</sup>, Audrey Leloire<sup>1,2</sup>, Emmanuelle Durand<sup>1,2</sup>, Bénédicte Toussaint<sup>1,2</sup>, Emmanuel Vaillant<sup>1,2</sup>, Hélène Loïsele<sup>1,2</sup>, Marlène Huyvaert<sup>1,2</sup>, Aurélie Dechaume<sup>1,2</sup>, Beverley Balkau<sup>3</sup>, Guillaume Charpentier<sup>4</sup>, Sylvia Franc<sup>4,5</sup>, Michel Marre<sup>6</sup>, Ronan Roussel<sup>6,7,8</sup>, Souhila Amanzougarene<sup>1,2</sup>, Raphaël Scharfmann<sup>9</sup>, Mathilde Boissel<sup>1,2†</sup>, Mickaël Canouil<sup>1,2†</sup>, Morgane Baron<sup>1,2†</sup>, Philippe Froguel<sup>1,2,10\*</sup>, Amélie Bonnefond<sup>1,2,10\*</sup>

<sup>1</sup>Inserm UMR1283, CNRS UMR8199, European Genomic Institute for Diabetes (EGID), Institut Pasteur de Lille, Lille University Hospital; Lille, 59000, France.

<sup>2</sup>Université de Lille, Lille, 59000, France.

<sup>3</sup>Paris-Saclay University, Paris-Sud University, UVSQ, Center for Research in Epidemiology and Population Health, Inserm U1018 Clinical Epidemiology; Villejuif, 94807, France.

<sup>4</sup>CERITD (*Centre d'Étude et de Recherche pour l'Intensification du Traitement du Diabète*); Evry, 91000, France.

<sup>5</sup>Department of Diabetes, Sud-Francilien Hospital, Paris-Sud University; Corbeil-Essonnes, 91100, France.

<sup>6</sup>Inserm U1138, Centre de Recherche des Cordeliers; Paris, 75006, France.

<sup>7</sup>Department of Diabetology Endocrinology Nutrition, Hôpital Bichat, DHU FIRE, Assistance Publique Hôpitaux de Paris; Paris, 75018, France.

<sup>8</sup>UFR de Médecine, University Paris Diderot, Sorbonne Paris Cité; Paris, 75005, France.

<sup>9</sup>Institut Cochin, Inserm U1016, CNRS UMR8104, Université de Paris; Paris, 75014, France.

<sup>10</sup>Department of Metabolism, Imperial College London; London, W12 0NN, United Kingdom.

\*Corresponding authors. Email: [amelie.bonnefond@inserm.fr](mailto:amelie.bonnefond@inserm.fr); [p.froguel@imperial.ac.uk](mailto:p.froguel@imperial.ac.uk)

†These authors contributed equally to this work

**Short title:** Delta opioid receptor & metabolic disorders

## **Abstract**

*Objective:* Functional genetics has been proven to be instrumental in the identification of drug targets against metabolic disorders. Opioids consumption impacts metabolic homeostasis, although mechanisms remain elusive. Here, we used functional genetics focused on opioid receptor delta gene (*OPRD1* encoding DOP) to decipher its impact on metabolism.

*Research Design and Methods:* *OPRD1* was sequenced in 6,971 participants from RaDiO study. Functional investigation of each *OPRD1* variant was done *in vitro*. Association analyses between *OPRD1* variants and metabolic traits were performed in RaDiO and confirmed in up to 222,054 participants (from UK Biobank and AMP T2D knowledge portal). Following RNA sequencing analysis in human pancreatic islets and beta cells, effect of DOP activation on insulin secretion was assessed in beta cells.

*Results:* In RaDiO, 34 *OPRD1* variants were detected and functionally analyzed. We found 12 rare, loss-of-function variants and 7 gain-of-function variants, including p.I52V variant that is common in Africans and that was analyzed in the AMP portal. In UK Biobank, we found 12 additional null *OPRD1* variants that we considered loss-of-function. Loss-of-function *OPRD1* variants were associated with higher adiposity and lower hyperglycemia risk. In contrast, gain-of-function *OPRD1* variants were associated with lower adiposity and higher type 2 diabetes risk, which was in line with observational studies of opium addicts. We found that *OPRD1* was expressed in human pancreatic islets, especially beta cells, and showed that DOP activation in human beta cells inhibited insulin secretion.

*Conclusions:* Our study highlights DOP as a major link between opioids and metabolic homeostasis in humans.

The failure to elucidate the pathophysiology of type 2 diabetes and obesity has frustrated current efforts to develop improved therapeutic and preventive strategies. Type 2 diabetes and obesity are both complex genetic disorders with an estimated heritability ranging between 40 and 70% (1,2). So far, genome-wide association studies have identified several hundreds of susceptibility genes for type 2 diabetes risk and body mass index (BMI), which were found to primarily impact pancreatic islets, especially pancreatic beta cells, and the central nervous system, respectively (3,4).

The identification and careful characterization of rare mutations causing metabolic disorders has been proven to be a fantastic tool for the identification of new drugs or new drug targets in the field of diabetes and other metabolic diseases (5–7). Among the privileged drug targets, G-protein coupled receptors (GPCRs) represent the largest family of proteins targeted by approved drugs (8). Our present study has stemmed from the observed correlation between the consumption of opioids and metabolic trait abnormalities. There are various sources of opioids: natural (*e.g.*, opium), semi-synthetic (*e.g.*, heroin), fully synthetic (*e.g.*, fentanyl), and endogenous (*e.g.*, enkephalin, endorphin) (9). In humans, a preponderance of studies has reported lower BMI among opioid users (10,11). In the same line, the majority of studies has shown that opioids cause decreased food intake in mice (12). Paradoxically, in the late 70s, a first study reported that enkephalin (*i.e.*, endogenous opioid) might be a cause of type 2 diabetes in humans (13). Since then, several studies have shown a significant association between opium and heroin addictions on altered metabolic traits including glycated hemoglobin A1c (HbA1c) (14) and plasma glucose (15,16), although these results were not always consistent (10).

Delta opioid receptor (DOP), encoded by *OPRD1*, is one of the three classical opioid receptors that are GPCRs coupled to inhibitory G proteins ( $G\alpha_i$ ). Opioid receptors are mostly expressed in

the central nervous system (17). In contrast with the other opioid receptor genes (*i.e.*, *OPRM1* and *OPRK1*), *OPRD1* is also expressed in metabolic tissues, especially in pancreas (17), and has a high genetically constraint score according to the Genome Aggregation Database (gnomAD; v2.1.1). Here, in cases with type 2 diabetes and/or obesity and controls, we performed functional genetics of *OPRD1* to decipher the putative link between *OPRD1* mutations and metabolic traits.

## RESEARCH DESIGN AND METHODS

**Study participants.** In the RaDiO study, we analyzed up to 6,971 blood DNA samples (that were successfully sequenced) from several population studies that are described in the **Supplementary Material**. Normal weight was defined as  $\text{BMI} < 25 \text{ kg m}^{-2}$ , overweight as  $25 \leq \text{BMI} < 30 \text{ kg m}^{-2}$  and obesity as  $\text{BMI} \geq 30 \text{ kg m}^{-2}$ . Hyperglycemia was defined as fasting plasma glucose levels  $\geq 5.6 \text{ mmol l}^{-1}$  and/or use of drug therapy for hyperglycemia. High-density lipoprotein (HDL) cholesterol and triglycerides were only assessed in participants who were not under lipid-lowering agents. Systolic blood pressure and diastolic blood pressure were only assessed in participants who were not under high blood pressure medications.

In UK Biobank, we analyzed up to 187,242 samples, with available exome sequencing data. This research is part of UK Biobank research application #67575. Normal weight was defined as  $\text{BMI} < 25 \text{ kg m}^{-2}$ , overweight as  $25 \leq \text{BMI} < 30 \text{ kg m}^{-2}$  and obesity as  $\text{BMI} \geq 30 \text{ kg m}^{-2}$ . The inclusion and exclusion criteria for the definition of cases with type 2 diabetes and controls are provided in the **Supplementary Material**.

***OPRD1* sequencing and detection of variants.** In the RaDiO study, DNA sequencing of *OPRD1* (NM\_000911.4) was performed by next-generation sequencing as previously described (18). Briefly, NimbleGen SeqCap EZ Choice XL target enrichment (Roche, Pleasanton, USA) was

performed according to the manufacturer's protocol for next-generation sequencing on the HiSeq 4000 system (Illumina, San Diego, USA), using a paired-end 2×150 bp protocol. The demultiplexing of sequence data was performed using bcl2fastq Conversion Software (Illumina; version 2.17). Sequence reads were then mapped to the human genome (hg19/GRCh37) using Burrows-Wheeler Aligner (version 0.7.13). The variant calling was performed using Genome Analysis ToolKit (GATK; version 3.3). Only variants with a coverage higher than 8 reads were kept for further analyses. The annotation of variants was performed using the Ensembl Perl Application Program Interfaces (version 75) and custom Perl scripts to include data from both dbSNP (version 135) and dbNSFP (version 3.0) databases. All rare coding variants (with a minor allele frequency [MAF] < 1%) which were detected in *OPRD1* had a QUAL score higher than 50. Furthermore, no variant had more than 5% of missing genotype (with a coverage below 8 reads or a QUAL score below 50) across the participants. We removed from the analyses 289 participants who had more than 5% of missing genotypes (with a coverage below 8 reads or a QUAL score below 50) across *OPRD1*, leading to 6,971 participants who were successfully sequenced and analyzed in the present study.

In UK Biobank, we used exome data from pVCF format (field #23156). Only variants with a coverage higher than 10 reads and quality GQ score higher than 20 were kept for further analyses. Annotation of variants in *OPRD1* (NM\_000911.4) was done using the Ensembl Variant Effect Predictor (VEP) tool version 103 (RefSeq). Subsequently, the analysis was focused on loss-of-function variants only, namely the variants that were genetically null (*i.e.* nonsense, frameshift, canonical  $\pm 1$  or 2 splice sites, start lost) or that were loss-of-function according to our present functional analyses. No loss-of-function variant had more than 5% of missing genotype (*i.e.* with

a coverage below 10 reads or a GQ score below 20) across the participants. All loss-of-function variants that were detected in *OPRDI* actually had a mean GQ score higher than 48.

**Ancestry assessment.** In the RaDiO study, the ancestry of participants was assessed using the first five genotypic principal components (from PC1 to PC5) calculated from >10,000 frequent single-nucleotide polymorphisms that were accurately mapped in the RaDiO study and in the 1,000 Genomes Project, as previously described (19). In UK Biobank, we used the first five genetic principal components using field #22009.

**Statistical analyses for association studies.** In the RaDiO study and UK Biobank, the rare variants were analyzed as a single cluster using the mixed-effects score test (MiST) method (20). MiST provides a score statistic  $S(\pi)$  for the mean effect ( $\pi$ ) of the cluster, and a score statistic  $S(\tau)$  for the heterogeneous effect ( $\tau$ ) of the cluster. Let the equation of the model be:  $Y = \alpha X + \pi GZ$ , where  $Y$  is the trait of interest,  $X$  is the matrix of covariates (*i.e.* age, sex, BMI, PC1 to PC5 for assessing hyperglycemia or type 2 diabetes, cholesterol levels, and blood pressure, and age, sex, PC1 to PC5 for assessing adiposity),  $G$  is the matrix of *OPRDI* variants and  $Z$  is a vector of ones repeated  $n$  times, with  $n$  the number of rare *OPRDI* variants, leading to:  $\pi GZ = \pi \sum_{i=1}^n G_i$ . As none of the association studies had significant heterogeneity, we only showed the  $P$ -values associated with the mean effect ( $\pi$ ) of the cluster. These statistical analyses were performed using the R software (version 4.0.2).

In the Accelerating Medicines Partnership (AMP) T2D knowledge portal (that was being accessed on September 2020), the association analyses for assessing the effect of the *OPRDI* variant encoding p.I52V were performed using an additive regression model adjusted for age, sex, BMI, and PC1 to PC5 (for assessing type 2 diabetes, cholesterol levels and blood pressure) or adjusted for age, sex, and PC1 to PC5 (for assessing BMI). For the analyses of cholesterol levels and blood

pressure, we only included participants without type 2 diabetes. For the analyses of type 2 diabetes, we only included controls older than 50 years.

**Plasmids.** Description of wild-type *OPRD1* plasmid, negative control *OPRD1* plasmid ('STOP'), and plasmids including each *OPRD1* mutation is provided in **Supplementary Material**.

**Culture of human embryonic kidney 293 (HEK293) cells.** HEK293 cells were cultured in Dulbecco's Modified Eagles' Medium (Gibco / Thermo Fisher Scientific, Waltham, USA) supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin (Gibco / Thermo Fisher Scientific) at 37 °C, 5% CO<sub>2</sub>.

**Luciferase assays.** Using FuGENE HD (Promega, Madison, USA), HEK293 cells were transfected with three plasmids: 375 ng ml<sup>-1</sup> of *OPRD1* plasmid (wild-type or with a variant), 225 ng ml<sup>-1</sup> of plasmid including the firefly luciferase gene under the control of Gα<sub>i</sub>-dependent serum response element (SRE), and 150 ng ml<sup>-1</sup> of plasmid including the β-galactosidase gene. Transfected cells were treated for 5 h with increasing concentrations (1, 10 and 100 nM) of either [D-Ala<sup>2</sup>]-Deltorphin II (DII) (Tocris Bioscience, Bristol, UK) or [D-Pen<sup>2</sup>,D-Pen<sup>5</sup>]-Enkephalin (DPDPE) (Tocris Bioscience) which are both DOP agonists. The experiments were performed in technical triplicate and each experiment was repeated at least four times. Further details are provided in **Supplementary Material**.

**Immunofluorescence assays.** HEK293 cells were transfected in suspension using FuGENE HD (Promega) with 375 ng ml<sup>-1</sup> of *OPRD1* plasmid (wild-type or with a variant). For immunofluorescence assays, we used the anti-DOP primary antibody (ab176324 [Abcam, Cambridge, United Kingdom]; diluted at 1/100 in blocking buffer) and a fluorescent secondary antibody (A-21206 [Thermo Fisher]; diluted at 1/500 in blocking buffer). F-actin was labelled with fluorescent phalloidin (#21836 [Thermo Fisher Scientific]; diluted at 1/300 in blocking buffer),

and nuclei were labelled with DAPI (R-37606 [Thermo Fisher Scientific]). Further details are provided in **Supplementary Material**.

**Western blot assays.** HEK293 cells were transfected in suspension using FuGENE HD (Promega), with 375 ng ml<sup>-1</sup> of *OPRD1* plasmid (wild-type or with a variant). For the Western blot assays, we used the anti-DOP antibody (ab176324 [Abcam]; diluted at 1/1000 in the blocking buffer) and a fluorescent anti-rabbit antibody (SA5-35571 [Invitrogen / Thermo Fisher Scientific]; diluted at 1/5000 in the blocking buffer). To normalize our data,  $\beta$ -actin was used as a loading control (anti- $\beta$ -actin primary antibody: #3700 [Cell Signaling Technology, Leiden, The Netherlands] diluted at 1/5000 in the blocking buffer; fluorescent anti-mouse secondary antibody: #35518 [Thermo Fisher Scientific] diluted at 1/5000 in the blocking buffer or HRP anti-mouse secondary antibody: #NA931 [Merck / Sigma-Aldrich] diluted at 1/5000 in the blocking buffer; revelation via the Odyssey CLx imaging system at 700 nm). At least, three independent experiments for each mutant were performed. Further details are provided in **Supplementary Material**.

**RNA-sequencing (RNA-seq) analysis.** We analyzed RNA-seq data from human pancreatic islets ( $n = 11$ ) and mouse pancreatic islets ( $n = 14$ ), which were available through the GEO data repository (21–24). Fastq files were downloaded from GEO using fastq-dump (version 2.9.6). The quality control of Fastq files was performed using FastQC (version 0.11.9). Sequencing reads were aligned to the human genome (hg38) or the mouse genome (mm10) using STAR aligner (version 2.7.1a). On average, 51.6 million reads and 52.0 million reads accurately mapped against the genomes of human and mouse pancreatic islets, respectively. Processed count tables and transcripts per million (TPM) were then obtained per gene using RSEM (v1.3.0) and Ensembl gene annotation (Ensembl Archive Release 100) using biomaRT package (v2.45) with R (v3.6.3).

***Culture of EndoCβH1 cells.*** EndoCβH1 cells were cultured in Optiβ1 medium (Univercell Biosolutions, Toulouse, France), at a density of  $2.5 \times 10^6$  in T25 flasks primarily coated with βcoat (Univercell Biosolutions) according to manufacturer's instructions (25). These cells were cultured at 37 °C, 5% CO<sub>2</sub>.

***Electroporation of EndoCβH1 cells.*** EndoCβH1 cells were electroporated using the Cell Line Nucleofactor Kit V (Lonza, Basel, Switzerland). Further details are provided in **Supplementary Material**.

***Insulin secretion assays in EndoCβH1 cells.*** At two days after electroporation, EndoCβH1 cells seeded in the 96-well plate were incubated overnight in a starving medium, and were rinsed at three days after electroporation. We subsequently performed insulin secretion assays as previously described (18). We used the R package called INsulin Secretion ANalysEr (Insane; R package version 0.1.0), which performs critical quality control (QC) steps to monitor experimental stages in addition to statistical analyses, adjusting for technical and confounding variables. The fold change in insulin secretion for cells overexpressing (or not) *OPRDI* was analyzed using linear regression adjusted for experimental conditions, from five independent experiments. Further details are provided in **Supplementary Material**.

## RESULTS

### Detection of 34 coding *OPRD1* variants in RaDiO

Among 6,971 participants from the RaDiO study (including 97.2% of European-descent individuals) (19), we detected 31 rare coding *OPRD1* variants, including one nonsense variant (encoding p.E355\*), in a total of 36 carriers, and three coding *OPRD1* variants that were common (*i.e.*, with a MAF above 1%) in at least one population from gnomAD (v2.1.1) (**Fig.1, Supplementary Table 1**). Among the 31 rare variants, 14 were novel (*i.e.*, not reported in gnomAD [v2.1.1]; **Supplementary Table 1**). Using a pathogenicity threshold of 50%, six missense *OPRD1* variants were predicted to be deleterious according to REVEL that is based on 13 software programs for predicting the pathogenicity of missense variants (**Supplementary Table 1**) (26). Only two variants (encoding p.D95N and p.R160G) were predicted to be deleterious when using a more stringent threshold (75%) (**Supplementary Table 1**). Before functional investigations, the burden of rare *OPRD1* variants was not associated with any metabolic traits (**Supplementary Table 2**).

### Identification of loss- and gain-of-function *OPRD1* variants via functional genetics

Each mutant or wild-type *OPRD1* was overexpressed in human embryonic kidney 293 (HEK293) cells, and the SRE luciferase activity was analyzed in response to increasing concentrations of DII and DPDPE, which are two specific agonists for DOP (27). When compared to wild-type *OPRD1*, we found that 12 *OPRD1* variants encoding p.P14R, p.G36E, p.K81N, p.D95N, p.A96V, p.N131D, p.M132I, p.R160G, p.R192Q, p.A221S, p.G248V and p.D322E (loss-of-function variants) significantly decreased SRE luciferase activity in response to both DOP agonists (**Supplementary Fig.1**). In contrast, seven *OPRD1* variants encoding p.P24S, p.A32G, p.I52V, p.I183V, p.K250R, p.T361N and p.P366L (gain-of-function variants) significantly increased the

SRE luciferase activity in response to both DOP agonists when compared to wild-type *OPRD1* (**Supplementary Fig.2**). The other *OPRD1* variants (neutral variants) did not have significant effect on SRE luciferase activity in response to both DOP agonists (**Supplementary Fig.3**). Immunofluorescence assays showed that all the mutants were effectively expressed and localized at the plasma membrane of HEK293 cells, except for two loss-of-function mutants p.P14R and p.G36E (**Supplementary Fig.4**). The variants encoding these two mutants were both located in the extracellular terminus region (**Fig.1**). Western blots showed that these two mutants tended to have a lower expression than wild-type DOP, while the other DOP mutants were expressed at the same magnitude as wild-type DOP (**Supplementary Fig.5**).

The amino acids affected by missense loss-of-function *OPRD1* variants were significantly more conserved than the amino acids affected by gain-of-function variants (83% versus 29%,  $P = 0.027$ ; **Fig.1**). In this regard, half of the loss-of-function variants and only one gain-of-function variant were novel (**Supplementary Table 1**). The burden of gain-of-function *OPRD1* variants tended to be higher in extracellular and intracellular terminus regions ( $P = 0.085$ ; **Fig.1**), where amino acids were significantly less conserved than in transmembrane domains and loops (16% versus 79%,  $P = 5.7 \times 10^{-22}$ ; **Fig.1**). Based on our results, the REVEL tool for predicting the pathogenicity of missense variants had a poor sensitivity (33% for loss-of-function variants versus neutral variants; and 0% for gain-of-function variants versus neutral variants) but high specificity (86%) using a pathogenicity threshold of 50% (**Supplementary Table 1**).

### **Mirror effect of loss- and gain-of-function *OPRD1* variants on metabolic homeostasis**

We then analyzed the effect of the burden of rare, gain- or loss-of-function variants on metabolic homeostasis. In these analyses, we also considered the rare, null *OPRD1* variant encoding p.E355\*

as loss-of-function. In the RaDiO study, all carriers of rare, loss- or gain-of-function *OPRD1* variants were heterozygote and were of European ancestry except for one individual.

All carriers of a rare, loss-of-function variant were overweight or obese, except for one individual (**Supplementary Table 3**). Using the MiST method adjusted for age, sex and ancestry, the burden of rare, loss-of-function *OPRD1* variants was indeed significantly higher among individuals with overweight or obesity (0.31% of carriers among cases *versus* 0.038% of carriers among controls;  $P = 0.0054$  with an odds ratio [OR] of 11; 95% confidence interval [CI], 2.1–194; **Table 1**) and its effect was significant on a higher BMI ( $P = 0.044$  with an effect [ $\pi$ ] of  $0.35 \pm 0.17$ ; **Table 1**). In up to 187,242 participants from UK Biobank, we detected 12 additional rare, null variants (encoding p.M1?, p.E10\*, p.S44\*, p.V59Cfs\*22, p.S106Vfs\*5, p.P115Afs\*31, p.W114\*, p.C232Vfs\*80, p.Y233\*, p.I277Hfs\*33, p.F320Sfs\*158 and p.R356Pfs\*119) that we considered loss-of-function, as well as 7 loss-of-function variants previously detected in RaDiO (**Supplementary Table 4**). We confirmed that the burden of rare, loss-of-function *OPRD1* variants in UK Biobank was significantly higher among individuals with overweight or obesity ( $P = 0.0047$  with an OR of 2.6; 95% confidence interval [CI], 1.4–5.5; **Table 1**) and its effect was significant on a higher BMI ( $P = 0.0028$  with a  $\pi$  of  $0.081 \pm 0.027$ ; **Table 1**) using MiST adjusted for age, sex and ancestry.

In contrast, except for one individual, all carriers from RaDiO of rare, gain-of-function *OPRD1* variants were not obese but the effect of the burden on decreased risk of overweight or obesity was not significant (0.083% of carriers among cases *versus* 0.19% of carriers among controls; **Table 2**). The gain-of-function *OPRD1* variant encoding p.I52V (rs111622802; that is common in Africans only) was independently analyzed in the Accelerating Medicines Partnership (AMP) T2D Knowledge Portal which enabled the assessment of its metabolic effect in African individuals

including cases with type 2 diabetes (28), while the MAF of rs111622802 was too low in both RaDiO and UK Biobank. We found that this variant was significantly associated with a lower BMI ( $P = 0.0038$  with an effect  $[\beta]$  of  $-0.027 \pm 0.0093$ ; **Table 2**), when adjusting for age, sex and ancestry.

In RaDiO, the burden of rare, loss-of-function variants tended to be lower among individuals with hyperglycemia or type 2 diabetes (0.063% of carriers among cases *versus* 0.25% of carriers among controls;  $P = 0.054$  with an OR of 0.23; 95% CI, 0.033–0.95; **Table 1**), using MiST adjusted for BMI, age, sex and ancestry. In the same line, none carriers of a rare, loss-of-function *OPRD1* variant in UK Biobank presented with type 2 diabetes (**Table 1**). In contrast, in the AMP T2D knowledge portal, we found that the gain-of-function *OPRD1* variant encoding p.I52V was significantly associated with a higher risk of type 2 diabetes in 34,812 participants ( $P = 6.1 \times 10^{-5}$  with an OR of 1.7; 95% CI, 1.3–2.2; **Table 2**), when adjusting for age, sex, BMI and ancestry. This association signal was most probably driven by the results found in 4,966 African participants ( $P = 3.6 \times 10^{-6}$  with an OR of 2.0; 95% CI, 1.5–2.6; **Table 2**).

In RaDiO, the burden of rare, gain-of-function variants improved lipid homeostasis via higher HDL cholesterol levels and lower triglycerides levels ( $P = 0.049$  with a  $\pi$  of  $0.33 \pm 0.17$  and  $P = 0.046$  with a  $\pi$  of  $-0.37 \pm 0.18$ , respectively; **Table 2**), using MiST methods adjusted for BMI, age, sex and ancestry. In the same line, the gain-of-function variant encoding p.I52V was associated with improved lipid homeostasis, via a significant effect on lower triglyceride levels and lipoprotein (LDL) cholesterol levels ( $P = 0.018$  with a  $\beta$  of  $-0.13 \pm 0.054$ ,  $P = 0.0026$  with a  $\beta$  of  $-0.37 \pm 0.12$ , respectively; **Table 2**) and a nominal effect on increased HDL levels in African participants AMP T2D knowledge portal ( $P = 0.055$  with a  $\beta$  of  $0.095 \pm 0.049$ ; **Table 2**), when adjusting for age, sex, BMI and ancestry.

### **Impaired glucose-stimulated insulin secretion by DOP activation in human beta cells**

Higher adiposity is usually strongly associated with higher risk of hyperglycemia and type 2 diabetes (29). However, according to the present data, *OPRD1* loss of function associates with a higher adiposity and a lower risk of hyperglycemia, while *OPRD1* gain of function has opposite effects, in line with long-term consumption of opium (10). We hypothesized that this discrepancy might be due to independent effects of DOP in the central nervous system *versus* the pancreas. We analyzed RNA-seq data from human and mouse pancreatic islets which were available through the gene expression omnibus (GEO) data repository (21–24,30). *OPRD1*, but not the other opioid receptor genes (*i.e.*, *OPRK1* or *OPRM1*), was expressed in human pancreatic islets (**Fig.2A**). Of note, neither *Oprd1*, *Oprm1* nor *Oprk1* was expressed in mouse pancreatic islets (**Fig.2A**). According to other RNA-seq data (31), *OPRD1* was expressed in human purified beta cells, and this expression was significantly enriched in these purified beta cells when compared to beta-cell depleted islet preparations ( $P = 0.0075$ , accounting for islet purity and adjusted for multiple testing using the Benjamini-Hochberg method) (**Fig.2B**). As the human EndoC $\beta$ H1 beta-cell line did not express *OPRD1*, its expression was enhanced by the electroporation of wild-type *OPRD1* plasmid (**Fig.2C**). The null *OPRD1* expression in EndoC $\beta$ H1 cells that are fetal cells was in line with previously published RNA-seq data from human fetal beta cells (32). In EndoC $\beta$ H1 overexpressing *OPRD1*, we found that DII DOP agonist significantly impaired glucose-stimulated insulin secretion ( $P = 0.018$ ; **Fig.2D**). Furthermore, the enhancement of *OPRD1* expression impaired glucose-stimulated insulin secretion in the presence of DII DOP agonist ( $P = 0.0076$ ; **Fig.2D**).

## DISCUSSION

Based on functional human genetic experiments, we found that *OPRD1* loss of function led to higher adiposity but tended to decrease the risk of hyperglycemia. In contrast, *OPRD1* gain of function was associated with lower adiposity and improved lipid metabolism (regardless of BMI) but led to markedly higher risk of type 2 diabetes.

The present results related to *OPRD1* gain of function are in line with several observational studies of people addicted to opium and its derivatives, which preponderantly reported an association between opioid consumption and lower adiposity but higher risk of hyperglycemia (10–12,14–16). However, to our knowledge, no studies had incriminated the putative impact of DOP on type 2 diabetes in humans before the present study. In fact, the pathophysiological mechanisms explaining the effects of opioids on type 2 diabetes and related traits in human addicts have been surprisingly unexplored in the last three decades (12), while morphine that is the most prevalent and pharmacologically active alkaloid in opium, has high affinity for DOP (33,34). Further, the lack of data might be due to the fact that *Oprd1* is not expressed in mouse pancreatic islets (**Fig.2A**). In humans, early studies showed that met-enkephalin and  $\beta$ -endorphin (*i.e.*, high-affinity, endogenous ligands for DOP (33)) inhibited insulin secretion, and consequently increased plasma glucose concentrations (35–37). Here, we found that DOP activation through its opioid agonist DII significantly decreased glucose-stimulated insulin secretion from human beta cells. Therefore, we suggest that the contribution of opium and its derivatives to increased type 2 diabetes risk is at least partly explained by their stimulating effect on DOP in pancreatic islets.

Regarding adiposity, mice deleted for *Oprd1* were hyperphagic, which was in line with the fact that opioids decreased food intake in mice (12), but gained less weight when compared to wild-type mice fed an high-energy diet (38). This was due to an enhancement of the thermogenic

program in brown adipose tissue (38). Although not reaching genome-wide significance, several genetic associations between common single nucleotide polymorphisms (SNPs) in *OPRD1* and anorexia nervosa have been shown (39,40). In this respect, the anorexia nervosa associated risk allele of rs569356 located in the promoter of *OPRD1* increased *OPRD1* expression (41).

Our present study highlights DOP as a new key GPCR involved in metabolic disorders – in particular type 2 diabetes, which might represent a promising therapeutic target.

The present study is limited by the fact that we did not functionally investigate the other classical opioid receptor genes (*i.e.*, *OPRM1* and *OPRK1* encoding mu and kappa opioid receptors, respectively). Therefore, we cannot rule out their putative involvement in the link between opioids and metabolism, in particular adiposity.

Another limitation of the present study is the fact that we analyzed only one signaling pathway of DOP in response to two agonists for the definition of gain- and loss-of-function *OPRD1* mutations, while GPCRs trigger various pathways when activated. Gain- or loss-of-function mutations in a GPCR gene can differently affect these pathways, as found for deleterious mutations of *MTNR1B* encoding melatonin receptor 1B (42). In this regard, the presently observed genetic associations between gain- or loss-of-function *OPRD1* mutations and metabolic traits are primarily supported by our observed mutation effects on the  $G\alpha_i$ -dependent SRE pathway.

## References

1. Bonnefond A, Froguel P. Rare and common genetic events in type 2 diabetes: what should biologists know? *Cell Metab.* 2015 Mar 3;21(3):357–68.
2. El-Sayed Moustafa JS, Froguel P. From obesity genetics to the future of personalized obesity therapy. *Nat Rev Endocrinol.* 2013 Jul;9(7):402–13.
3. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature.* 2015 Feb 12;518(7538):197–206.
4. Ndiaye FK, Ortalli A, Canouil M, Huyvaert M, Salazar-Cardozo C, Lecoeur C, et al. Expression and functional assessment of candidate type 2 diabetes susceptibility genes identify four new genes contributing to human insulin secretion. *Mol Metab.* 2017;6(6):459–70.
5. Babenko AP, Polak M, Cavé H, Busiah K, Czernichow P, Scharfmann R, et al. Activating mutations in the ABCC8 gene in neonatal diabetes mellitus. *N Engl J Med.* 2006 Aug 3;355(5):456–66.
6. Kühnen P, Clément K, Wiegand S, Blankenstein O, Gottesdiener K, Martini LL, et al. Proopiomelanocortin Deficiency Treated with a Melanocortin-4 Receptor Agonist. *N Engl J Med.* 2016 Jul 21;375(3):240–6.
7. Fitzgerald K, White S, Borodovsky A, Bettencourt BR, Strahs A, Clausen V, et al. A Highly Durable RNAi Therapeutic Inhibitor of PCSK9. *N Engl J Med.* 2017 05;376(1):41–51.
8. Sriram K, Insel PA. G Protein-Coupled Receptors as Targets for Approved Drugs: How Many Targets and How Many Drugs? *Mol Pharmacol.* 2018;93(4):251–8.
9. Pathan H, Williams J. Basic opioid pharmacology: an update. *Br J Pain.* 2012 Feb;6(1):11–6.
10. Najafipour H, Beik A. The Impact of Opium Consumption on Blood Glucose, Serum Lipids and Blood Pressure, and Related Mechanisms. *Front Physiol.* 2016;7:436.

11. Elman I, Howard M, Borodovsky JT, Mysels D, Rott D, Borsook D, et al. Metabolic and Addiction Indices in Patients on Opioid Agonist Medication-Assisted Treatment: A Comparison of Buprenorphine and Methadone. *Sci Rep.* 2020 27;10(1):5617.
12. Vuong C, Van Uum SHM, O'Dell LE, Lutfy K, Friedman TC. The effects of opioids and opioid analogs on animal and human endocrine systems. *Endocr Rev.* 2010 Feb;31(1):98–132.
13. Leslie RD, Pyke DA, Stubbs WA. Sensitivity to enkephalin as a cause of non-insulin dependent diabetes. *Lancet.* 1979 Feb 17;1(8112):341–3.
14. Karam GA, Reisi M, Kaseb AA, Khaksari M, Mohammadi A, Mahmoodi M. Effects of opium addiction on some serum factors in addicts with non-insulin-dependent diabetes mellitus. *Addict Biol.* 2004 Mar;9(1):53–8.
15. Gozashti MH, Yazdi F, Salajegheh P, Dehesh MM, Divsalar K. Fasting Blood Glucose and Insulin Level in Opium Addict versus Non-Addict Individuals. *Addict Health.* 2015 Winter-Spring;7(1–2):54–9.
16. Ceriello A, Giugliano D, Passariello N, Quatraro A, Dello Russo P, Torella R, et al. Impaired glucose metabolism in heroin and methadone users. *Horm Metab Res.* 1987 Sep;19(9):430–3.
17. Peng J, Sarkar S, Chang SL. Opioid receptor expression in human brain and peripheral tissues using absolute quantitative real-time RT-PCR. *Drug Alcohol Depend.* 2012 Aug 1;124(3):223–8.
18. Baron M, Maillet J, Huyvaert M, Dechaume A, Boutry R, Loisele H, et al. Loss-of-function mutations in MRAP2 are pathogenic in hyperphagic obesity with hyperglycemia and hypertension. *Nat Med.* 2019;25(11):1733–8.
19. Bonnefond A, Boissel M, Bolze A, Durand E, Toussaint B, Vaillant E, et al. Pathogenic variants in actionable MODY genes are associated with type 2 diabetes. *Nat Metab.* 2020 Oct;2(10):1126–34.
20. Sun J, Zheng Y, Hsu L. A unified mixed-effects model for rare-variant association in sequencing studies. *Genet Epidemiol.* 2013 May;37(4):334–44.

21. Bone RN, Oyebamiji O, Talware S, Selvaraj S, Krishnan P, Syed F, et al. A Computational Approach for Defining a Signature of  $\beta$ -Cell Golgi Stress in Diabetes. *Diabetes*. 2020 Nov;69(11):2364–76.
22. Kolic J, Beet L, Overby P, Cen HH, Panzhinskiy E, Ure DR, et al. Differential Effects of Voclosporin and Tacrolimus on Insulin Secretion From Human Islets. *Endocrinology*. 2020 Nov 1;161(11).
23. Gharavy SNM, Owen B, Millership SJ, Chabosseau P, Pizza G, Martinez-Sanchez A, et al. Sexually dimorphic roles for the type 2 diabetes-associated *C2cd4b* gene in murine glucose homeostasis [Internet]. *Genetics*; 2020 May [cited 2020 Oct 27]. Available from: <http://biorxiv.org/lookup/doi/10.1101/2020.05.18.099200>
24. Berthault C, Staels W, Scharfmann R. Purification of pancreatic endocrine subsets reveals increased iron metabolism in beta cells. *Mol Metab*. 2020 Aug 5;42:101060.
25. Ravassard P, Hazhouz Y, Pechberty S, Bricout-Neveu E, Armanet M, Czernichow P, et al. A genetically engineered human pancreatic  $\beta$  cell line exhibiting glucose-inducible insulin secretion. *J Clin Invest*. 2011 Sep;121(9):3589–97.
26. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am J Hum Genet*. 2016 Oct 6;99(4):877–85.
27. Raynor K, Kong H, Chen Y, Yasuda K, Yu L, Bell GI, et al. Pharmacological characterization of the cloned kappa-, delta-, and mu-opioid receptors. *Mol Pharmacol*. 1994 Feb;45(2):330–4.
28. Flannick J, Mercader JM, Fuchsberger C, Udler MS, Mahajan A, Wessel J, et al. Exome sequencing of 20,791 cases of type 2 diabetes and 24,440 controls. *Nature*. 2019;570(7759):71–6.
29. American Diabetes Association. 8. Obesity Management for the Treatment of Type 2 Diabetes: Standards of Medical Care in Diabetes-2020. *Diabetes Care*. 2020 Jan;43(Suppl 1):S89–97.

30. Rachdi L, Maugein A, Pechberty S, Armanet M, Hamroune J, Ravassard P, et al. Regulated expression and function of the GABAB receptor in human pancreatic beta cell line and islets. *Sci Rep*. 2020 Aug 10;10(1):13469.
31. Nica AC, Ongen H, Irminger J-C, Bosco D, Berney T, Antonarakis SE, et al. Cell-type, allelic, and genetic signatures in the human pancreatic beta cell transcriptome. *Genome Res*. 2013 Sep;23(9):1554–62.
32. Blodgett DM, Nowosielska A, Afik S, Pechhold S, Cura AJ, Kennedy NJ, et al. Novel Observations From Next-Generation RNA Sequencing of Highly Purified Human Adult and Fetal Islet Cell Subsets. *Diabetes*. 2015 Sep;64(9):3172–81.
33. Borsodi A, Bruchas M, Caló G, Chavkin C, Christie MJ, Civelli O, et al. Opioid receptors (version 2019.4) in the IUPHAR/BPS Guide to Pharmacology Database. *GtoPdb CITE* [Internet]. 2019 Sep 16 [cited 2020 Nov 17];2019(4). Available from: <http://journals.ed.ac.uk/gtopdb-cite/article/view/3202>
34. Carlin MG, Dean JR, Ames JM. Opium Alkaloids in Harvested and Thermally Processed Poppy Seeds. *Front Chem*. 2020;8:737.
35. Giugliano D, Quatraro A, Consoli G, Ceriello A, Torella R, D’Onofrio F. Inhibitory effect of enkephalin on insulin secretion in healthy subjects and in non insulin-dependent diabetic subjects. *Metabolism*. 1987 Mar;36(3):286–9.
36. Giugliano D, Cozzolino D, Ceriello A, Salvatore T, Paolisso G, Torella R. Beta-endorphin and islet hormone release in humans: evidence for interference with cAMP. *Am J Physiol*. 1989 Sep;257(3 Pt 1):E361-366.
37. Feldman M, Kiser RS, Unger RH, Li CH. Beta-endorphin and the endocrine pancreas. Studies in healthy and diabetic human beings. *N Engl J Med*. 1983 Feb 17;308(7):349–53.

38. Czyzyk TA, Romero-Picó A, Pintar J, McKinzie JH, Tschöp MH, Statnick MA, et al. Mice lacking  $\delta$ -opioid receptors resist the development of diet-induced obesity. *FASEB J*. 2012 Aug;26(8):3483–92.
39. Brown KMO, Bujac SR, Mann ET, Campbell DA, Stubbins MJ, Blundell JE. Further evidence of association of OPRD1 & HTR1D polymorphisms with susceptibility to anorexia nervosa. *Biol Psychiatry*. 2007 Feb 1;61(3):367–73.
40. Wang K, Zhang H, Bloss CS, Duvvuri V, Kaye W, Schork NJ, et al. A genome-wide association study on common SNPs and rare CNVs in anorexia nervosa. *Mol Psychiatry*. 2011 Sep;16(9):949–59.
41. Zhang H, Gelernter J, Gruen JR, Kranzler HR, Herman AI, Simen AA. Functional impact of a single-nucleotide polymorphism in the OPRD1 promoter region. *J Hum Genet*. 2010 May;55(5):278–84.
42. Karamitri A, Plouffe B, Bonnefond A, Chen M, Gallion J, Guillaume J-L, et al. Type 2 diabetes-associated variants of the MT2 melatonin receptor affect distinct modes of signaling. *Sci Signal*. 2018 Aug 28;11(545).

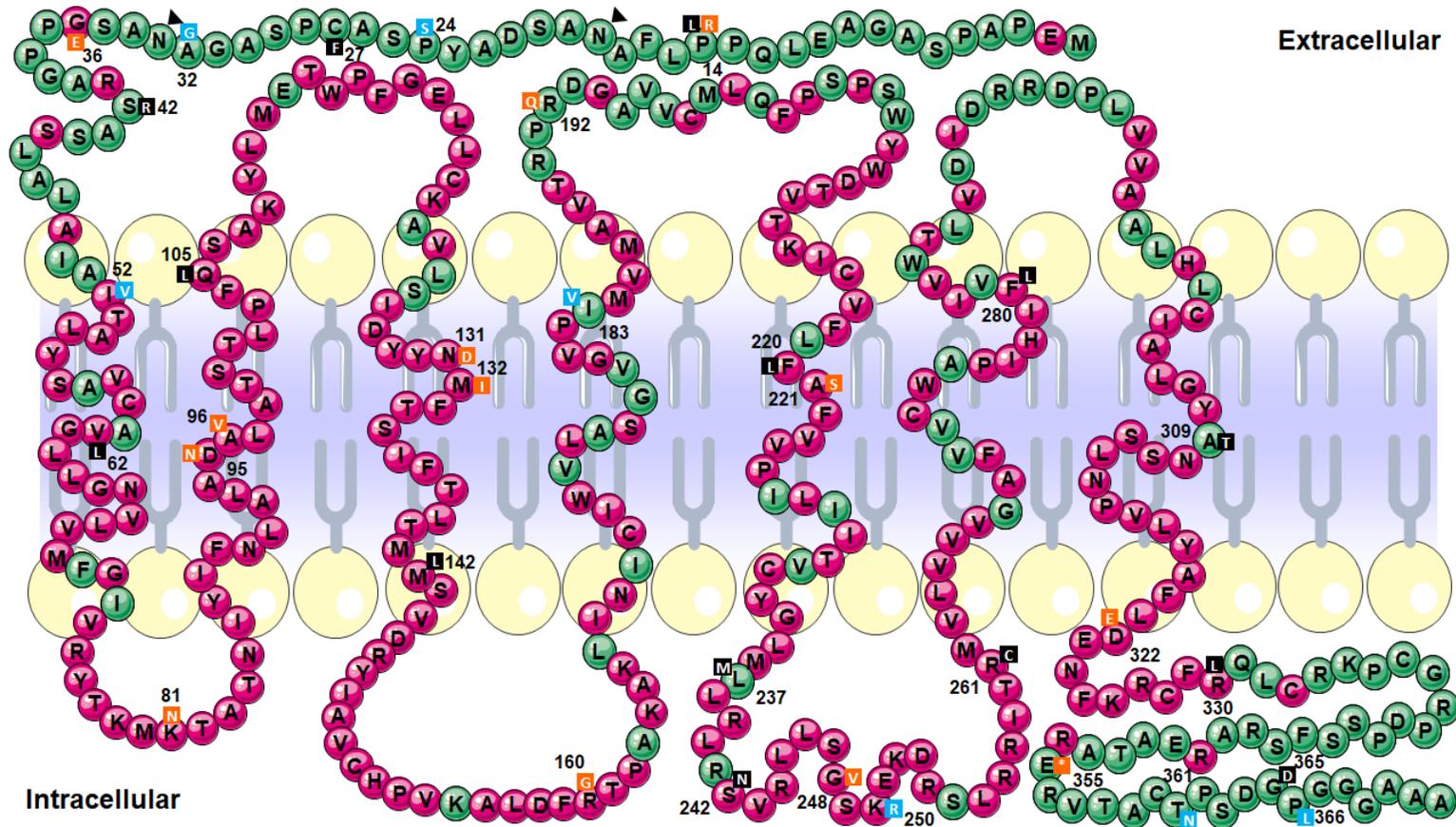
**Acknowledgments:** We are grateful to all individuals included in the different cohort studies. We are grateful to Julie Maillet, Frédéric Allegaert and Nicolas Larcher for technical assistance. We thank the Genome Aggregation Database (gnomAD) and the groups that provided exome and genome variant data to this resource. A full list of contributing groups can be found at <https://gnomad.broadinstitute.org/about>. We would like to thank the Type 2 Diabetes Knowledge Portal and the groups that provided data to this resource. This research has been conducted using the UK Biobank Application #67575. This study was funded by the French National Research Agency (ANR-10-LABX-46 [European Genomics Institute for Diabetes] to PF and AB), the French National Research Agency (ANR-10-EQPX-07-01 [LIGAN-PM] to PF and AB), the European Research Council (ERC GEPIDIAB – 294785, to PF; ERC Reg-Seq – 715575, to AB), the National Center for Precision Diabetic Medicine – PreciDIAB, which is jointly supported by the French National Agency for Research (ANR-18-IBHU-0001), by the European Union (FEDER), by the Hauts-de-France Regional Council and by the European Metropolis of Lille (MEL).

**Author contributions:** Conceptualization: PF, ABonnefond; Methodology/Investigation: SM, GQ, RB, AL, HL, MH, AD, MBaron, ED, BT, EV, SA, MD, ABadreddine, MBoissel, MC, ABonnefond; Sample acquisition: BB, GC, SF, MM, RR, RS and PF; Visualization: SM, ABonnefond; Funding acquisition: PF, ABonnefond; Supervision: PF, ABonnefond; Writing – original draft: SM, ABonnefond; Writing – review & editing: all authors

ABonnefond is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

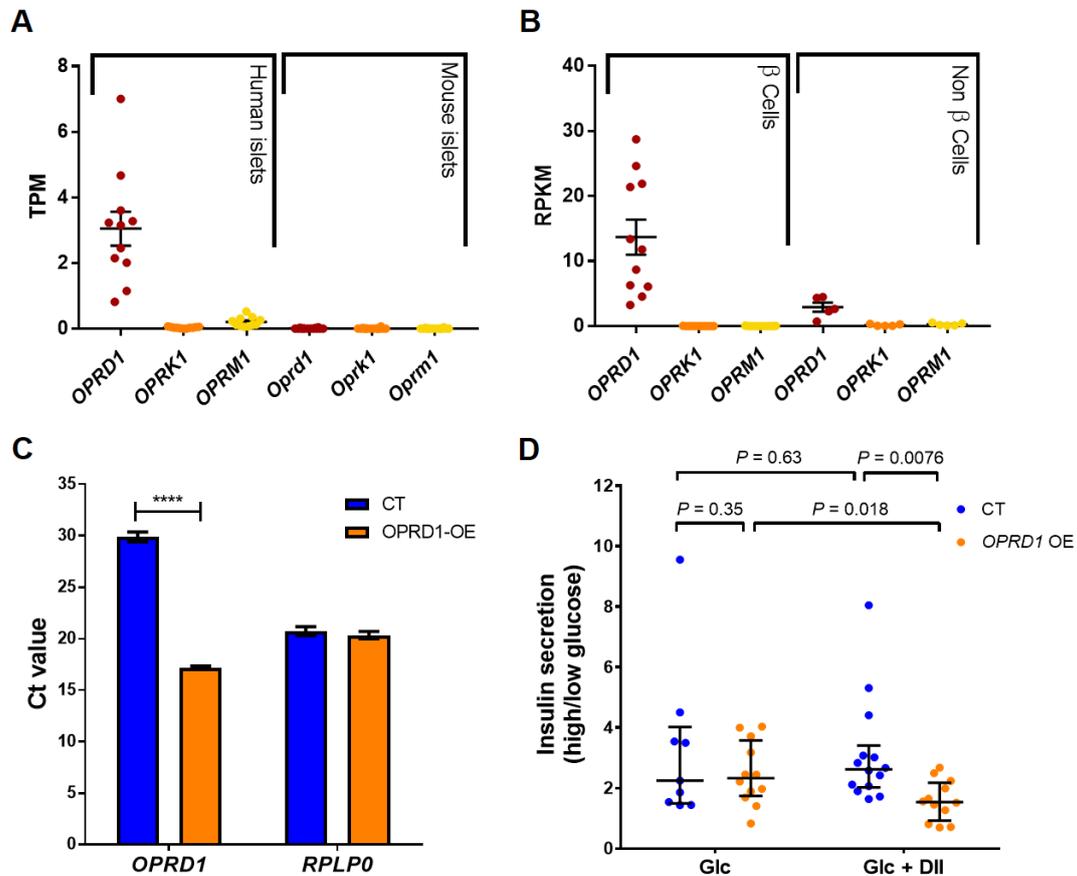
**Conflict-of-interest statement:** Authors declare that they have no competing interests.

Fig. 1. Detection of 34 *OPRD1* coding variants.



Pink bubbles represent amino acids in loci of high sequence homology (in human versus chimpanzee, Northern white-cheeked gibbon, macaque, olive baboon, rat, mouse, dog, platypus, chicken, frog and zebrafish) according to Ensembl. Other amino acids are represented by green bubbles. The black triangles pinpoint the putative sites of N-linked glycosylation. Black squares represent amino acid changes linked to neutral variants according to our *in vitro* assays; orange squares represent amino acid changes linked to loss-of-function variants which are genetically null or deleterious according to our *in vitro* assays; blue squares represent amino acid changes linked to gain-of-function variants according to our *in vitro* assays. The numbers indicate the amino acid position of *OPRD1* coding variants.

**Fig. 2. *OPRD1* is expressed in human pancreatic beta cells, and DOP activation in human beta cells impairs glucose-stimulated insulin secretion.**



**(A) *OPRD1* is expressed in human pancreatic islets.** Expression of *OPRD1*, *OPRK1* and *OPRM1* in human pancreatic islets was assessed via RNA-seq done from 11 independent experiments (21,22), and expression of *Oprd1*, *Oprk1* and *Oprm1* in mouse pancreatic islets was assessed via RNA-seq done from 14 independent experiments (23,24). Data are the mean  $\pm$  SEM of TPM. **TPM**, transcripts per million.

**(B) *OPRD1* expression is significantly enriched in human pancreatic beta cells.** Expression of *OPRD1*, *OPRK1* and *OPRM1* was assessed in human purified pancreatic beta cells and beta-cell depleted islet preparations (*i.e.*, non-beta cells) (31). Data were directly uploaded from <ftp://jungle.unige.ch/BetaCellRNA-seq/>. Data are the mean  $\pm$  SEM of RPKM. **RPKM**, reads per kilobase million.

**(C) Overexpression of *OPRD1* in EndoC $\beta$ H1.** *OPRD1* expression in control EndoC $\beta$ H1 cells (CT; blue) and in EndoC $\beta$ H1 cells electroporated with *OPRD1* plasmid (OPRD1-OE; orange). *OPRD1* gene expression was measured by RT-qPCR. *RPLP0* was used as housekeeping gene. Data are the mean  $\pm$  SEM of Ct values from five independent experiments. A decrease in Ct values corresponds to an increase in gene expression. **Ct**, cycle threshold; \*\*\*\* $P < 0.0001$  (unpaired t-test).

**(D) DOP activation impairs glucose-stimulated insulin secretion from EndoC $\beta$ H1 cells.** Insulin secretion was assessed in control EndoC $\beta$ H1 cells (CT; blue) and in EndoC $\beta$ H1 cells overexpressing *OPRD1* (*OPRD1* OE; orange). Fold change in insulin secretion was analyzed in response to 60 min of incubation with 0.5 mM glucose  $\pm$  100 nM DII, followed by 60 min of incubation with 16.7 mM glucose  $\pm$  100 nM DII. Data are median (interquartile range) of five independent experiments. Insulin secretion (% of content or fold change) was analyzed using a linear regression adjusted for experimental conditions. Non-normalized insulin levels in each condition are available in **Supplementary Fig.6**. *DII*, [D-Ala<sup>2</sup>]-Deltorphan II; *Glc*, glucose.

**Table 1. Effect of rare, loss-of-function *OPRD1* variants on metabolic homeostasis in RaDiO and UK Biobank.**

Trait	Study	<i>n</i>	%, mean ± SD or median [IQR] of the trait in carriers	%, mean [± SD] or median [IQR] of the trait in non-carriers	OR [CI]	$\pi$ ± SE	<i>P</i>
Hyperglycemia or T2D	RaDiO	6,745	15%	47%	0.23 [0.033–0.95]	-	0.054
	UK Biobank	171,667	0%	3.4%	3.2×10 <sup>-5</sup> [2.0×10 <sup>-14</sup> –9.5×10 <sup>-4</sup> ]	-	0.13
Overweight or obesity	RaDiO	6,211	92%	58%	11 [2.1–194]	-	0.0054
	UK Biobank	187,242	85%	67%	2.6 [1.4–5.5]	-	0.0047
BMI (kg m <sup>-2</sup> )	RaDiO	6,971	27 [25–30]	26 [23–29]	-	0.35 ± 0.17	0.044
	UK Biobank	187,242	28 [26–31]	27 [24–30]	-	0.081 ± 0.027	0.0028
HDL* (mmol l <sup>-1</sup> )	RaDiO	6,008	1.3 ± 0.23	1.5 ± 0.45	-	-0.029 ± 0.16	0.86
	UK Biobank	164,358	1.5 ± 0.37	1.5 ± 0.38	-	0.10 ± 0.058	0.073
TG* (mmol l <sup>-1</sup> )	RaDiO	6,051	0.90 [0.81–1.1]	1.1 [0.76–1.6]	-	-0.13 ± 0.17	0.44
	UK Biobank	178,797	1.5 [1.1–1.9]	1.5 [1.0–2.1]	-	-0.0052 ± 0.027	0.85
SBP** (mmHg)	RaDiO	5,257	129 ± 20	130 ± 15	-	3.7 ± 6.4	0.56
	UK Biobank	180,474	136 ± 18	140 ± 20	-	1.9 ± 3.0	0.51
DBP** (mmHg)	RaDiO	4,693	77 ± 11	78 ± 9.4	-	1.4 ± 4.4	0.75
	UK Biobank	180,514	81 ± 8.2	82 ± 11	-	2.3 ± 2.0	0.25

In the RaDiO study, loss-of-function variants included the 12 rare variants that decreased the activity of DOP in response to two DOP agonists, and the rare null variant encoding p.E355\* (**Fig.1**). In UK Biobank, loss-of-function variants included 13 rare null variants and 6 rare variants that were found loss-of-function according to the present *in vitro* analyses (**Supplementary Table 4**). BMI and TG data were log-transformed before analysis. \*HDL and TG were only assessed in participants who were not under lipid-lowering agents in the RaDiO study. \*\*SBP and DBP were only assessed in participants who were not under high blood pressure medications in the RaDiO study. **BMI**, body mass index; **CI**, confidence interval; **DBP**, diastolic blood pressure; **HDL**, high-density lipoprotein; **IQR**, interquartile range;  $\pi$ , mean effect; **OR**, odds ratio; **SBP**, systolic blood pressure; **SD**, standard deviation; **SE**, standard error; **T2D**, type 2 diabetes; **TG**, triglycerides.

**Table 2. Effect of gain-of-function *OPRD1* variants on metabolic homeostasis in RaDiO and the AMP T2D knowledge portal.**

Trait	Study	<i>n</i>	%, mean ± SD or median [IQR] of the trait in carriers	%, mean [± SD] or median [IQR] of the trait in non-carriers	OR [CI]	Π or β ± SE	<i>P</i>
Hyperglycemia or T2D	RaDiO		50%	47%	0.89 [0.15–4.7]	-	0.89
	AMP-All	34,812	67%	57%	1.7 [1.3–2.2]	-	6.1×10 <sup>-5</sup>
	AMP-Africans	4,966	69%	54%	2.0 [1.5–2.6]	-	3.6×10 <sup>-6</sup>
BMI (kg m <sup>-2</sup> )	RaDiO		25 [24–27]	26 [23–29]	-	0.040 ± 0.22	0.86
	AMP-All	17,073	-	-	-	-0.027 ± 0.0093	0.0038
	AMP-Africans	1,988	-	-	-	-0.011 ± 0.013	0.41
HDL* (mmol l <sup>-1</sup> )	RaDiO		1.8 ± 0.39	1.5 ± 0.45	-	0.33 ± 0.17	0.049
	AMP-All	7,427	-	-	-	0.048 ± 0.046	0.30
	AMP-Africans	492	-	-	-	0.095 ± 0.049	0.055
TG* (mmol l <sup>-1</sup> )	RaDiO		0.55 [0.54–0.92]	1.1 [0.76–1.6]	-	-0.37 ± 0.18	0.046
	AMP-All	7,974	-	-	-	-0.13 ± 0.054	0.018
	AMP-Africans	492	-	-	-	-0.11 ± 0.060	0.069
LDL* (mmol l <sup>-1</sup> )	RaDiO	-	-	-	-	-	-
	AMP-All	5,288	-	-	-	-0.37 ± 0.12	0.0026
	AMP-Africans	489	-	-	-	-0.48 ± 0.16	0.0026
SBP** (mmHg)	RaDiO		120 ± 17	130 ± 15	-	-1.3 ± 8.3	0.88
	AMP-All	6,551	-	-	-	-0.91 ± 2.3	0.70
	AMP-Africans	493	-	-	-	-0.79 ± 3.3	0.81
DBP** (mmHg)	RaDiO		70 ± 11	78 ± 9.4	-	-2.3 ± 5.7	0.69
	AMP-All	6,552	-	-	-	2.6 ± 1.4	0.066
	AMP-Africans	493	-	-	-	2.5 ± 2.0	0.22

In the RaDiO study, gain-of-function variants included the 7 rare variants encoding p.P24S, p.A32G, p.I52V, p.I183V, p.K250R, p.T361N and p.P366L which increased the activity of DOP in response to two DOP agonists. In the AMP T2D knowledge portal, the gain-of-

function variant was the variant encoding p.I52V that increased the activity of DOP in response to two DOP agonists and that was common in Africans. \*To avoid possible biases due to treatment, these traits were only assessed in participants without type 2 diabetes in the AMP T2D knowledge portal or in participants who were not under lipid-lowering agents in the RaDiO study. \*\*To avoid possible biases due to treatment, these traits were only assessed in participants without type 2 diabetes in the AMP T2D knowledge portal or in participants who were not under high blood pressure medications in the RaDiO study.  $\beta$ , mean effect; **BMI**, body mass index; **CI**, confidence interval; **DBP**, diastolic blood pressure; **HDL**, high-density lipoprotein; **LDL**, low-density lipoprotein; **OR**, odds ratio;  $\pi$ , mean effect; **SBP**, systolic blood pressure; **SE**, standard error; **T2D**, type 2 diabetes; **TG**, triglycerides.

## Supplementary Methods

**Study participants.** In the RaDiO study, we analyzed up to 6,971 blood DNA samples (that were successfully sequenced) from several population studies (1): 1/ the D.E.S.I.R. 9-year prospective study including middle-aged men and women from western France (2); 2/ participants who were recruited and followed-up either by the UMR 1283/8199 unit (Lille, France), by the Department of Nutrition of Hotel-Dieu Hospital (Paris, France), or by the Centre d'Etude du Polymorphisme Humain (CEPH, Saint-Louis Hospital, Paris, France) (3,4); 3/ participants who were recruited from the Department of Endocrinology of the Corbeil-Essonnes Hospital (Corbeil-Essonnes, France) (3); 4/ the French Fleurbaix-Laventie Ville Santé study (5). All cohort studies followed ethical principles defined in the Helsinki declaration (revised in 1996), and they were approved by local ethical committees from Corbeil-Essonnes hospital (France), Comité Consultatif de Protection des Personnes se prêtant à des Recherches Biomédicales (CCPPRB) of Lille - Lille Hospital (Lille, France), Hotel-Dieu hospital (France), Bicêtre hospital (France). All participants signed an informed consent form. Normal weight was defined as  $BMI < 25 \text{ kg m}^{-2}$ , overweight as  $25 \leq BMI < 30 \text{ kg m}^{-2}$  and obesity as  $BMI \geq 30 \text{ kg m}^{-2}$ . Hyperglycemia was defined as fasting plasma glucose levels  $\geq 5.6 \text{ mmol l}^{-1}$  and/or use of drug therapy for hyperglycemia (6). HDL cholesterol and triglycerides were only assessed in participants who were not under lipid-lowering agents. Systolic blood pressure and diastolic blood pressure were only assessed in participants who were not under high blood pressure medications.

In UK Biobank, we analyzed up to 187,242 samples, with available exome sequencing data. This research is part of UK Biobank research application #67575. The participants were recruited by UK Biobank from across the United Kingdom between 13 March 2006 and 1 October 2010. Normal weight was defined as  $BMI < 25 \text{ kg m}^{-2}$ , overweight as  $25 \leq BMI < 30 \text{ kg m}^{-2}$  and obesity as  $BMI \geq 30 \text{ kg m}^{-2}$ .

People with T2D in UK Biobank were defined as follows:

Inclusion criteria: 'HbA1c  $\geq 48 \text{ mmol mol}^{-1}$ ' (field #30750) or 'use of insulin' (fields #6153 or #6177, code [3]) or 'diabetes diagnosed by doctor' (field #2443)

Exclusion criteria: 'at least one International Classification of Diseases (ICD)-10 code starting with E10' (type 1 diabetes), 'at least one ICD-10 code starting with O24' (diabetes mellitus in pregnancy).

Controls for the T2D case-control study in UK Biobank were defined as follows:

Inclusion criteria: 'HbA1c  $< 42 \text{ mmol mol}^{-1}$ ' (field #30750) and 'age  $\geq 40$  years'

Exclusion criteria: 'at least one ICD-10 code starting with E10' (type 1 diabetes), 'at least one ICD-10 code starting with E11' (non-insulin-dependent diabetes mellitus), 'at least one ICD-10 code starting with E11' (non-insulin-dependent diabetes mellitus), 'at least one ICD-10 code starting with E13' (other specified diabetes mellitus), 'at least one ICD-10 code starting with R73' (elevated blood glucose level), 'at least one ICD-10 code starting with O24' (diabetes mellitus in pregnancy), 'gestational diabetes only' (field #4041), 'diabetes diagnosed by doctor' (field #2443), 'use of insulin' (fields #6153 or #6177, code [3]).

**Plasmids.** Wild-type *OPRD1* plasmid, a negative control *OPRD1* plasmid ('STOP'; c.3\_4insTAA/p.E2\*), and three plasmids containing respectively the rare *OPRD1* variants encoding p.F220L, p.F280L, and p.G365D were purchased from e-Zyvec (Loos, France). The plasmids including each of the remaining 31 coding variants were generated using the QuickChange site-directed mutagenesis kit (Stratagene, San Diego, USA). The sequence of each plasmid was verified by Sanger sequencing.

**Luciferase assays.** Using FuGENE HD (Promega, Madison, USA), HEK293 cells were transfected with three plasmids:  $375 \text{ ng ml}^{-1}$  of *OPRD1* plasmid (wild-type or with a variant),

225 ng ml<sup>-1</sup> of plasmid including the firefly luciferase gene under the control of Gα<sub>i</sub>-dependent serum response element (SRE), and 150 ng ml<sup>-1</sup> of plasmid including the β-galactosidase gene. 100 μl of these cells (1×10<sup>6</sup> cells ml<sup>-1</sup>) were then seeded in a poly-L-lysine coated 96-well plate. The day after, the medium was replaced by a starving medium containing 0.1% fetal bovine serum. Cells were starved overnight. Two days after transfection, cells were treated for 5 h with increasing concentrations (1, 10 and 100 nM) of either [D-Ala<sup>2</sup>]-Deltorphin II (DII) (Tocris Bioscience, Bristol, UK) or [D-Pen<sup>2</sup>,D-Pen<sup>5</sup>]-Enkephalin (DPDPE) (Tocris Bioscience) which are both DOP agonists. Then, cells were lysed in 100 μl of luciferase cell culture lysis Buffer (Promega) for 30 minutes. Luciferase activity was measured by adding 25 μl of luciferase assay system reagent (Promega) to 40 μl of cell lysate. Luminescence was read with a GloMax Luminometer (Promega). Simultaneously, β-galactosidase activity was measured by adding 100 μl of β-galactosidase buffer (NaHPO<sub>4</sub> 0.479 g, NAH<sub>2</sub>PO<sub>4</sub> – 2 H<sub>2</sub>O 1.060 g, KCl 2M 0.5 ml, MgCl<sub>2</sub> 1M 0.1 ml, Q.S. to 1 ml water; extemporaneous addition of β-mercaptoethanol [3.25 μl ml<sup>-1</sup>] and ortho-nitrophenyl-β-galactoside 4 mM [250 μl ml<sup>-1</sup>]) to 40 μl of cell lysate. Absorbance was read at 450 nm. The experiments were performed in technical triplicate and each experiment was repeated at least four times. Luciferase measurements in relative luminescence units were normalized using β-galactosidase measures. Fold change (FC) was computed by dividing the normalized luciferase activity by the mean of the baseline luciferase activity (*i.e.* under dimethyl sulfoxide [DMSO] 0.1%). The effect of each *OPRD1* variant (or wild-type) was analyzed using the following ANOVA model:  $FC = \beta_0 + \beta_1C + \delta_1V + \delta_2C \times V + \beta_2E + \epsilon$ , where *C* was the concentration of DII or DPDPE (as a categorical variable), *V* was the *OPRD1* variant (or wild-type) and *E* was the experiment ID.

**Immunofluorescence assays.** HEK293 cells were transfected in suspension using FuGENE HD (Promega) with 375 ng ml<sup>-1</sup> of *OPRD1* plasmid (wild-type or with a variant). 2 ml of these cells (5×10<sup>5</sup> cells ml<sup>-1</sup>) were then seeded in a 6-well plate containing poly-L-lysine coated coverslips. Two days after transfection, cells were fixed in paraformaldehyde 4% (Thermo Fisher Scientific) for 20 minutes, washed three times with phosphate buffer saline (PBS), and permeabilized in PBS – Triton X-100 0.1% for 20 minutes. The unoccupied binding surfaces on the membrane were blocked in a blocking buffer (PBS, donkey serum 5% (Millipore / Sigma-Aldrich, St. Louis, USA), Triton X-100 0.1%) for 1 h. Cells were then incubated with the anti-DOP primary antibody (ab176324 [Abcam, Cambridge, United Kingdom]; diluted at 1/100 in blocking buffer) overnight at 4 °C. After three washes with PBS – Triton X-100 0.1%, cells were incubated with a fluorescent secondary antibody (A-21206 [Thermo Fisher]; diluted at 1/500 in blocking buffer) for 2 h at 4 °C in the dark. After three washes with PBS – Triton X-100 0.1%, F-actin was labelled with fluorescent phalloidin (#21836 [Thermo Fisher Scientific]; diluted at 1/300 in blocking buffer), and nuclei were labelled with DAPI (R-37606 [Thermo Fisher Scientific]) for 30 min at room temperature in the dark. Cells were then washed twice with PBS – Triton X-100 0.1%, and slides were mounted with Prolong Gold Antifade Mountant (Thermo Fisher Scientific). Cells were observed with a Cell Observer Spinning Disk microscope (Zeiss, Oberkochen, Germany).

**Western blot assays.** HEK293 cells were transfected in suspension using FuGENE HD (Promega), with 375 ng ml<sup>-1</sup> of *OPRD1* plasmid (wild-type or with a variant). 2 ml of these cells (1×10<sup>6</sup> cells ml<sup>-1</sup>) were then seeded in a poly-L-lysine coated 6-well plate. Two days after transfection, proteins were extracted in Pierce RIPA buffer (Thermo Fisher Scientific), supplemented with dithiothreitol 1 mM (Invitrogen / Thermo Fisher Scientific), protease and phosphatase inhibitors (Roche Molecular Diagnostics, Pleasanton, USA). Protein concentration was measured via the Pierce BCA protein assay kit (Thermo Fisher Scientific), according to manufacturer's instructions. 40 μg of proteins were denatured in Laemmli Buffer 1× (Alfa Aesar, Haverhill, USA) for 5 minutes at 95 °C, loaded for migration on a concentration gel

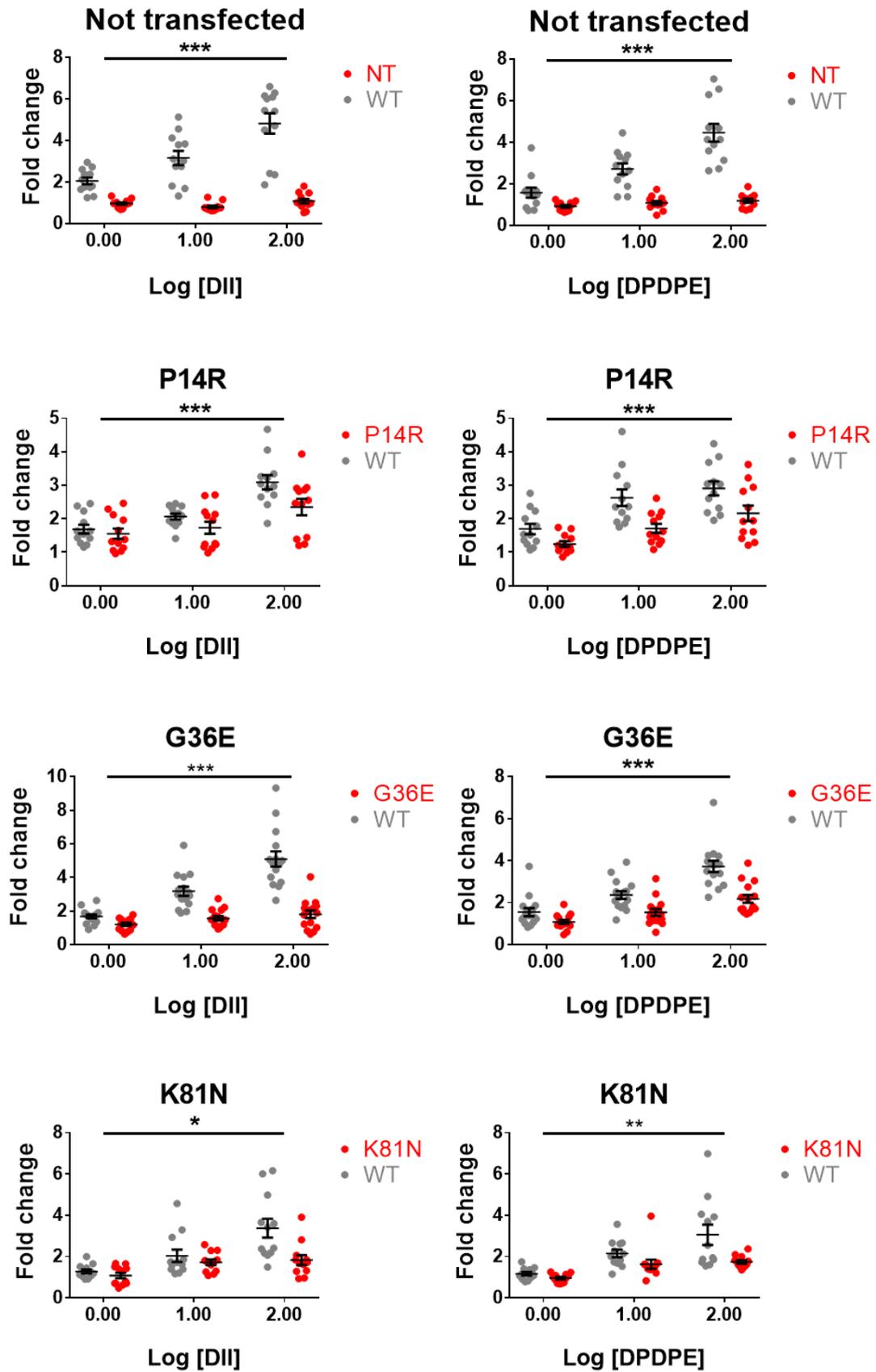
(Acrylamide 4% [Dutscher, Brumath, France], Tris 125 mM [Dutscher], Sodium Dodecyl Sulfate [SDS] 0.04% [Sigma-Aldrich], Ammonium Persulfate 0.08% [Sigma-Aldrich], Temed 0.004% [Dutscher]) followed by a migration on a 12% separation gel (Acrylamide 10%, Tris 375 mM, SDS 0.2%, Ammonium Persulfate 0.1%, Temed 0.05% [Dutscher]) in Tris-Glycine-SDS buffer (Eurobio Scientific, Les Ulis, France). Proteins were then transferred on a nitrocellulose membrane (GE Healthcare, Chicago, USA) in Tris-Glycine buffer (Eurobio Scientific). Ponceau Red (Sigma-Aldrich) was used to confirm the presence of the proteins on the membrane. The unoccupied binding surfaces on the membrane were blocked in a blocking buffer (Tris buffer saline [TBS] 1× [Euromedex, Souffelweyersheim, France], Tween 20 0.1% [Biosolve, Dieuze, France] and skimmed milk 5% [Dutscher]) for 1 h, and then incubated overnight under agitation with the anti-DOP antibody (ab176324 [Abcam]; diluted at 1/1000 in the blocking buffer). After the incubation, the membrane was washed three times with TBS – Tween 20 0.1%, and incubated with a fluorescent anti-rabbit antibody (SA5-35571 [Invitrogen / Thermo Fisher Scientific]; diluted at 1/5000 in the blocking buffer). The membrane was then washed three times with TBS – Tween 20 0.1%, and the protein of interest was revealed via the Odyssey CLx imaging system (LI-COR [Lincoln, USA]; reading at 800 nm). To normalize our data,  $\beta$ -actin was used as a loading control, and was revealed as described above (anti- $\beta$ -actin primary antibody: #3700 [Cell Signaling Technology, Leiden, The Netherlands] diluted at 1/5000 in the blocking buffer; fluorescent anti-mouse secondary antibody: #35518 [Thermo Fisher Scientific] diluted at 1/5000 in the blocking buffer or HRP anti-mouse secondary antibody: #NA931 [Merck / Sigma-Aldrich] diluted at 1/5000 in the blocking buffer; revelation via the Odyssey CLx imaging system at 700 nm). At least, three independent experiments for each mutant were performed. Two-by-two expression comparisons of each DOP mutant *versus* wild-type were performed using a Wilcoxon rank sum test.

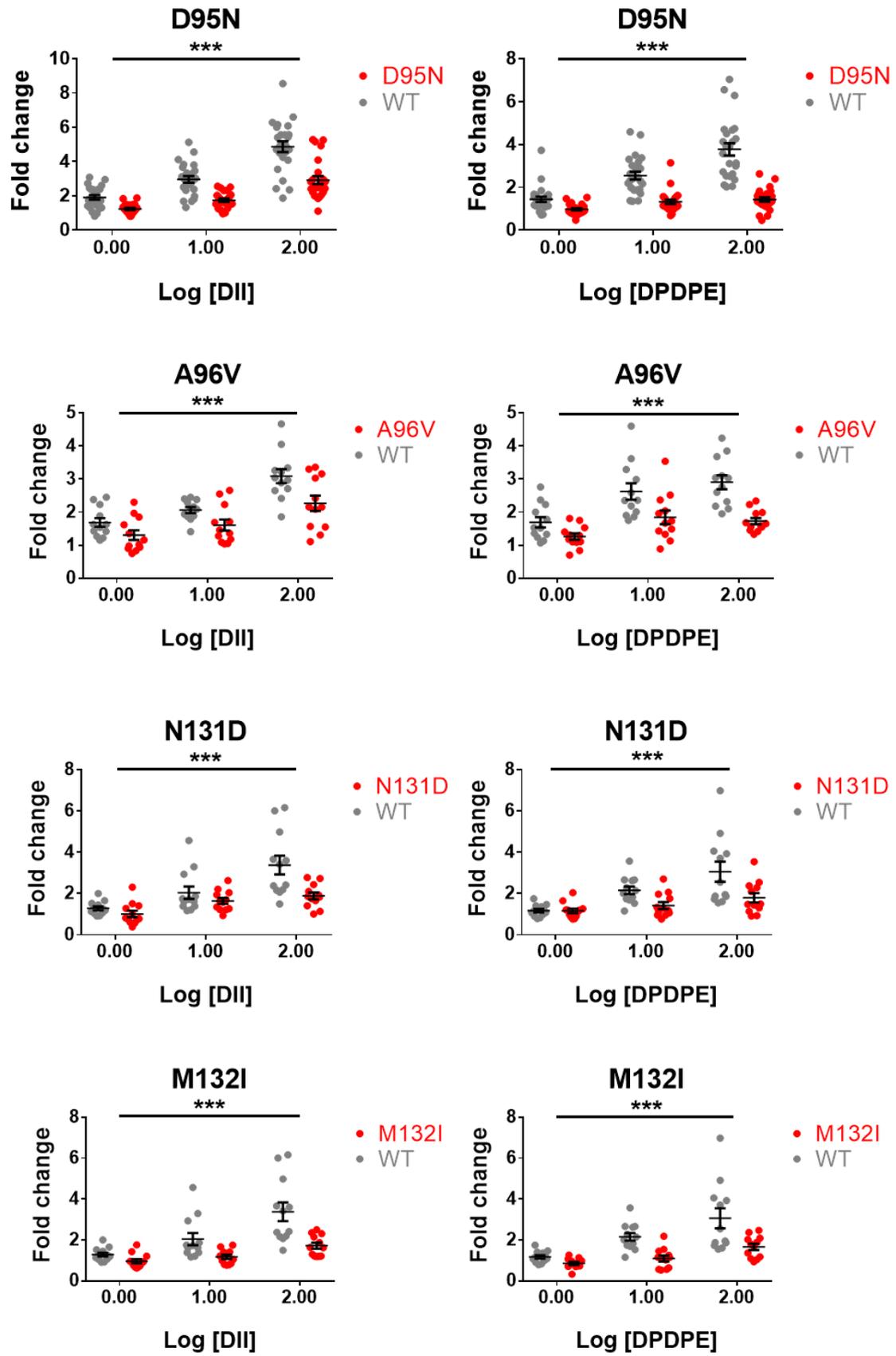
***Electroporation of EndoC $\beta$ H1 cells.*** EndoC $\beta$ H1 cells were electroporated using the Cell Line Nucleofactor Kit V (Lonza, Basel, Switzerland). Briefly,  $2.5 \times 10^6$  EndoC $\beta$ H1 cells were placed in suspension in 100  $\mu$ l of electroporation solution  $\pm$  3.125  $\mu$ l of wild-type *OPRD1* plasmid (at a concentration of 0.2  $\mu$ g  $\mu$ l<sup>-1</sup>). The electroporation was done on the Nucleofactor 2b device (Lonza) using the G-016 program, and then resuspended in 2 ml of Opti $\beta$ 1 medium. 100  $\mu$ l of cells, at a concentration of  $1 \times 10^6$  cells ml<sup>-1</sup>, were then seeded in triplicate in wells of a 96-well plate primarily coated with  $\beta$ coat. To assess the efficiency of the electroporation of *OPRD1* plasmid, 1.5 ml of cells, at a concentration of  $1 \times 10^6$  cells ml<sup>-1</sup> were seeded in a 6-well plate coated with  $\beta$ coat, Q.S. to 2 ml Opti $\beta$ 1 medium. 6 h after the electroporation, the medium in the 96-well plate was replaced with 100  $\mu$ l of medium including 50  $\mu$ l of fresh Opti $\beta$ 1 medium and 50  $\mu$ l of conditioned medium, and with 2 ml of this medium in the 6-well plate. The day after, this medium was replaced by the same amount of fresh Opti $\beta$ 1 medium. At three days after electroporation, cells seeded in the 6-well plate were used to assess the efficiency of the electroporation. For this purpose, RNA was extracted using TRIzol reagent (Thermo Fisher Scientific) and was retrotranscribed using the High Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific). Quantitative PCR was then done using the Brilliant III Ultra-Fast SybrGreen QPCR Master mix with low ROX (Agilent Technologies, Santa Clara, USA), on the Viia 7 Real-Time PCR System (Thermo Fisher Scientific). We used *RPLP0* as housekeeping gene.

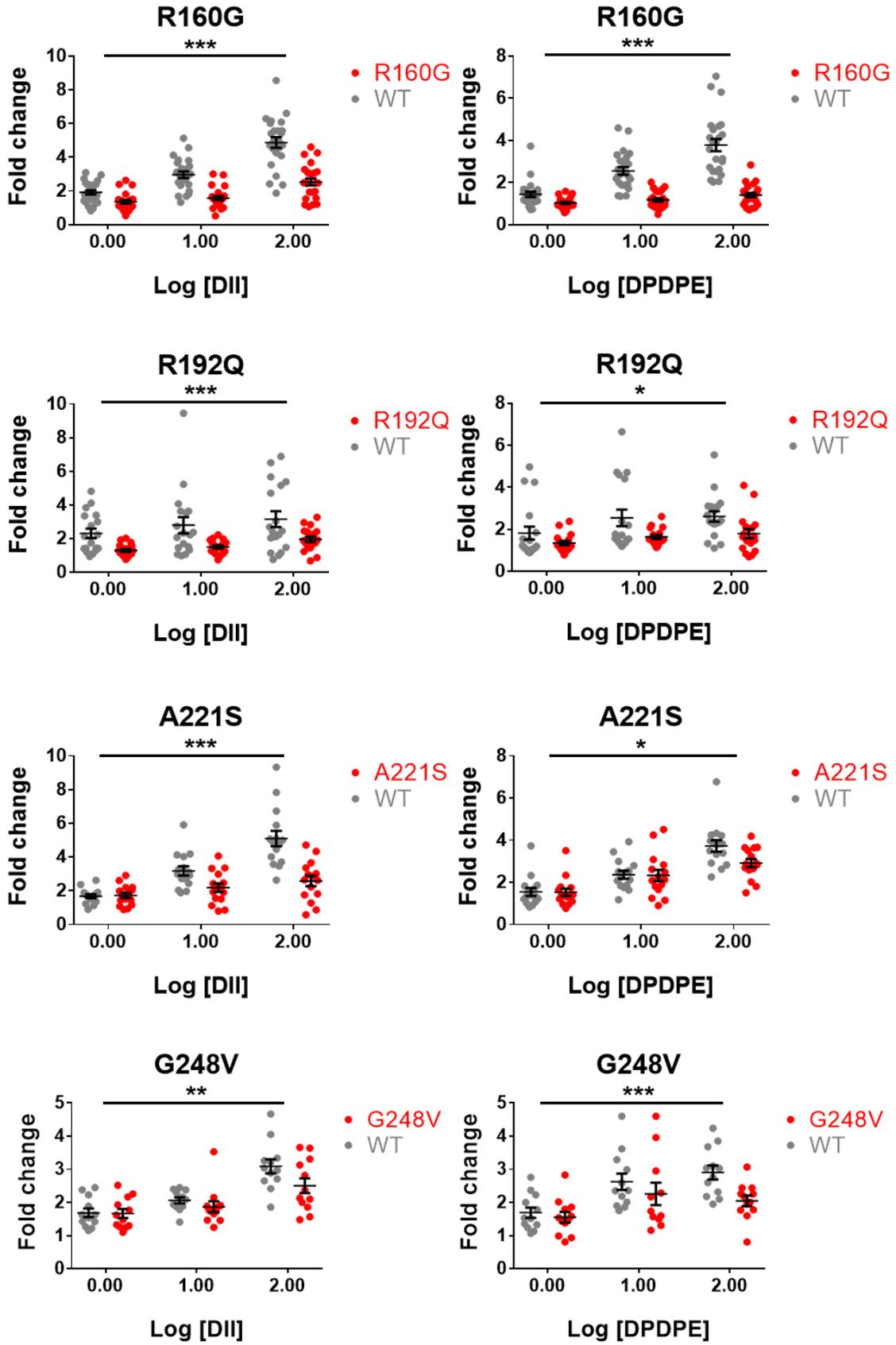
***Insulin secretion assays in EndoC $\beta$ H1 cells.*** At two days after electroporation, EndoC $\beta$ H1 cells seeded in the 96-well plate were incubated overnight in a starving medium, and were rinsed at three days after electroporation. We subsequently performed insulin secretion assays as previously described (7). Briefly, EndoC $\beta$ H1 cells were incubated with the  $\beta$ Krebs Buffer (Univercell Biosolution) containing 0.5 mM glucose  $\pm$  100 nM DII (Tocris Bioscience) for 60 minutes at 37 °C. The supernatant (called supernatant #1) was collected and was replaced by

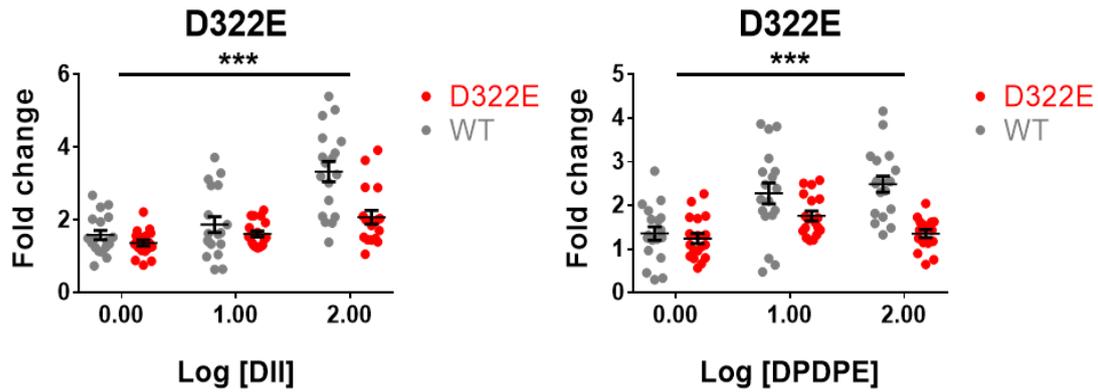
the  $\beta$ Krebs Buffer containing 16.7 mM glucose  $\pm$  100 nM DII for 60 minutes at 37 °C. The supernatant (called supernatant #2) was collected. Cells were then lysed and the lysates were collected. Supernatants and lysates were centrifuged at 700 g for 5 minutes at 4 °C and insulin was assessed after proper dilution of the samples in water (1:12 for supernatants and 1:500 for lysates), using an ELISA Human Insulin kit (Merckodia, Uppsala, Sweden). We then used the R package called INsulin Secretion ANalysEr (Insane; R package version 0.1.0), which performs critical quality control (QC) steps to monitor experimental stages in addition to statistical analyses, adjusting for technical and confounding variables. These critical QC steps were previously described (7). The threshold to define accurate insulin secretion in cells was set at 1.5. Experiments in which cells did not secrete insulin, or in which the blank value was higher than 1.5 times the interquartile range above the 75<sup>th</sup> percentile or lower than 1.5 times the interquartile range below the 25<sup>th</sup> percentile, were excluded. Outliers values, corresponding to an optical density higher than 1.5 times the interquartile range above the 75<sup>th</sup> percentile or lower than 1.5 times the interquartile range below the 25<sup>th</sup> percentile were excluded. The fold change in insulin secretion for cells overexpressing (or not) *OPRDI* was analyzed using linear regression adjusted for experimental conditions, from five independent experiments.

Supplementary Fig. 1. Activity of loss-of-function *OPRD1* variants in response to two DOP agonists, according to luciferase assays



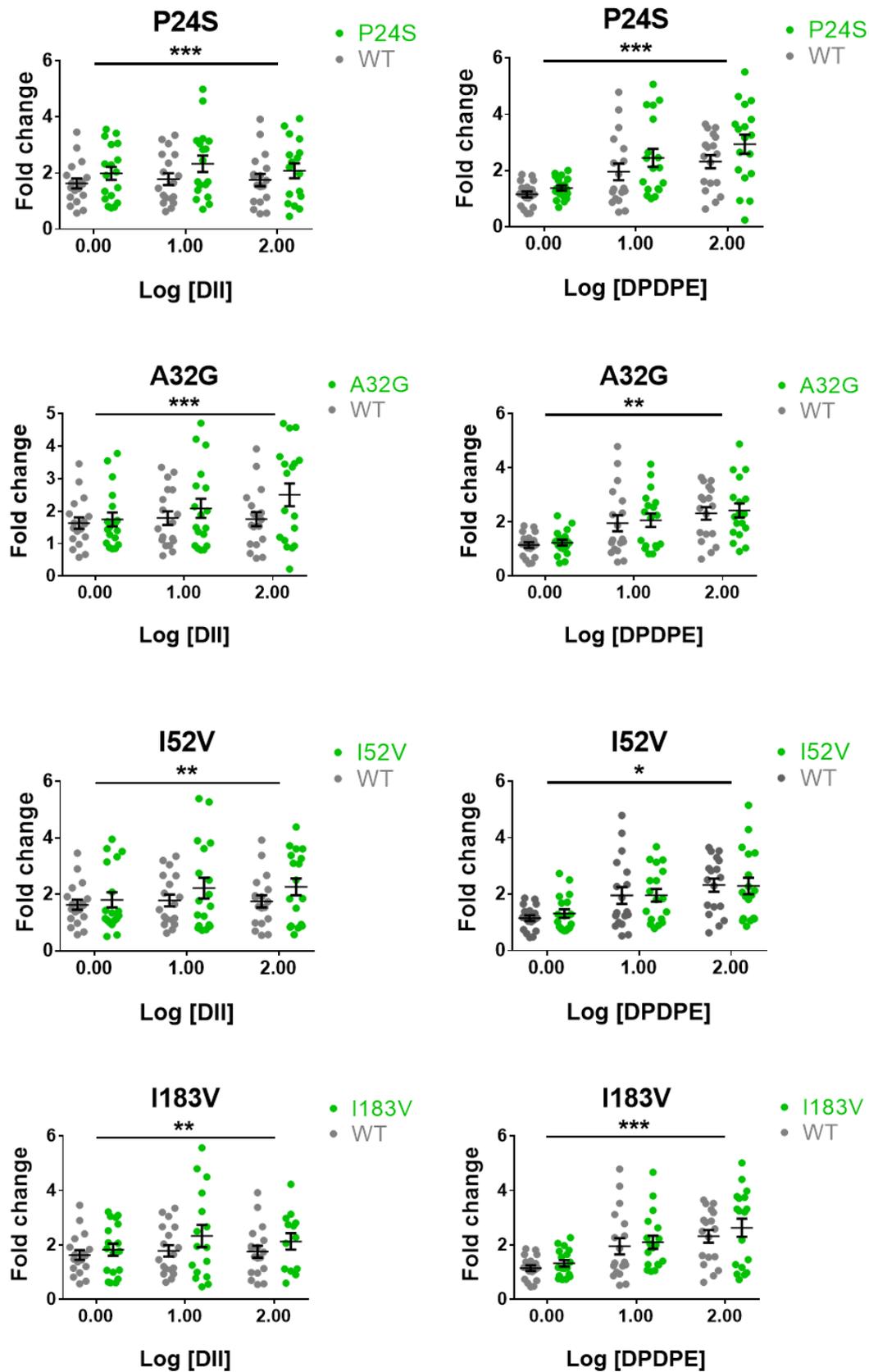


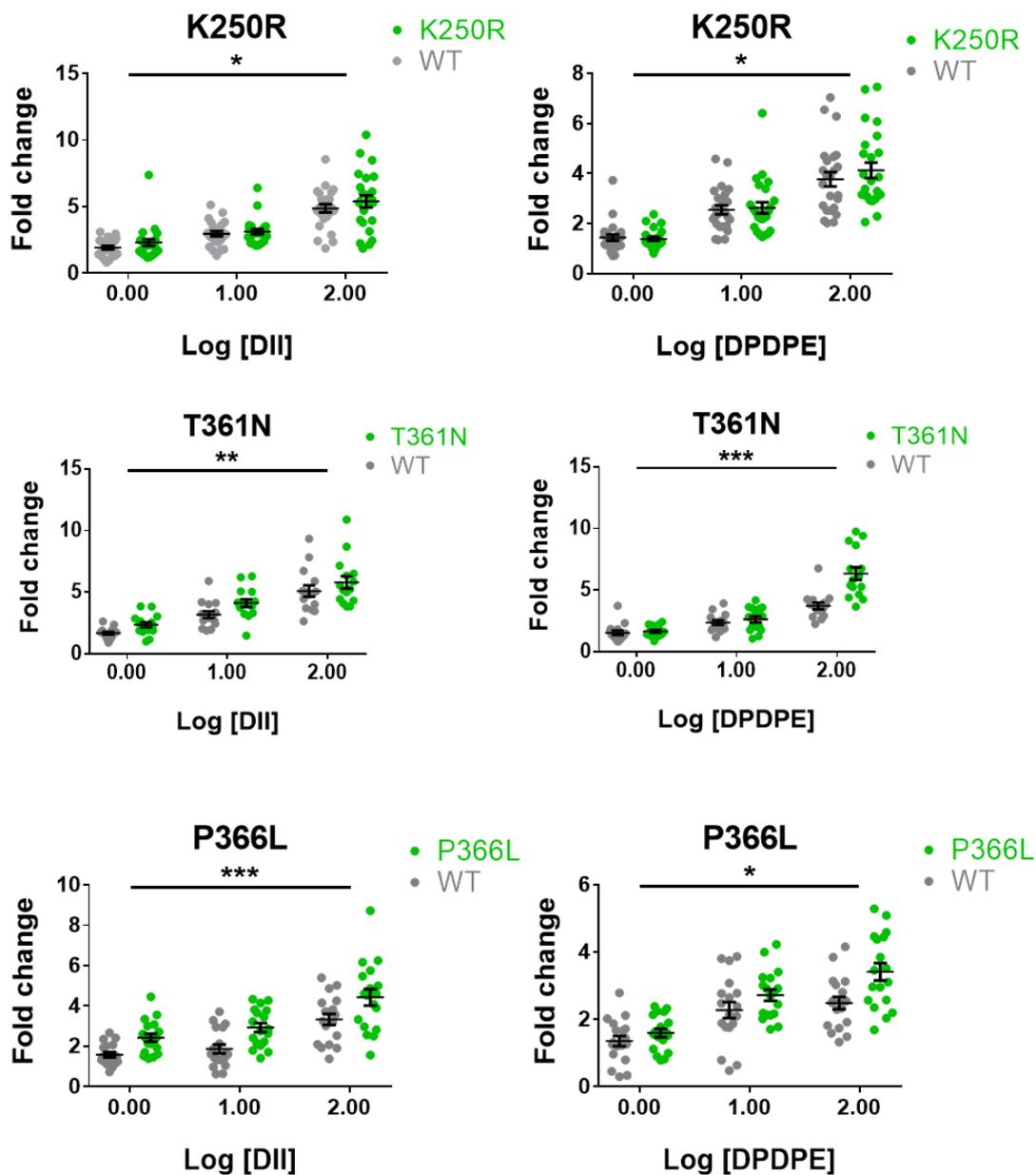




The figures show fold changes (FCs) of SRE luciferase activity normalized with  $\beta$ -galactosidase in HEK293 cells cotransfected or not (*i.e.* ‘not transfected’ condition) with wild-type or mutant *OPRD1* plasmid, in response to 1, 10 and 100 nM DII or DPDPE, relative to the baseline activity of WT DOP. Data are the mean  $\pm$  SEM of the FCs from at least four independent experiments in technical triplicate. The effect of each *OPRD1* variant was analyzed using the following ANOVA model:  $FC = \beta_0 + \beta_1C + \delta_1V + \delta_2C \times V + \beta_2E + \epsilon$ , where  $C$  is the concentration of DII or DPDPE (as a categorical variable),  $V$  is the *OPRD1* variant and  $E$  is the experiment ID. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$  (*versus* wild-type). **DII**, [D-Ala2]-Deltorphan II; **DPDPE**, [D-Pen2,D-Pen5]-Enkephalin; **NT**, not transfected; **WT**, wild-type.

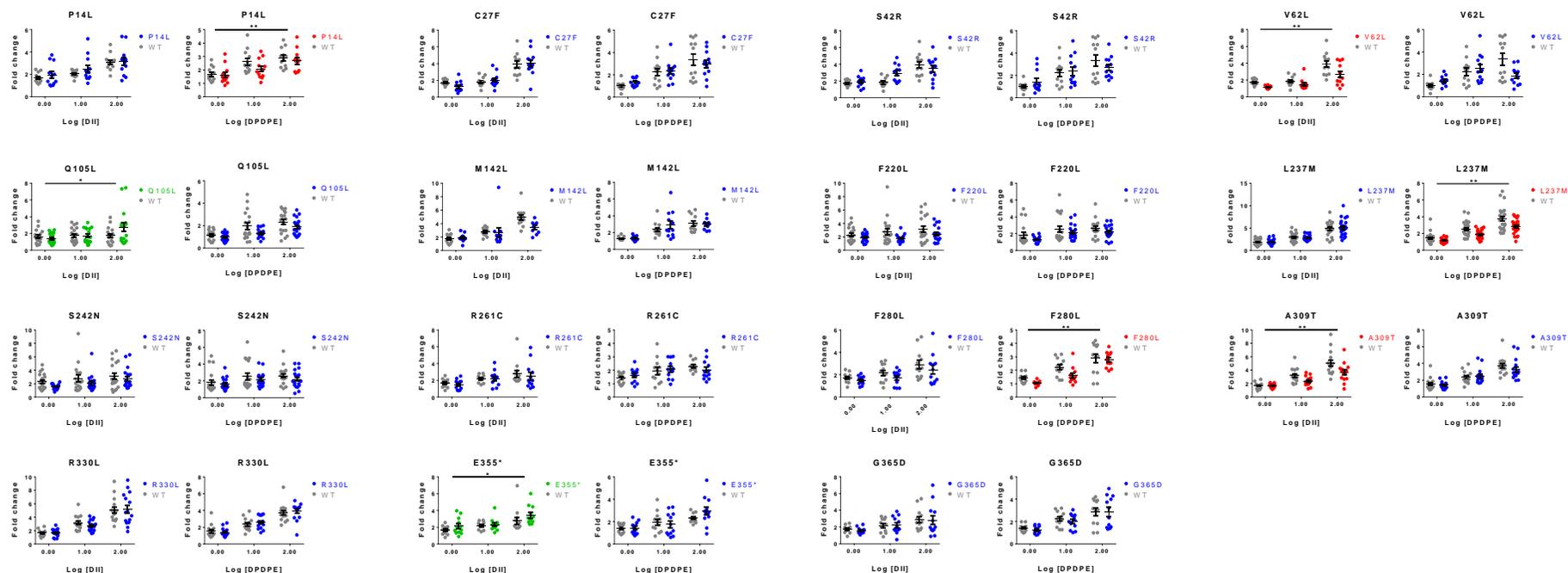
Supplementary Fig. 2. Activity of gain-of-function *OPRD1* variants in response to two DOP agonists, according to luciferase assays





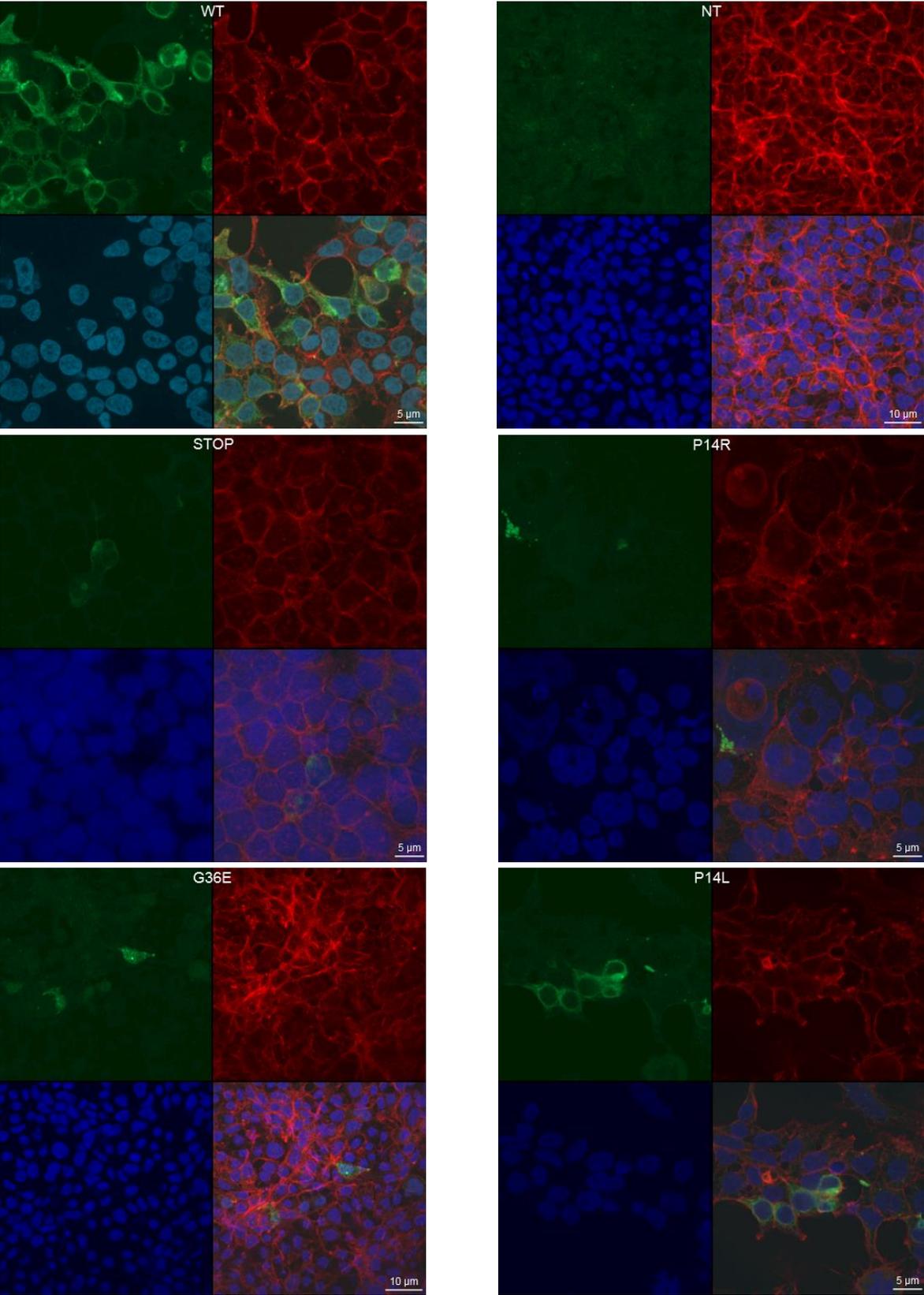
The figures show fold changes (FCs) of SRE luciferase activity normalized with  $\beta$ -galactosidase in HEK293 cells cotransfected with wild-type or mutant *OPRD1* plasmid, in response to 1, 10 and 100 nM DII or DPDPE, relative to the baseline activity of WT DOP. Data are the mean  $\pm$  SEM of the FCs from at least four independent experiments in technical triplicate. The effect of each *OPRD1* variant was analyzed using the following ANOVA model:  $FC = \beta_0 + \beta_1 C + \delta_1 V + \delta_2 C \times V + \beta_2 E + \epsilon$ , where  $C$  is the concentration of DII or DPDPE (as a categorical variable),  $V$  is the *OPRD1* variant and  $E$  is the experiment ID. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$  (versus wild-type). *DII*, [D-Ala<sup>2</sup>]-Deltorphan II; *DPDPE*, [D-Pen<sup>2</sup>,D-Pen<sup>5</sup>]-Enkephalin; *WT*, wild-type.

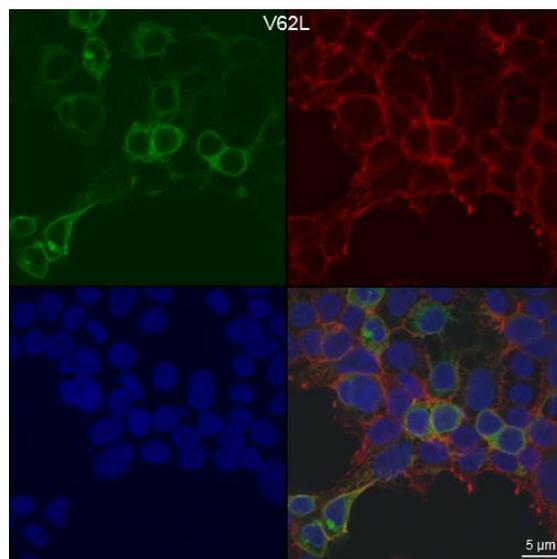
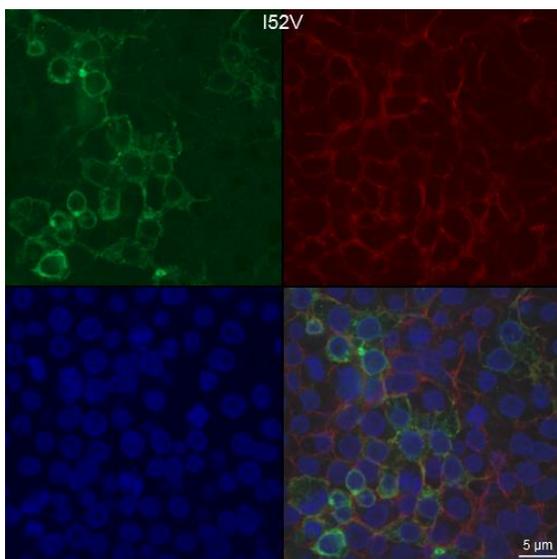
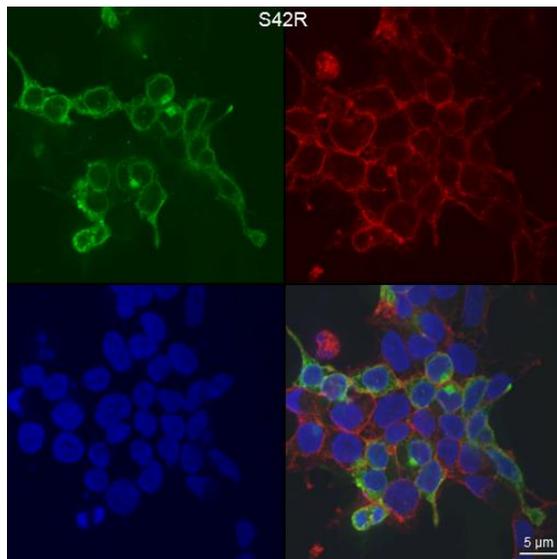
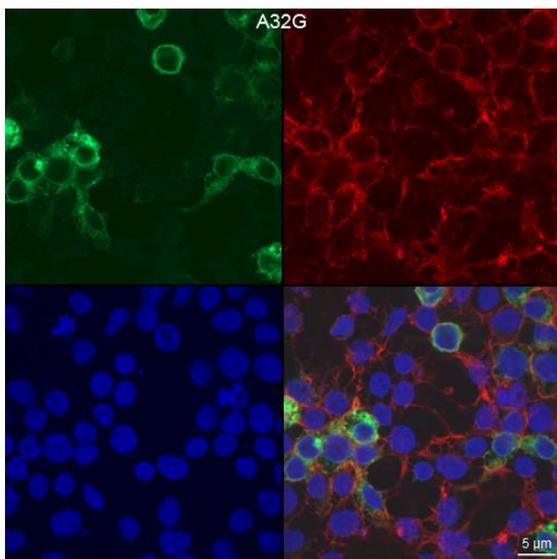
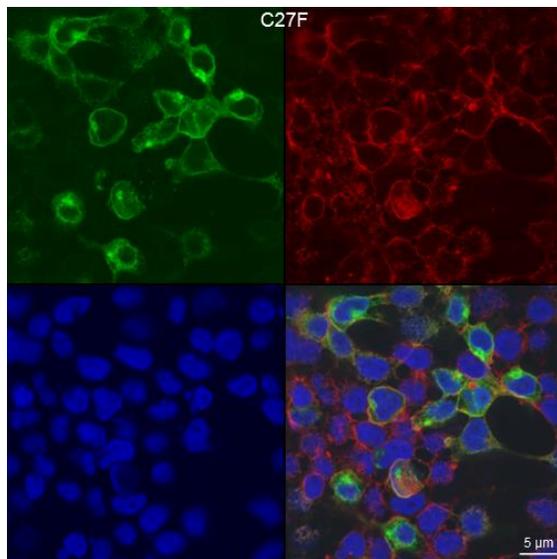
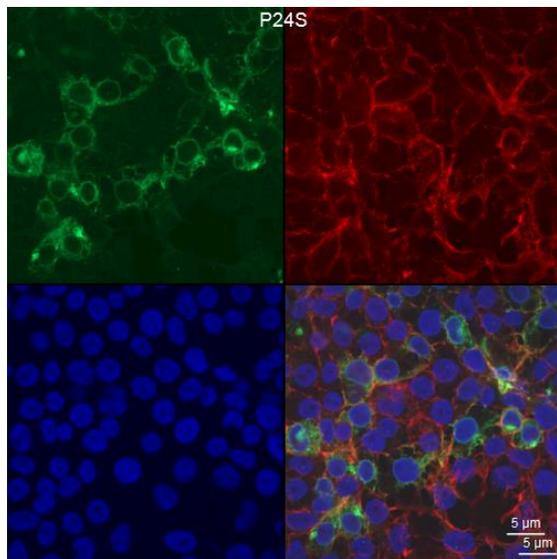
### Supplementary Fig. 3. Activity of neutral *OPRD1* variants in response to two DOP agonists, according to luciferase assays

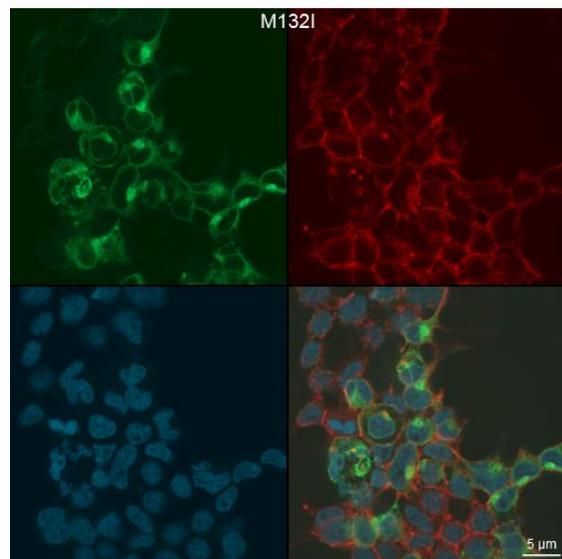
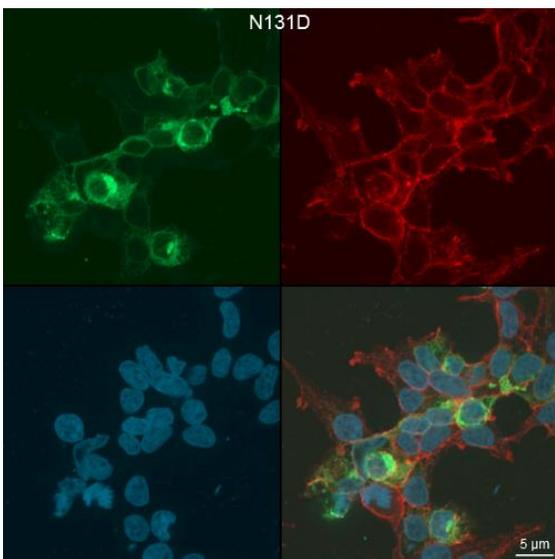
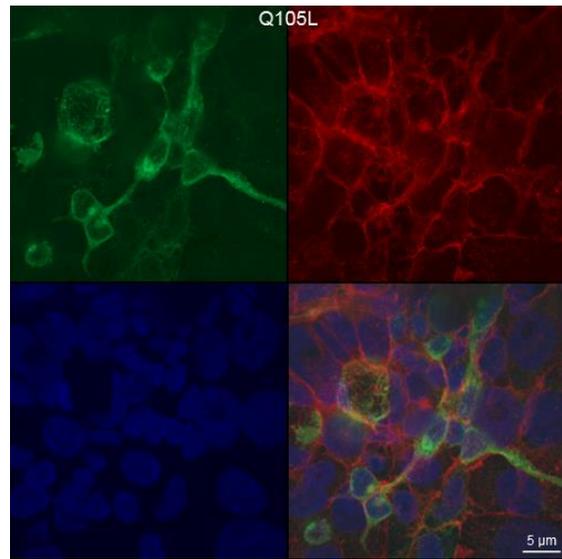
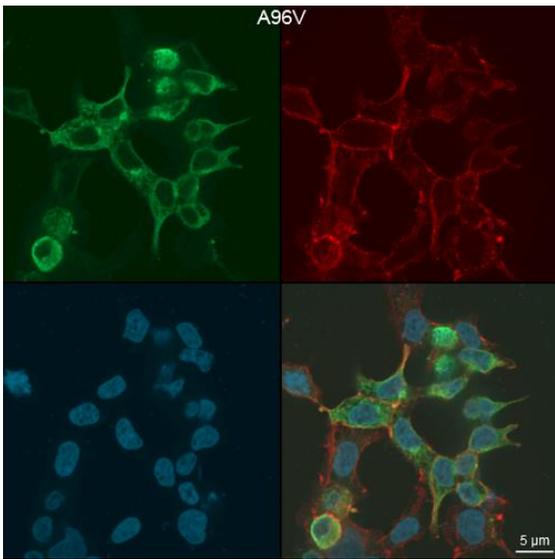
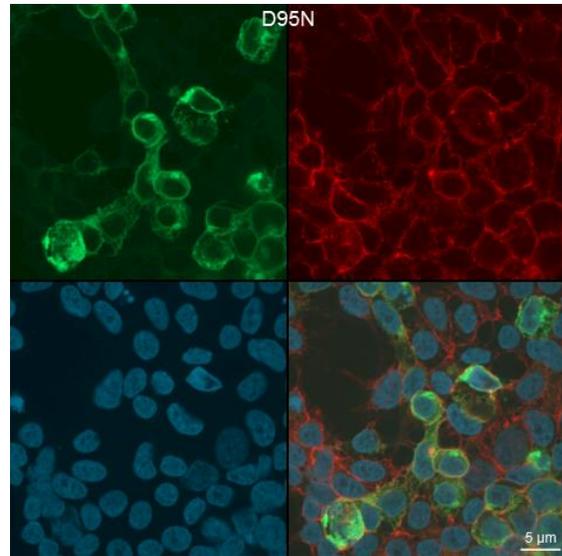
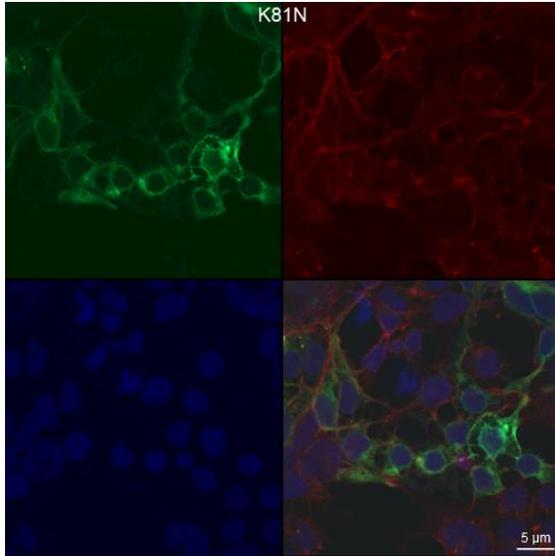


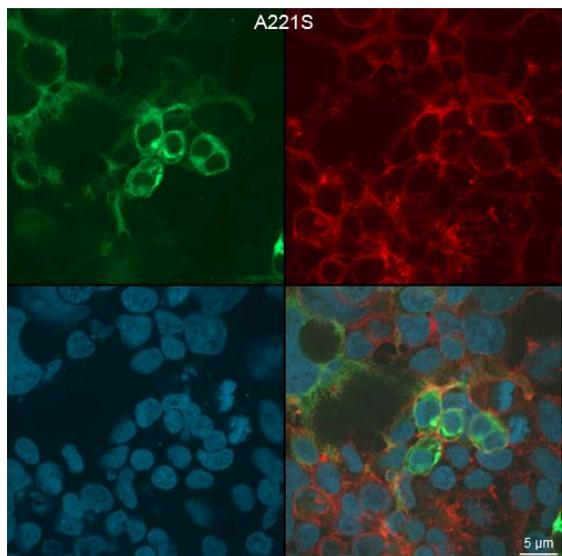
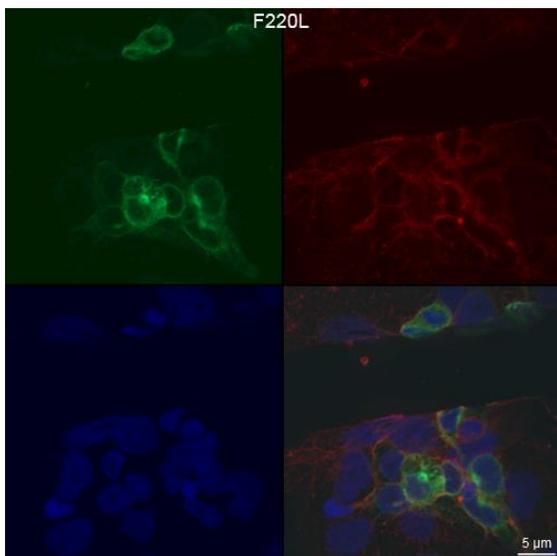
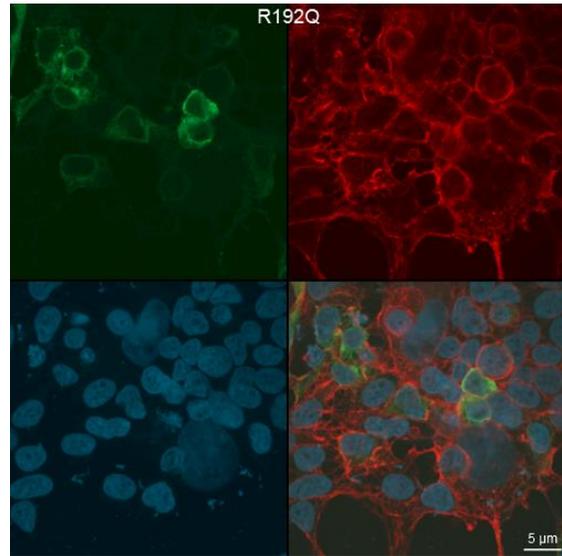
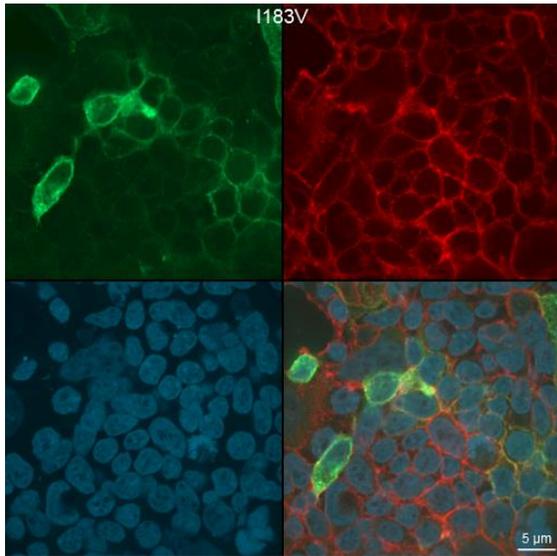
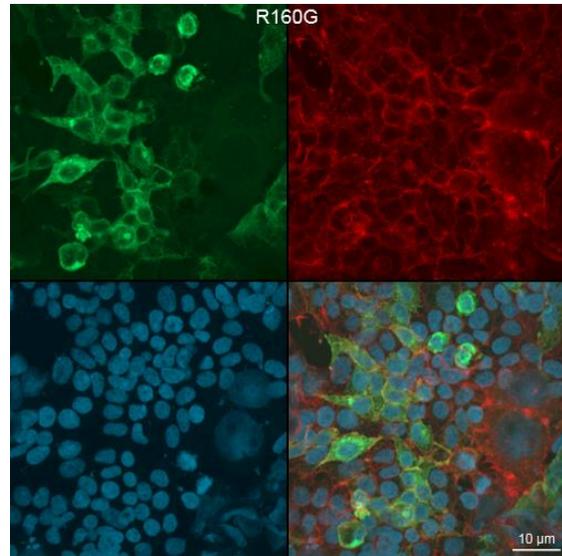
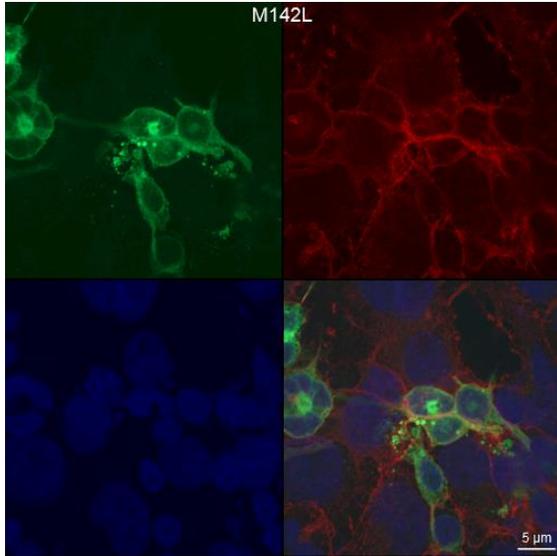
The figures show fold changes (FCs) of SRE luciferase activity normalized with  $\beta$ -galactosidase in HEK293 cells cotransfected with wild-type or mutant *OPRD1* plasmid, in response to 1, 10 and 100 nM DII or DPDPE, relative to the baseline activity of WT DOP. Data are the mean  $\pm$  SEM of the FCs from at least four independent experiments in technical triplicate. The effect of each *OPRD1* variant was analyzed using the following ANOVA model:  $FC = \beta_0 + \beta_1 C + \delta_1 V + \delta_2 C \times V + \beta_2 E + \epsilon$ , where  $C$  is the concentration of DII or DPDPE (as a categorical variable),  $V$  is the *OPRD1* variant and  $E$  is the experiment ID. \* $P < 0.05$ ; \*\* $P < 0.01$  (versus wild-type). **DII**, [D-Ala2]-Deltorphan II; **DPDPE**, [D-Pen2,D-Pen5]-Enkephalin; **WT**, wild-type.

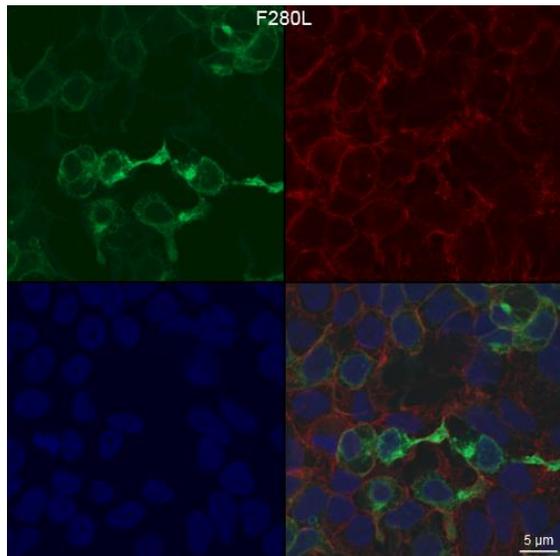
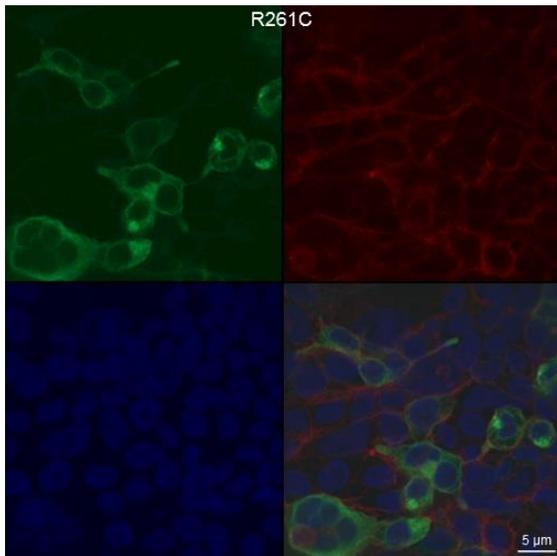
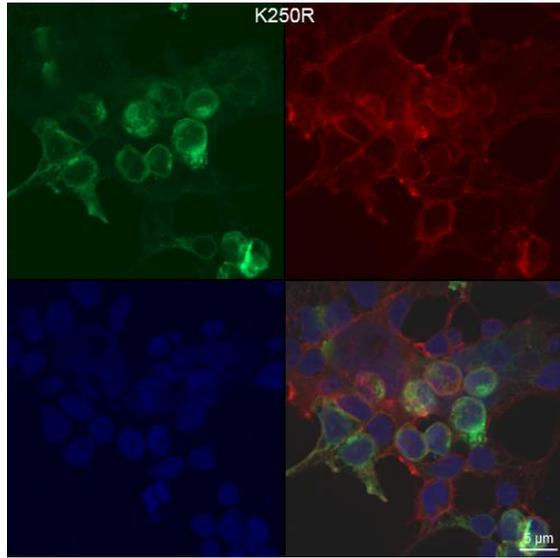
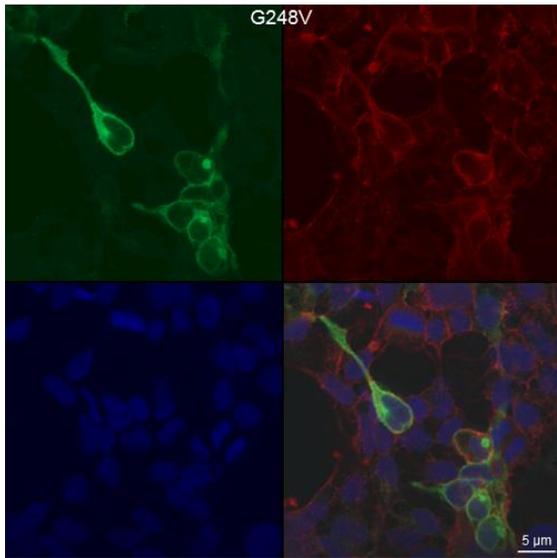
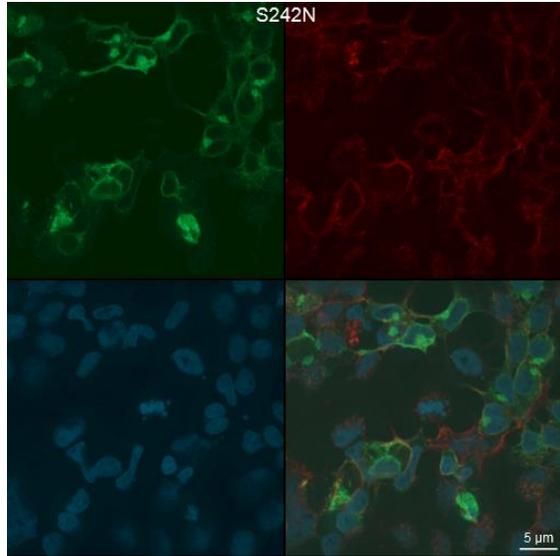
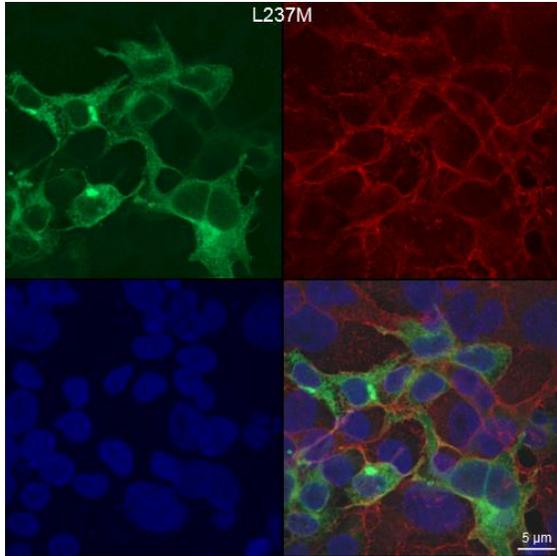
**Supplementary Fig. 4. Cellular localization of the DOP mutants**

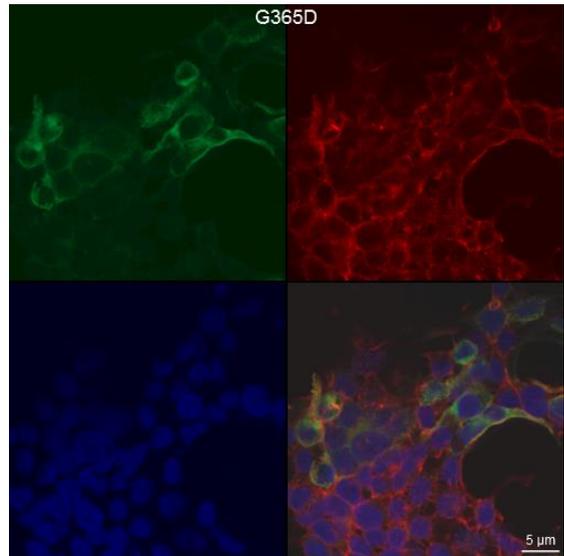
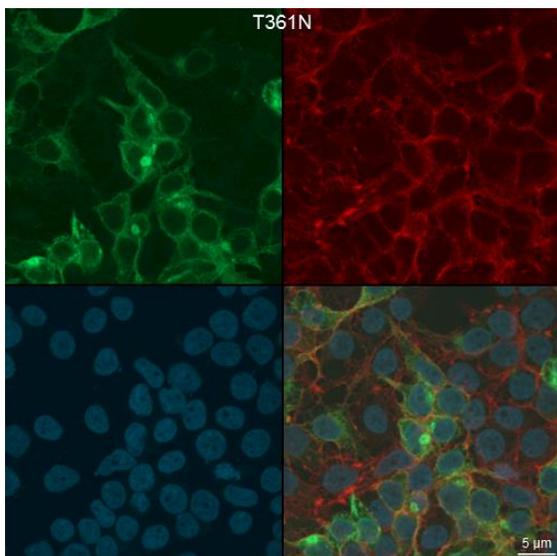
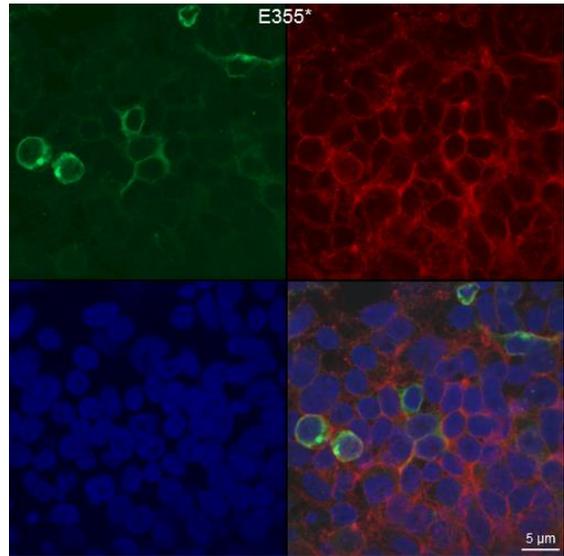
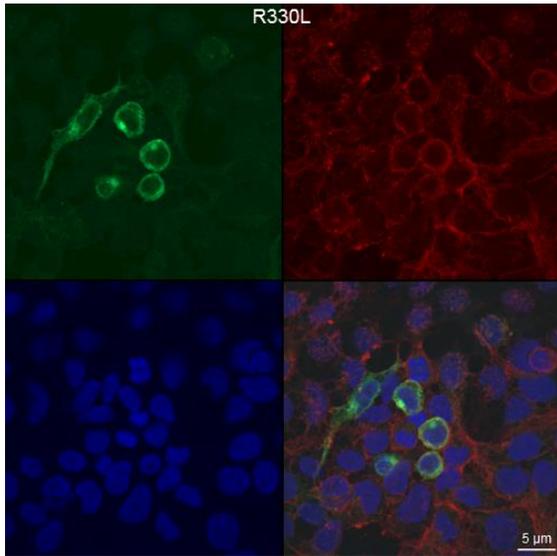
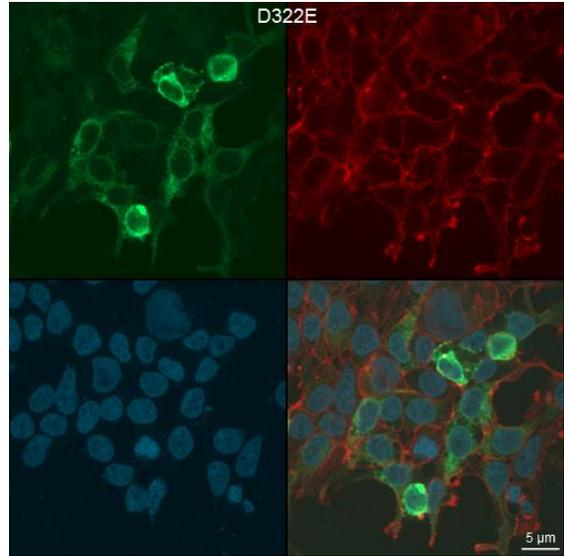
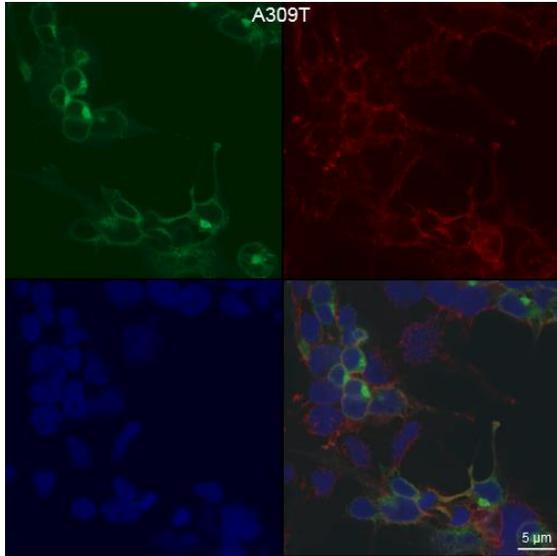


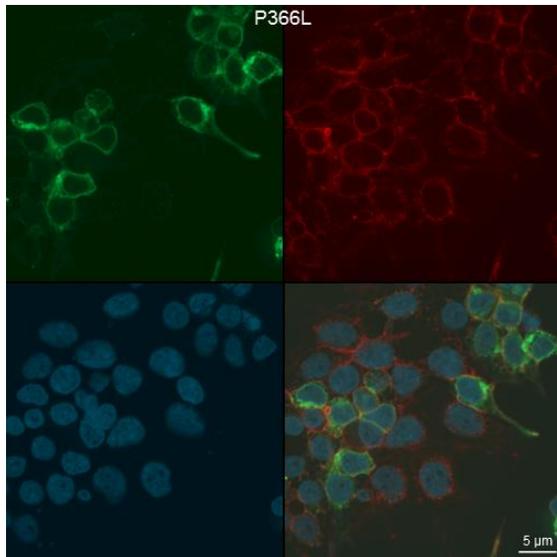






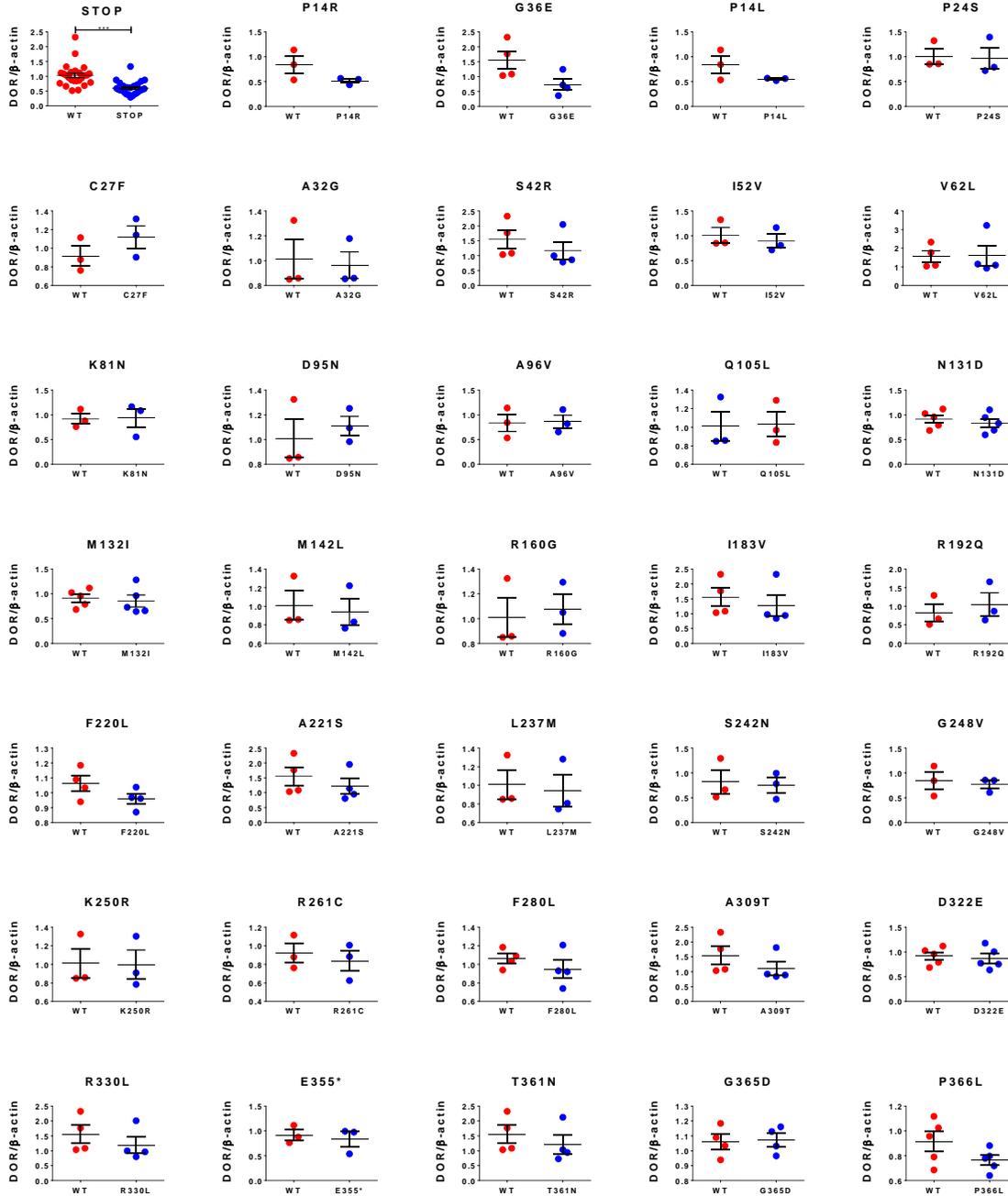






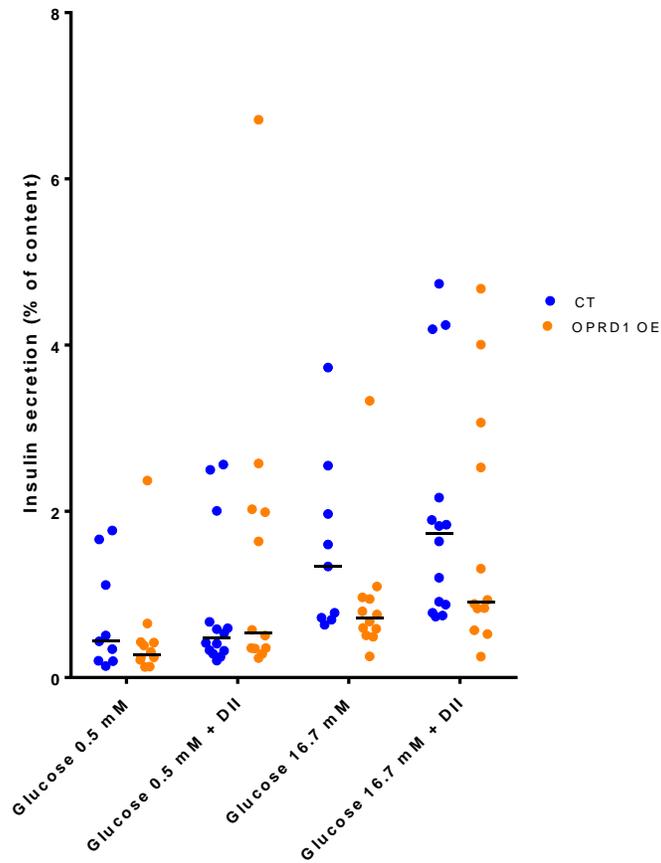
Confocal microscopy on HEK293 cells overexpressing or not (*i.e.* ‘not transfected’ condition) wild-type *OPRD1*, STOP *OPRD1*, and the *OPRD1* variants. For the STOP *OPRD1* plasmid, we used a plasmid containing *OPRD1* mutated for an early nonsense variant (c.3\_4insTAA/p.E2\*). The altered expression of the two variants encoding p.P14R and p.G36E was confirmed in a total of three independent experiments. DOP in green, extracellular matrix (F-actin) in red, and nuclei (DAPI) in blue. *NT*, not transfected; *WT*, wild-type.

## Supplementary Fig. 5. Expression of DOP mutants through Western blots



Expression of the STOP DOP, DOP mutants or wild-type DOP was assessed after overexpression of each related plasmid in HEK293 cells. For the STOP *OPRD1* plasmid, we used a plasmid containing *OPRD1* mutated for an early nonsense variant (c.3\_4insTAA/p.E2\*). Intensity of each band was normalized with the  $\beta$ -actin. Data of the mean normalized expression  $\pm$  SEM of each mutant or wild-type from at least three independent experiments. Two-by-two expression comparisons of each DOP mutant *versus* wild-type were performed using a Wilcoxon rank sum test. \*\*\* $P < 0.001$ . *WT*, wild-type.

Supplementary Fig. 6. Insulin levels in EndoC $\beta$ H1 cells (CT; blue) and in EndoC $\beta$ H1 cells overexpressing *OPRD1* (*OPRD1* OE; orange)



Insulin secretion was assessed in control EndoC $\beta$ H1 cells (CT; blue) and in EndoC $\beta$ H1 cells overexpressing *OPRD1* (*OPRD1* OE; orange). Insulin secretion was analyzed in response to 60 min of incubation with 0.5 mM glucose  $\pm$  100 nM DII, followed by 60 min of incubation with 16.7 mM glucose  $\pm$  100 nM DII. Data are median of five independent experiments. **DII**, [D-Ala<sup>2</sup>]-Deltorphan II; **Glc**, glucose.

**Supplementary Table 1. *OPRD1* variants detected in the RaDiO study including up to 6,971 participants**

Chr	Pos (Hg19)	Mutation	MAF in gnomAD	MAC in RaDiO	<i>In silico</i> prediction score <sup>†</sup>
1	29138936	c.41C>G, p.P14R	0	1	0.016
1	29138936	c.41C>T, p.P14L	0.00018	1	0.067
1	29138965	c.70C>T, p.P24S	0.000032	1	0.071
1	29138975	c.80G>T, p.C27F	0.91	common	0.10
1	29138990	c.95C>G, p.A32G	0	1	0.070
1	29139002	c.107G>A, p.G36E	0.0000063	1	0.075
1	29139021	c.126G>C, p.S42R	0	3	0.037
1	29139049	c.154A>G, p.I52V	0.0066 (0.074 in Africans)	28	0.16
1	29139079	c.184G>C, p.V62L	0	1	0.19
1	29185481	c.243G>T, p.K81N	0	1	0.23
1	29185521	c.283G>A, p.D95N	0.0000080	1	0.75
1	29185525	c.287C>T, p.A96V	0.000016	1	0.20
1	29185552	c.314A>T, p.Q105L	0	1	0.54
1	29185629	c.391A>G, p.N131D	0	1	0.51
1	29185634	c.396G>C, p.M132I	0.0000040	1	0.66
1	29185662	c.424A>T, p.M142L	0.000012	1	0.53
1	29185716	c.478C>G, p.R160G	0	1	0.76
1	29185785	c.547A>G, p.I183V	0.00018	1	0.035
1	29185813	c.575G>A, p.R192Q	0.000045	1	0.083
1	29189334	c.658T>C, p.F220L	0.000028	1	0.16
1	29189337	c.661G>T, p.A221S	0.000028	1	0.37
1	29189385	c.709C>A, p.L237M	0	1	0.21
1	29189401	c.725G>A, p.S242N	0	1	0.096
1	29189419	c.743G>T, p.G248V	0	1	0.45
1	29189425	c.749A>G, p.K250R	0.00013	1	0.15
1	29189457	c.781C>T, p.R261C	0.0000081	1	0.44
1	29189516	c.840C>G, p.F280L	0	1	0.39
1	29189601	c.925G>A, p.A309T	0.015	3	0.11
1	29189642	c.966C>G, p.D322E	0	1	0.27
1	29189665	c.989G>T, p.R330L	0.000032	1	0.33
1	29189739	c.1063G>T, p.E355*	0	1	NA
1	29189758	c.1082C>A, p.T361N	0.000057	1	0.075
1	29189770	c.1094G>A, p.G365D	0.00021	3	0.15
1	29189773	c.1097C>T, p.P366L	0.000039	3	0.066

*Chr*, chromosome; *gnomAD*, genome aggregation database browser (v2.1.1); *MAC*, minor allele count; *MAF*, minor allele frequency; *Pos*, position (according to the human alignment hg19/GRCh37).

<sup>†</sup>According to the rare exome variant ensemble learner (REVEL)

**Supplementary Table 2. Effect of rare *OPRD1* variants on metabolic homeostasis in RaDiO**

Trait	<i>n</i>	OR [CI]	$\pi \pm SE$	<i>P</i>
Hyperglycemia or T2D	6,745	0.80 [0.37–1.7]	-	0.56
Overweight or obesity	6,211	0.93 [0.48–1.9]	-	0.83
BMI (kg m <sup>-2</sup> )	6,971	-	0.13 ± 0.10	0.21
HDL* (mmol l <sup>-1</sup> )	6,008	-	0.12 ± 0.085	0.16
TG* (mmol l <sup>-1</sup> )	6,051	-	-0.098 ± 0.092	0.29
SBP** (mmHg)	5,257	-	5.0 ± 2.5	0.19
DBP** (mmHg)	4,693	-	0.21 ± 2.5	0.93

Association analyses were performed using the MiST method adjusted for age, sex, BMI and ancestry (for the analyses of hyperglycemia/T2D, HDL, TG, SBP and DBP) or for age, sex and ancestry (for the analyses of overweight/obesity and BMI). BMI and TG data were log-transformed before analysis. \*HDL and TG were only assessed in participants who were not under lipid-lowering agents. \*\*SBP and DBP were only assessed in participants who were not under high blood pressure medications. **BMI**, body mass index; **CI**, confidence interval; **DBP**, diastolic blood pressure; **HDL**, high-density lipoprotein; **IQR**, interquartile range;  $\pi$ , mean effect; **OR**, odds ratio; **SBP**, systolic blood pressure; **SD**, standard deviation; **SE**, standard error; **T2D**, type 2 diabetes; **TG**, triglycerides.

**Supplementary Table 3. Clinical data of carriers of a rare, loss- or gain-of-function *OPRD1* variant in the RaDiO study**

Clinical data	Carriers of a rare LOF variant	Carriers of a rare GOF variant	Non-carriers of a rare, LOF or GOF variant
<b><i>N</i></b>	13	8	6,950
<b>Sex</b>	M: 6 / F: 7	M: 3/ F: 5	M: 3,429/ F: 3,521
<b>Age at investigation</b>	47 ± 13	51 ± 12	51 ± 13
<b>BMI (kg/m<sup>2</sup>)</b>	27 [25–30]	25 [24–27]	26 [23–29]
<b>HDL* (mmol/l)</b>	1.3 ± 0.23	1.8 ± 0.39	1.5 ± 0.45
<b>TG* (mmol/l)</b>	0.90 [0.81–1.1]	0.55 [0.54–0.92]	1.1 [0.76–1.6]
<b>SBP** (mmHg)</b>	129 ± 20	120 ± 17	130 ± 15
<b>DBP** (mmHg)</b>	77 ± 11	70 ± 11	78 ± 9.4
<b><i>N</i> overweight/obesity (%)</b>	12 (92%)	3 (38%)	3,592 (58%)
<b><i>N</i> hyperglycemia (%)</b>	2 (15%)	4 (50%)	3,188 (47%)

Data are numbers, mean ± SD or median [interquartile range].

\*HDL and TG were only reported in participants who were not treated for lipid-lowering agents.

\*\*SBP and DBP were only reported in participants who were not treated for high blood pressure medications.

***BMI***, body mass index; ***DBP***, diastolic blood pressure; ***F***, female; ***GOF***, gain-of-function; ***HDL***, high-density lipoprotein; ***LOF***, loss-of-function; ***M***, male; ***SBP***, systolic blood pressure; ***TG***, triglycerides.

**Supplementary Table 4. Loss-of-function *OPRD1* variants detected in UK Biobank including up to 187,242 participants**

Chr	Pos (Hg38)	Mutation	MAF in gnomAD	MAC in UK Biobank	Detected in RaDiO as loss-of-function
1	28812384	c.1A>T, p.M1?	0	2	No
1	28812411	c.28G>T, p.E10*	0	1	No
1	28812424	c.41C>G, p.P14R	0	1	Yes
1	28812514	c.131C>A, p.S44*	0.0000053	13	No
1	28812555	c.174del, p.V59Cfs*22	0.000065	12	No
1	28859009	c.283G>A, p.D95N	0.0000080	1	Yes
1	28859013	c.287C>T, p.A96V	0.000016	9	Yes
1	28859041	c.316del, p.S106Vfs*5	0	1	No
1	28859066	c.342dup, p.P115Afs*31	0	4	No
1	28859067	c.341G>A, p.W114*	0.0000040	1	No
1	28859122	c.396G>C, p.M132I	0.0000040	1	Yes
1	28859301	c.575G>A, p.R192Q	0.000045	10	Yes
1	28862855	c.692dup, p.C232Vfs*80	0.0000040	1	No
1	28862863	c.699T>G, p.Y233*	0.0000040	1	No
1	28862989	c.828_832del, p.I277Hfs*33	0	1	No
1	28863120	c.959del, p.F320Sfs*158	0	1	No
1	28863130	c.966C>G, p.D322E	0	2	Yes
1	28863227	c.1063G>T, p.E355*	0	2	Yes
1	28863228	c.1067_1076del, p.R356Pfs*119	0.000011	1	No

*Chr*, chromosome; *gnomAD*, genome aggregation database browser (v2.1.1); *MAC*, minor allele count; *MAF*, minor allele frequency; *Pos*, position (according to the human alignment hg38/GRCh38).

## References

1. Bonnefond A, Boissel M, Bolze A, Durand E, Toussaint B, Vaillant E, et al. Pathogenic variants in actionable MODY genes are associated with type 2 diabetes. *Nat Metab.* 2020 Oct;2(10):1126–34.
2. Balkau B. [An epidemiologic survey from a network of French Health Examination Centres, (D.E.S.I.R.): epidemiologic data on the insulin resistance syndrome]. *Rev Epidemiol Sante Publique.* 1996 Aug;44(4):373–5.
3. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature.* 2007 Feb 22;445(7130):881–5.
4. Meyre D, Delplanque J, Chèvre J-C, Lecoœur C, Lobbens S, Gallina S, et al. Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. *Nat Genet.* 2009 Feb;41(2):157–9.
5. Romon M, Lafay L, Bresson JL, Oppert J-M, Borys J-M, Kettaneh A, et al. Relationships between physical activity and plasma leptin levels in healthy children: the Fleurbaix-Laventie Ville Santé II Study. *Int J Obes Relat Metab Disord.* 2004 Oct;28(10):1227–32.
6. American Diabetes Association. 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2019. *Diabetes Care.* 2019 Jan;42(Suppl 1):S13–28.
7. Baron M, Maillet J, Huyvaert M, Dechaume A, Boutry R, Loiselle H, et al. Loss-of-function mutations in MRAP2 are pathogenic in hyperphagic obesity with hyperglycemia and hypertension. *Nat Med.* 2019;25(11):1733–8.