

*ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTE*

IGBMC, CNRS UMR 7104 - Inserm U 1258

**THÈSE** présentée par :

**Colette GIETHLEN**

soutenue le : 23 janvier 2019

pour obtenir le grade de : **Docteur de l'Université de Strasbourg**

Discipline/ Spécialité : Aspects Moléculaires et Cellulaires de la Biologie

**Etude de la régulation transcriptionnelle  
de la différenciation des cellules  
entéroendocrines dans un modèle  
d'organoïde intestinal humain**

**THÈSE dirigée par :**

**Mr GRADWOHL Gérard**

Dr, Université de Strasbourg

**RAPPORTEURS :**

**Mme RIBEIRO-PILLET Agnès**

Dr, Université de Paris

**Mme VOZ Marianne**

Dr, Université de Liège

---

**AUTRES MEMBRES DU JURY :**

**Mme DANTZER Françoise**

Dr, Université de Strasbourg

Tout d'abord, je remercie les membres de mon jury d'avoir accepté d'évaluer ce travail de thèse.

Merci Gérard d'avoir supervisé cette thèse et de m'avoir accueillie dans l'équipe.

Merci à la fondation Novo Nordisk d'avoir financé ce projet de thèse.

Bien évidemment, merci à tous les membres de l'équipe, actuels et passés, pour vos conseils scientifiques et votre aide.

Aline, merci à toi de m'avoir encadrée pendant mon stage de master, de m'avoir transmis tes connaissances et de m'avoir initiée à de nombreuses techniques.

Adèle et Valérie, merci pour vos nombreux conseils et votre aide dans la préparation de la soutenance.

Sabitri, thank you for your help with the cell culture and everything else. I wish you all the best for your future.

Constance, merci de ton aide précieuse dans l'analyse des données de RNA-seq que je n'aurais pas pu faire sans toi. Je pense que ton absence va vite se faire ressentir au labo.

Florence, bon courage à toi pour la suite.

Je tiens à remercier les personnes des différentes plateformes ayant participé à ce projet. Merci particulièrement au service de culture cellulaire, Betty, Marion et Amélie ; à la plateforme de microscopie, Elvire, Yves, Marcel et Erwan ; à la plateforme de tri, Muriel et Claudine ; et à la plateforme de biopuces et séquençage, Christelle et Marie.

Thanks to Maria Elena's team, my lab neighbors for a few years, especially Diego (you little weirdo). It was a lot of fun kicking walls and stuff. I miss our parties! I wish you all the best and hope to see you soon.

Un énoooooorme merci à Céline & Julia, sans qui je n'écrirais probablement pas ces lignes aujourd'hui. C'était un bonheur de vous avoir comme collègues et amies. Vous allez me manquer même si je sais que nous nous reverrons souvent. Je vous souhaite le meilleur dans tous vos projets futurs.

Merci à Thomas & Laura pour les nombreuses soirées partagées, je vous souhaite beaucoup de bonheur et que notre belle amitié continue.

## Remerciements

---

Merci à Elise pour les nombreuses Kasteel rouges au Xanadu et les «crazy Sunday».

Merci aux « Vrais », qui se reconnaîtront.

Merci à tous les autres avec qui j'ai partagé des soirées et autres moments de convivialité, qui ont contribué à rendre ces trois années agréables.

Merci à mes parents de m'avoir toujours soutenue, notamment avec le traditionnel poulet-frites du samedi soir qui, j'en suis convaincue, a contribué au succès de cette thèse.

Merci mille fois à Etienne, qui m'a supporté pendant cette période (dieu sait que c'est pas toujours facile). Nul besoin de m'épancher ici, tu sais déjà tout ce que je pense.

& merci au petit Moumou pour son soutien inconditionnel !

# **TABLE DES MATIÈRES**

<b>Introduction</b>	<b>1</b>
<b>Partie 1. L'intestin, le principal organe endocrine</b>	<b>2</b>
1. Anatomie et histologie de l'intestin	2
1.1. Structure et rôle du tractus gastro-intestinal	2
1.2. Histologie de l'intestin	3
1.3. Les types cellulaires de l'épithélium intestinal	5
1.3.1. Les cellules souches	5
1.3.2. Les cellules du lignage absorbant	6
• Les entérocytes	6
1.3.3. Les cellules du lignage sécréteur	6
• Les cellules à mucus	6
• Les cellules de Paneth	6
• Les cellules entéroendocrines	7
1.3.4. Autres types cellulaires	7
• Les cellules M	7
• Les cellules « en touffe »	8
• Les cellules « en coupe »	8
2. Morphogenèse intestinale	8
2.1. Formation de l'endoderme	10
2.2. Régionalisation de l'endoderme	11
2.3. Formation des organes	12
3. Différenciation et renouvellement de l'épithélium intestinal	13
3.1. Les cellules souches intestinales : deux modèles décrits	13
3.2. Plasticité de l'épithélium intestinal	14
3.3. Maintien de l'homéostasie intestinale	15
3.3.1. Régulation de l'équilibre crypte/villosité	15
3.3.2. Régulation de la prolifération et différenciation cellulaire	16
4. Description des principales voies de signalisation moléculaire impliquées dans le développement et l'homéostasie intestinale	18
• Wnt	18
• Nodal	19
• Notch	20
• BMP	20
<b>Partie 2. Les cellules entéroendocrines</b>	<b>21</b>
1. Fonctions physiologiques et diversité des cellules entéroendocrines	21
1.1. GLP1 / GIP : l'effet incrétine	24
2. Mécanismes de sécrétion des hormones	25
3. Révision du dogme un sous-type de cellules entéroendocrines / un peptide produit	26
4. Régulation transcriptionnelle de la différenciation entéroendocrine	27
• Ngn3	29
• Arx et Pax4	30
• Rfx6	30
• NeuroD1	31
• FoxA1 et FoxA2	31
• Autres facteurs de transcription	31
5. Pathologies humaines associées à des défauts dans la régulation transcriptionnelle de la différenciation entéroendocrine	33
5.1. Mutation de <i>NGN3</i> : diarrhée congénitale malabsorptive	33

5.2. Mutation de <i>RFX6</i> : syndrome de Mitchell-Riley	34
5.3. Mutation de <i>ARX</i>	34
<b>Partie 3. Modèles d'étude du développement et de la différenciation intestinale</b>	<b>36</b>
1. Modèles <i>in vivo</i>	37
2. Modèles <i>in vitro</i>	38
3. Modèles <i>ex vivo</i>	41
<b>Partie 4. CRISPR/Cas9, outil de modification du génome</b>	<b>44</b>
1. Historique et rôle dans l'immunité adaptative chez les procaryotes	44
2. Applications dans la modification de génome	46
3. Les limites de l'outil CRISPR/Cas9	49
<b>Objectif du projet de thèse</b>	<b>49</b>
<b>Matériel &amp; méthodes</b>	<b>51</b>
1. CRISPR/Cas9	51
1.1. Plasmides utilisés	51
1.2. Clonage des sgRNAs dans les plasmides pX461	52
1.3. Surveyor®: détection de mutation	52
2. Matériel biologique	53
2.1. Cellules HEK293T	53
• Origine et culture	53
• Transfection	53
2.2. hiPSCs	54
• Origine et culture	54
• Transfection	54
• Génotypage des clones	55
2.3. Génération et maintien des HIOs	56
2.3.1. Différenciation intestinale <i>in vitro</i>	56
2.3.2. Passage des HIOs	56
3. Analyses transcriptomiques	56
3.1. Extraction d'ARN	56
• HIOs	56
• hiPSCs	58
3.2. RT-qPCR	58
3.3. RNA-Seq	58
4. Histologie	59
4.1. Immunofluorescences sur HIOs entiers/ <i>in toto</i> ( <i>whole-mount</i> )	59
4.2. Transparisation	61
4.3. Imagerie	61
5. Analyse au FACS	61
6. Analyses statistiques	61
<b>Résultats</b>	<b>63</b>
<b>Partie 1. Optimisation de la production d'organoïdes intestinaux humains (HIOs)</b>	<b>63</b>
1. Contrôle qualité des étapes clés du protocole de différenciation des HIOs <i>in vitro</i>	64
1.1. Influence de l'enzyme de dissociation sur la densité cellulaire	65
1.2. Formation de l'endoderme définitif	66

1.3. Formation de l'intestin « moyen » : les sphéroïdes intestinales	68
1.4. Croissance et expansion des sphéroïdes en HIOs	70
<b>Partie 2. Inactivation des gènes codant les facteurs de transcription NGN3, RFX6, ARX et PAX4</b>	<b>72</b>
1. Mise au point de l'outil CRISPR/Cas9	72
1.1. Approche utilisant une version mutée de la nucléase Cas9 : la double- <i>nickase</i>	73
1.1.1. Choix et validation des ARNs guides (sgRNAs)	73
1.2. Approche utilisant la Cas9 sauvage	76
2. Inactivation des facteurs de transcription dans les hiPSCs	76
2.1. Inactivation de <i>NGN3</i>	76
2.1.1. Utilisation de la stratégie de double- <i>nickase</i>	76
2.1.2. Seconde approche CRISPR/Cas9 : utilisation de la Cas9 sauvage	80
2.2. Inactivation de <i>ARX</i>	81
2.3. Inactivation de <i>RFX6</i>	83
2.4. Le cas de <i>PAX4</i>	85
2.5. Analyse de la pluripotence des hiPSCs LV37 et des clones <i>NGN3<sup>-/-</sup></i> et <i>RFX6<sup>-/-</sup></i> générés	87
<b>Partie 3. Etude des conséquences de l'inactivation de <i>NGN3</i> sur la différenciation entéroendocrine</b>	<b>89</b>
1. Analyses par RT-qPCR et immunofluorescences	90
1.1. RT-qPCR	90
1.2. Immunofluorescences <i>in toto</i>	93
1.3. Conclusion des analyses par RT-qPCR et immunofluorescences	95
2. Séquençage des ARNm (RNA-seq)	96
2.1. Différence entre chaque échantillon	97
2.2. Alignements des <i>reads</i> obtenus sur la séquence d'ADN codant <i>NGN3</i>	98
2.3. Nature intestinale et épithéliale des HIOs générés	99
2.4. Etude de l'expression de marqueurs généraux de cellules entéroendocrines	102
2.5. Impact de l'inactivation de <i>NGN3</i> sur la différenciation entéroendocrine	104
2.5.1. Expression des gènes codant les produits de sécrétion des cellules entéroendocrines ou leur enzyme de biosynthèse	104
2.5.2. Expression des gènes codant des facteurs de transcription régulant la différenciation entéroendocrine	105
2.6. Gènes les plus différenciellement exprimés de manière significative dans les HIOs <i>NGN3<sup>-/-</sup></i>	108
2.6.1. Gènes les plus sous-exprimés de manière significative dans les HIOs <i>NGN3<sup>-/-</sup></i>	109
2.6.2. Etude de la nature proximale des HIOs	110
2.6.3. Perturbation de la régulation de la prolifération cellulaire dans les HIOs <i>NGN3<sup>-/-</sup></i>	112
2.6.4. Gènes les plus surexprimés de manière significative dans les HIOs <i>NGN3<sup>-/-</sup></i>	113
2.7. Conclusions et points de discussion relevés	115

<b>Partie 4. Etude des conséquences de l'inactivation de <i>RFX6</i> sur la différenciation entéroendocrine chez l'homme</b>	<b>118</b>
1. Etudes des modèles murins déficients pour <i>Rfx6</i> (Piccand et al., manuscrit en préparation)	118
2. Analyse par RT-qPCR et immunofluorescences	121
2.1. RT-qPCR	121
2.2. Immunofluorescences <i>in toto</i>	124
2.3. Conclusion des analyses par RT-qPCR et immunofluorescences	127
3. Séquençage des ARNm (RNA-seq)	128
3.1. Différence entre chaque échantillon	128
3.2. Alignements des <i>reads</i> obtenus sur la séquence d'ADN codant <i>RFX6</i>	129
3.3. Nature épithéliale et intestinale des HIOs générés	131
3.4. Etude de l'expression de marqueurs généraux de cellules entéroendocrines	133
3.5. Impact de l'inactivation de <i>RFX6</i> sur la différenciation entéroendocrine	134
3.5.1. Expression des gènes codant les hormones ou leurs enzymes de biosynthèse exprimées par les cellules entéroendocrines	134
3.5.2. Expression des gènes codant des facteurs de transcription régulant la différenciation entéroendocrine	136
3.6. Comparaison des variations d'expression des gènes d'intérêt entre les HIOs <i>RFX6</i> <sup>-/-</sup> et les modèles murins	139
3.7. Analyses fonctionnelles des résultats de RNA-seq	141
3.7.1. Gènes impliqués dans le développement, la morphogenèse et la différenciation de l'intestin	141
3.7.2. Gènes impliqués dans le métabolisme et la physiologie des cellules épithéliales intestinales	142
3.7.3. Gènes codant des éléments initialement attribués au système nerveux	145
3.7.4. Analyse comparative des données transcriptomiques issues des différents modèles d'inactivation de <i>Rfx6/RFX6</i>	147
3.8. Conclusions et points de discussion relevés	149
<b>Discussion</b>	<b>152</b>
1. Discussion des résultats expérimentaux	152
1.1. Génération des modèles de perte de fonction dans les hiPSCs : difficultés et limites	152
• Récapitulatif des clones générés par les deux stratégies CRISPR/Cas9	152
• Limites de la technologie CRISPR/Cas9	152
1.2. Phénotype des HIOs <i>NGN3</i> <sup>-/-</sup>	153
• Les HIOs <i>NGN3</i> <sup>-/-</sup> ne présentent pas de cellules CHGA <sup>+</sup>	153
• L'expression des facteurs de transcription et hormones entéroendocrines ne semble pas affectée dans les HIOs <i>NGN3</i> <sup>-/-</sup>	154
• Implication de <i>NGN3</i> dans le lignage sécréteur ?	154
• Compensation partielle de l'absence de <i>NGN3</i> ?	155
• Les clones <i>NGN3</i> <sup>-/-</sup> se différencient en tissu épithélial intestinal de nature proximale plutôt que distale	155
• Le nombre de passage des hiPSCs, un potentiel biais expérimental ?	157

• Prolifération cellulaire aberrante et anomalies morphologiques associées à la mutation ?	157
1.3. Potentiels rôles de RFX6 dans les cellules intestinales chez l'homme	158
• La régulation de la différenciation entéroendocrine est affectée par la perte de RFX6	158
• RFX6 semble impliqué dans la différenciation et/ou fonction des cellules entérochromaffines et entérochromaffines-like	158
• Le rôle précis de RFX6 dans la différenciation entéroendocrine reste à déterminer	159
• Implication de RFX6 dans la formation des progéniteurs sécréteurs et absorbants ?	159
• Vers un rôle de RFX6 dans la fonction des cellules entéroendocrines ?	161
• La mutation de <i>RFX6</i> est associée à des anomalies morphologiques du tube digestif	163
• Des marqueurs de tissus nerveux sont détectés dans les HIOs et semblent affecter par la mutation de <i>RFX6</i>	163
• Conclusion et perspectives quant au rôle de RFX6 dans l'intestin	164
2. Limites du modèle HIOs	165
2.1. Variabilité de la proportion épithélium/mésenchyme au sein des échantillons	165
2.2. Variabilité dans la morphologie des HIOs	165
2.3. Effectifs limités et rareté des cellules entéroendocrines	167
2.4. Maturité des HIOs	168
2.5. Considérations méthodologiques concernant la culture cellulaire des hiPSCs	168
3. Les HIOs, un modèle pertinent ?	170

## Bibliographie

173

**Annexe 1:** Tableau récapitulatif des phénotypes observés dans les modèles déficients pour nos gènes d'intérêt (*Ngn3*, *Arx*, *Pax4*), décrits dans la littérature

**Annexe 2:** Protocole de différenciation intestinale *in vitro*

**Annexe 3:** Liste des potentiels sites de fixation non-spécifique des sgRNAs

**Annexe 4:** Gènes les plus différenciellement exprimés dans les HIOs RFX6<sup>-/-</sup>

**Annexe 5:** Gènes différenciellement exprimés communs aux 3 modèles: Rfx6<sup>-/-</sup>E18, Rfx6<sup>ΔAdInt</sup> et HIOs RFX6<sup>-/-</sup>

# **LISTE DES FIGURES**

<b>Figure 1.</b> Organisation générale du tractus digestif chez l'homme	2
<b>Figure 2.</b> Histologie du tractus digestif	3
<b>Figure 3.</b> Les différents types cellulaires de l'épithélium intestinal	4
<b>Figure 4.</b> Les cellules souches des cryptes intestinales	5
<b>Figure 5.</b> Schéma récapitulant les grandes étapes de la formation des organes d'origine endodermique au cours du développement embryonnaire	9
<b>Figure 6.</b> Facteurs impliqués dans la régionalisation de l'intestin primitif au cours du développement embryonnaire	11
<b>Figure 7.</b> Schéma illustrant la complexité de la régulation transcriptionnelle de la régionalisation des dérivés endodermiques	12
<b>Figure 8.</b> Schéma représentant la plasticité cellulaire homéostatique et en cas de dommage à l'épithélium intestinal	15
<b>Figure 9.</b> Voies de signalisation régulant l'homéostasie de l'épithélium intestinal	16
<b>Figure 10.</b> Voies de signalisation Wnt	19
<b>Figure 11.</b> Voie de signalisation Nodal	19
<b>Figure 12.</b> Voie de signalisation canonique Notch	20
<b>Figure 13.</b> Voie de signalisation BMP	20
<b>Figure 14.</b> Principales fonctions des cellules entéroendocrines et hormones impliquées	21
<b>Figure 15.</b> Morphologies et modes de signalisation adoptés par les cellules entéroendocrines	22
<b>Figure 16.</b> Patron d'expression des hormones entéroendocrines le long de l'axe proximo-distal chez la souris	22
<b>Figure 17.</b> Fonctions des hormones entéroendocrines sécrétées en réponse à divers stimuli	23
<b>Figure 18.</b> Structures du gène proglucagon, de son ARNm et de la protéine	25
<b>Figure 19.</b> Sécrétion des hormones par les cellules entéroendocrines	26
<b>Figure 20.</b> Principaux facteurs de transcription impliqués dans la différenciation des sous-types entéroendocrines sécréteurs d'hormones	28
<b>Figure 21.</b> Chronologie des découvertes majeures dans le domaine des cellules souches	36
<b>Figure 22.</b> Potentiel des hiPSCs : approches thérapeutiques et modélisation	38
<b>Figure 23.</b> Différenciation <i>in vitro</i> dirigée des cellules souches pluripotentes en divers tissus	39

<b>Figure 24.</b> La différenciation <i>in vitro</i> des cellules souches pluripotentes en organoïdes intestinaux humains (HIOs) mime le développement embryonnaire humain	40
<b>Figure 25.</b> Croissance d'une cellule souche Lgr5 <sup>+</sup> en entéroïde <i>ex vivo</i>	42
<b>Figure 26.</b> Mécanisme de l'immunité adaptative médiée par le système CRISPR/Cas de classe II chez les procaryotes	45
<b>Figure 27.</b> ARN guide chimérique (sgRNA) synthétisé pour remplir les fonctions du duplex crRNA-tracrRNA	46
<b>Figure 28.</b> Mécanismes de réparation de la coupure double-brin générée par la Cas9	47
<b>Figure 29.</b> Inactivation de gène : deux stratégies CRISPR/Cas9 possibles	48
<b>Figure 30.</b> Plasmides CRISPR/Cas9	51
<b>Figure 31.</b> Tableau récapitulatif des sgRNAs utilisés pour cibler les gènes d'intérêt	52
<b>Figure 32.</b> Amorces utilisées pour les amplifications PCR du génotypage des clones hiPSCs et pour le séquençage du plasmide pX461	55
<b>Figure 33.</b> Protocole de différenciation <i>in vitro</i> des hiPSCs en HIOs	57
<b>Figure 34.</b> Liste des produits utilisés pour la culture et la différenciation des hiPSCs	57
<b>Figure 35.</b> Amorces UPL utilisées pour les expériences de RT-qPCR	58
<b>Figure 36.</b> Anticorps utilisés pour les immunofluorescences et analyses au FACS	60
<b>Figure 37.</b> Chronologie des étapes de la différenciation intestinale <i>in vitro</i>	64
<b>Figure 38.</b> Contrôle de l'efficacité des étapes de la différenciation intestinale <i>in vitro</i>	67
<b>Figure 39.</b> Les HIOs expriment des marqueurs d'épithélium intestinal au stade J28	71
<b>Figure 40.</b> Mise au point de l'inactivation de <i>NGN3</i> par la stratégie de double- <i>nickase</i>	74
<b>Figure 41.</b> Représentation schématique de la structure des gènes cibles et de leur protéine associée	77
<b>Figure 42.</b> Inactivation de <i>NGN3</i> par l'approche de double- <i>nickase</i>	79
<b>Figure 43.</b> Génération de deux mutants <i>NGN3</i> <sup>-/-</sup> par délétion de la séquence codante	80
<b>Figure 44.</b> Génération d'un mutant <i>ARX</i> <sup>-/-</sup> et d'un mutant <i>ARX</i> <sup>+/-</sup> par délétion de l'exon 1	82
<b>Figure 45.</b> Génération de trois mutants <i>RFX6</i> <sup>-/-</sup> par délétion des exons 1 et 2	84

<b>Figure 46.</b> Tentative d'inactivation de <i>PAX4</i> par délétion de l'exon 1 à 3 codant le domaine <i>paired</i>	85
<b>Figure 47.</b> Analyse du caryotype des hiPSCs LV37 et de la pluripotence des clones <i>NGN3</i> <sup>-/-</sup> , <i>RFX6</i> <sup>-/-</sup> et des LV37	88
<b>Figure 48.</b> Analyse des HIOs <i>NGN3</i> <sup>-/-</sup> au stade J28 par RT-qPCR sur divers marqueurs intestinaux d'intérêt	90
<b>Figure 49.</b> Immunofluorescences du marqueur entéroendocrine CHGA et <i>NGN3</i> sur les HIOs <i>NGN3</i> <sup>-/-</sup> et CTRL au stade J28	93
<b>Figure 50.</b> Analyse de la différence entre chaque échantillon des expériences de RNA-seq	97
<b>Figure 51.</b> Alignement des <i>reads</i> sur <i>NGN3</i>	99
<b>Figure 52.</b> Tableau récapitulatif de l'expression de marqueurs spécifiques aux sous-types cellulaires de l'épithélium intestinal	100
<b>Figure 53.</b> Expression de marqueurs généraux de cellules entéroendocrines	103
<b>Figure 54.</b> Niveau d'expression des gènes codant les hormones produites dans le tractus digestif ou leur enzyme de biosynthèse	104
<b>Figure 55.</b> Tableau récapitulatif de l'expression de facteurs de transcription régulant la différenciation des cellules de l'épithélium intestinal	106
<b>Figure 56.</b> Gènes les plus différenciellement exprimés dans les HIOs <i>NGN3</i> <sup>-/-</sup>	108
<b>Figure 57.</b> Tableau présentant des gènes significativement sous-exprimés dans les HIOs <i>NGN3</i> <sup>-/-</sup>	109
<b>Figure 58.</b> Tableau représentant le niveau d'expression de certains marqueurs proximaux du tube digestif	110
<b>Figure 59.</b> <i>Heatmap</i> représentant la différence de nature proximo-distale des HIOs <i>NGN3</i> <sup>-/-</sup> en comparaison aux CTRL	111
<b>Figure 60.</b> Tableau représentant l'expression de gènes impliqués dans la régulation de la prolifération cellulaire de l'épithélium intestinal	112
<b>Figure 61.</b> Tableau présentant des gènes significativement surexprimés dans les HIOs <i>NGN3</i> <sup>-/-</sup>	114
<b>Figure 62.</b> Etude de l'expression de <i>Rfx6</i> et des conséquences de son inactivation sur la différenciation entéroendocrine chez la souris	119
<b>Figure 63.</b> Analyse des HIOs <i>RFX6</i> <sup>-/-</sup> au stade J28 par RT-qPCR sur divers marqueurs intestinaux d'intérêt	122
<b>Figure 64.</b> Transparisation au fructose des HIOs CTRL après immunofluorescence sur CHGA	124
<b>Figure 65.</b> Immunofluorescences du marqueur entéroendocrine CHGA sur les HIOs entiers <i>RFX6</i> <sup>-/-</sup> et CTRL au stade J28	125

<b>Figure 66.</b> Analyse de la différence entre chaque échantillon des expériences de RNA-seq	129
<b>Figure 67.</b> Alignement des reads sur <i>RFX6</i>	130
<b>Figure 68.</b> Tableau récapitulatif de l'expression de marqueurs spécifiques aux sous-types cellulaires de l'épithélium intestinal	133
<b>Figure 69.</b> Expression de marqueurs généraux de cellules entéroendocrines	135
<b>Figure 70.</b> Niveau d'expression des gènes codant les hormones produites dans le tractus digestif ou leur enzyme de biosynthèse	136
<b>Figure 71.</b> Tableau récapitulatif de l'expression de facteurs de transcription régulant la différenciation des cellules de l'épithélium intestinal	137
<b>Figure 72.</b> Tableau comparatif des phénotypes entéroendocrines observés dans différents modèles déficients pour <i>Rfx6/RFX6</i>	140
<b>Figure 73.</b> Gènes différentiellement exprimés dans les HIOs <i>RFX6<sup>-/-</sup></i> et impliqués dans la morphologie cellulaire et dans le développement du système digestif	141
<b>Figure 74.</b> Gènes différentiellement exprimés dans les HIOs <i>RFX6<sup>-/-</sup></i> et impliqués dans le métabolisme des lipides, le métabolisme du tryptophane et la glycolyse et gluconéogenèse	143
<b>Figure 75.</b> Gènes différentiellement exprimés dans les HIOs <i>RFX6<sup>-/-</sup></i> et impliqués dans le métabolisme du glucose et la sécrétion hormonale par les cellules entéroendocrines	144
<b>Figure 76.</b> Gènes différentiellement exprimés dans les HIOs <i>RFX6<sup>-/-</sup></i> et impliqués dans l'activité des canaux cationiques et dans la signalisation synaptique	146
<b>Figure 77.</b> Comparaison des gènes différentiellement exprimés entre les différents modèles d'inactivation de <i>Rfx6/RFX6</i>	147
<b>Figure 78.</b> Rôle de <i>Ngn3</i> dans la différenciation endocrine de l'épithélium gastrique chez la souris	156
<b>Figure 79.</b> Schéma récapitulatif des conséquences de l'inactivation de <i>RFX6</i> sur la différenciation endocrine au sein des HIOs	160
<b>Figure 80.</b> Similitudes des mécanismes de sécrétion hormonale en réponse aux nutriments entre les cellules $\beta$ pancréatiques et les cellules entéroendocrines	162
<b>Figure 81.</b> Variabilité de la proportion d'épithélium et de mésenchyme entre les différents échantillons	166
<b>Figure 82.</b> Hétérogénéité morphologique des HIOs	167

# **LISTE DES ABRÉVIATIONS**

3D : 3 dimensions  
5-HT : Sérotonine  
A-P : Antéro-Postérieur  
a1/a2 : allèle 1/allèle 2  
ADN : Acide Désoxyribonucléique  
AIP : Anterior Intestinal Portal  
AR : Acide Rétinoïque  
ARN : Acide Ribonucléique  
ARNm : ARN messenger  
ARNr : ARN ribosomique  
ARX: Aristaless Related Homeobox  
ATP : Adénosine triphosphate  
BER : Base Excision Repair  
bHLH : basic Helix-Loop-Helix  
BMP : Bone Morphogenic Protein  
Cas9n : Cas9 *nickase*  
CBC : Crypt Base Columnar  
CCK : Cholecystokinine  
CDX1/2 : Caudal Type Homeobox 1/2  
CHGA : Chromogranin A  
CIP : Caudal Intestinal Portal  
CRISPR : Clustered Regularly Interspaced Palindromic Repeats  
crRNA : CRISPR RNA  
crRNP : CRISPR Ribonucleoprotein  
CTRL : contrôle  
dCas9 : dead Cas9  
D-V : Dorso-Ventral  
DAPI : 4',6-diamidino-2-phénylindole  
DE : Definitive Endoderm  
DLL : Delta Like Ligand  
DMEM : Dulbecco's Modified Eagle Medium  
DMSO : Diméthylsulfoxyde  
DS : Donkey Serum  
EC : Entérochromaffine  
ECL : Entérochromaffine-like  
EDTA : Éthylène diamine tétraacétique  
EGF : Epithelial Growth Factor  
ELISA : Enzyme-Linked Immunosorbent Assay  
ENS : Enteric Nervous System

F/R : Forward/Reverse  
FABP : Fatty Acid-Binding Protein  
FACS : Fluorescence Activated Cell Sorting  
FC : Fold Change  
FCS : Fetal Calf Serum  
FFA : Free Fatty Acids  
Fg : Foregut  
FGF : Fibroblast Growth Factor  
FOXA1/2 : Forkhead Box A1/2  
Fz : Frizzled  
GAST : Gastrine  
GFP : Green Fluorescent Protein  
GHRL : Ghréline  
GIP : Glucose-dependent Insulinotropic Polypeptide  
GLP1 : Glucagon-Like Peptide 1  
GO : Gene Ontology  
GPCR : Récepteurs couplés aux protéines G  
GSIS : Glucose-Stimulated Insulin Secretion  
GSK3 $\beta$  : Glycogen Synthase Kinase 3 beta  
GRPP : Glicentin-Related Pancreatic Polypeptide  
HBSS : Hank's Balanced Salt Solution  
HCO : Human Colonic Organoid  
HDR : Homology Directed Repair  
HEK293T : Human Embryonic Kidney 293T  
hESC : cellules souches embryonnaires humaines  
Hg : Hindgut  
HGO : Human Gastric Organoid  
HH : Hedgehog  
HIO: Human Intestinal Organoid  
hiPSC : cellules souches pluripotentes humaines induites  
HNMT : Histamine N-Methyltransferase  
IDM : Intestinal Differentiation Medium  
iHH : Indian Hedgehog  
indel : insertion-délétion  
INS : insuline  
InsI5 : Insulin-like peptide 5  
IP : Intervening Peptide  
KEGG : Kyoto Encyclopedia of Genes and Genomes  
LFC : Log Fold Change

LGR5 : Leucine-rich repeat-containing G-protein coupled Receptor 5  
LRC : Label Retaining Cell  
MCI : Masse Cellulaire Interne  
mESC : cellules souches embryonnaires de souris  
Mg : Midgut  
MHG : Mid-Hindgut  
MLN : Motiline  
MPGF : major proglucagon fragment  
NGN3: Neurogenin 3  
NHEJ : Non Homologous End-Joining  
NLS : Nuclear Localization Signal  
NPY : Neuropeptide Y  
NTS : Neurotensine  
nucléase S : nucléase Surveyor  
OP : Octapeptide  
OXM : Oxyntomodulin  
P : passages  
PAM : Protospacer Adjacent Motif  
PAX4: Paired Box 4  
pb/kb : paire de bases / kilobase  
PBS : Phosphate Buffer Saline  
PCA/ACP : Analyse en Composantes Principales  
PCP : Planar Cell Polarity  
PCR : Polymerase Chain Reaction  
PFA : Paraformaldehyde  
PPY : Pancreatic Polypeptide Y  
PS : Penicilline-Streptomycine  
PYY : Peptide YY  
RFX6 : Regulatory Factor X6  
RPMI : Roswell Park Memorial Institute medium  
RT-qPCR : Real Time - quantitative Polymerase Chain Reaction  
SCT : Sécrétine  
SDS : Sodium Dodecyl Sulfate  
SGLT1 : Cotransporteur sodium-glucose 1  
sgRNA : single guide RNA  
SHH : Sonic Hedgehog  
shRNA : short hairpin RNA  
SST : Somatostatine  
SV40 : Virus simien 40

TALEN : Transcription Activator-Like Effector Nuclease

TCF : T-Cell Factor

TGF- $\beta$  : Transforming Growth Factor beta

TE : Tris-EDTA

TPH1 : Tryptophan Hydroxylase 1

tracrRNA : trans-activating CRISPR RNA

TSA : Tyramide Signal Amplification

UPL : Universal Probe Library

VIP : Vasoactive Intestinal Peptide

WT : Wild Type (sauvage)

ZFN : Zinc Finger Nuclease

# INTRODUCTION

# Introduction

L'intestin est connu pour son rôle dans la digestion et notamment dans l'absorption des nutriments. Chez l'homme, des anomalies dans le développement embryonnaire du tractus digestif peuvent provoquer de sévères malformations, menant à des problèmes d'absorption des nutriments. Ainsi, la biologie développementale du tube digestif est un domaine de fort potentiel dans l'identification de cibles thérapeutiques notamment dans le cadre des maladies intestinales congénitales. Comprendre les mécanismes régulant le développement de cet organe semble donc essentiel et pourrait permettre de développer des approches thérapeutiques de réparation et de régénération de l'intestin.

L'épithélium intestinal constitue une barrière physiologique et immunologique séparant le contenu luminal des tissus sous-jacents et les cellules le composant sont capables d'émettre des signaux mécaniques, nerveux et hormonaux en réponse à des stimuli divers et variés. Parmi celles-ci, les cellules entéroendocrines, qui représentent seulement 1% des cellules de l'épithélium, sont des régulateurs essentiels du métabolisme énergétique et l'altération de leur différenciation provoque des pathologies comme la diarrhée congénitale malabsorptive (Sansbury et al., 2015; Wang et al., 2006).

C'est dans ce contexte que s'inscrit ce projet de thèse qui vise à élucider les mécanismes régulant la différenciation entéroendocrine dans un modèle *in vitro* humain. Les processus de différenciation des cellules entéroendocrines impliquent une cascade de régulations transcriptionnelles dont l'ensemble des mécanismes n'est pas encore compris. L'objectif de ce projet est donc de déterminer le rôle de différents facteurs de transcription (NGN3, RFX6, ARX, PAX4) impliqués dans cette régulation.

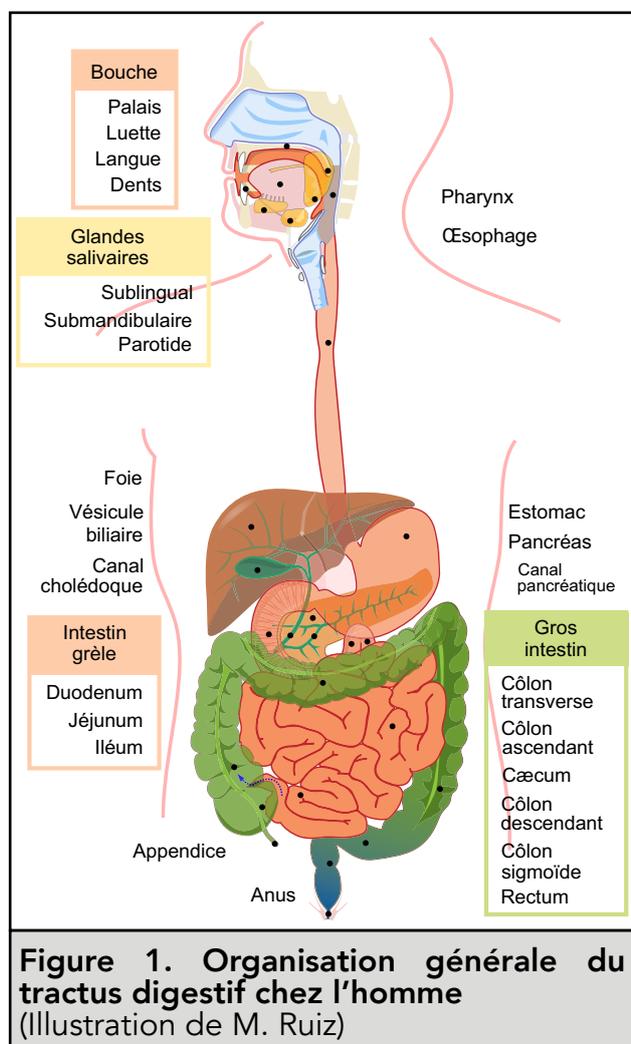
## Partie 1. L'intestin, le principal organe endocrine

La complexité fonctionnelle et structurelle de l'intestin témoigne de l'ampleur des tâches qu'il a à accomplir. Ces diverses fonctions comprennent la protection de l'organisme contre le contenu luminal, la digestion et l'absorption des nutriments, le maintien de l'équilibre hydrique et électrolytique ainsi que l'intégration des stimuli reliés à l'alimentation.

### 1. Anatomie et histologie de l'intestin

#### 1.1. Structure et rôle du tractus gastro-intestinal

Le tractus gastro-intestinal est une structure tubulaire qui s'étend de la bouche à l'anus et qui comprend la cavité buccale, l'œsophage, l'estomac, l'intestin et les organes associés comme le pancréas, le foie et la vésicule biliaire (**Figure 1**). L'intestin est lui-même divisé en deux grandes parties : l'intestin grêle composé du duodénum, jéjunum et iléon et le gros intestin composé du caecum, côlon et rectum. La nourriture est ingérée par la bouche et réduite en petits fragments par la mastication et l'action de la salive, qui constituent les premières étapes de la digestion. Les aliments mâchés passent par l'œsophage pour entrer dans l'estomac où ils vont subir une digestion mécanique (brassage) et chimique par les sucs gastriques. Le chyme formé va ensuite passer dans l'intestin grêle où les nutriments sont absorbés notamment sous l'action des sucs pancréatiques et biliaires. Les aliments non absorbés entrent dans le caecum où ils seront fermentés et décomposés par le microbiote intestinal. Enfin, les fluides et électrolytes sont absorbés dans le gros intestin et la matière non digérée restante est éliminée par le rectum et l'anus.

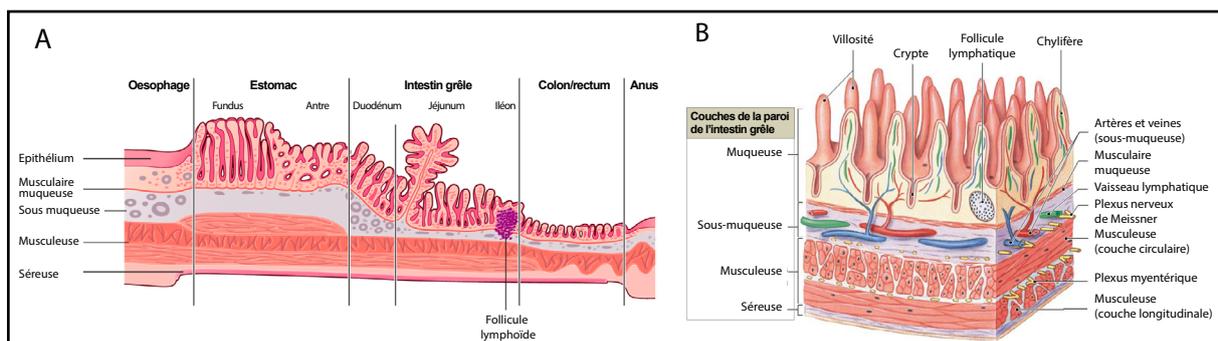


## 1.2. Histologie de l'intestin

L'intestin humain se compose, de la lumière vers la périphérie, d'une :

- muqueuse qui comprend l'épithélium unistratifié tapissant le côté luminal de la paroi intestinale, la *lamina propria* contenant des vaisseaux sanguins et lymphatiques, des fibres nerveuses et des fibroblastes, et une couche de muscle lisse séparant la muqueuse de la sous-muqueuse;
- sous-muqueuse qui est un tissu conjonctif dense contenant de nombreux vaisseaux sanguins et lymphatiques ainsi que le plexus nerveux sous-muqueux, appelé plexus de Meissner;
- musculieuse composée de cellules musculaires lisses disposées en deux sous-couches: une couche interne (côté luminal) où la disposition des cellules est circulaire et une couche externe où la disposition est longitudinale;
- séreuse qui est une couche de tissu conjonctif de cellules mésothéliales (**Figure 2A**) (Netter, 1996 ; Madara et al., 1990).

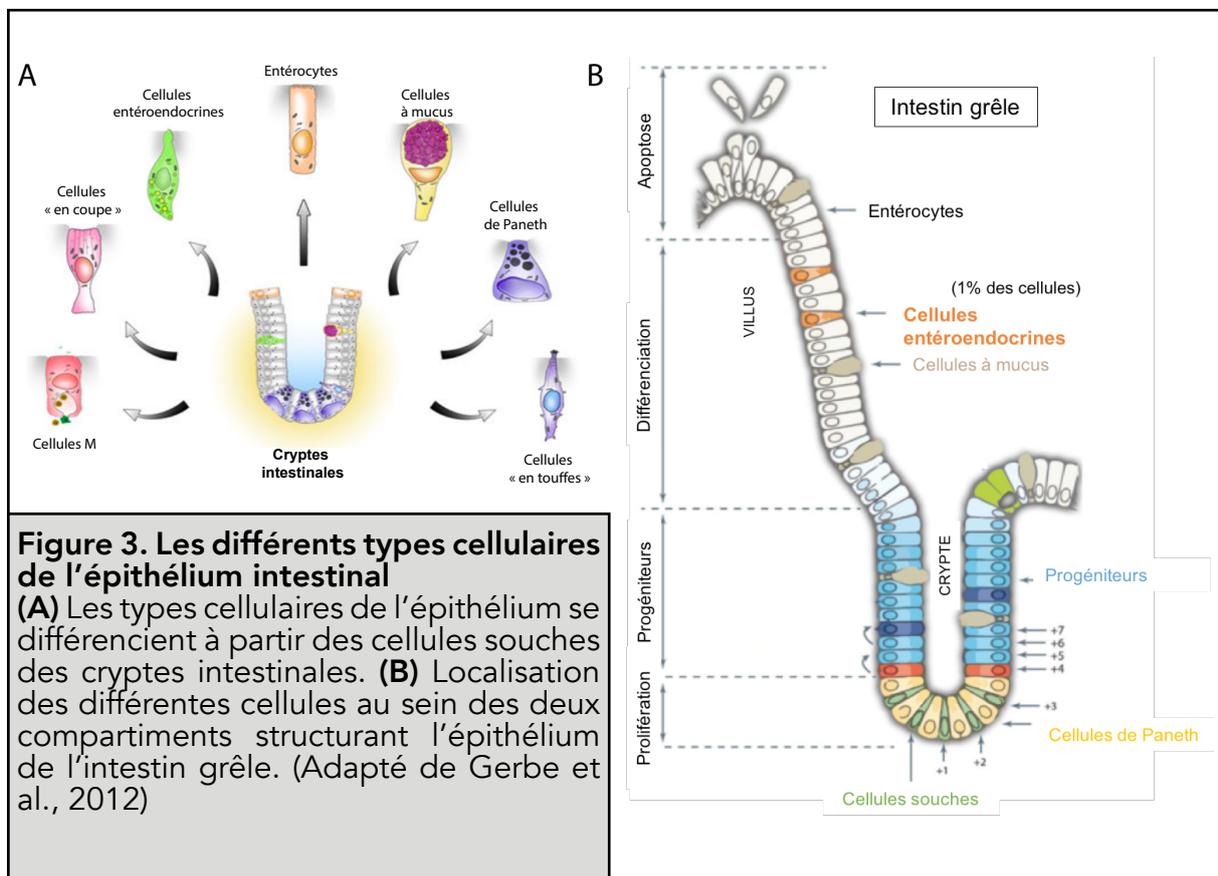
L'épithélium de l'intestin grêle présente une structure caractéristique (**Figure 2B**) avec la présence de deux compartiments morphologiquement et fonctionnellement distincts : le compartiment prolifératif des cryptes et le compartiment différencié des villosités. Le compartiment des cryptes de Lieberkühn contient les cellules souches intestinales dont descendent toutes les cellules de l'épithélium, des cellules en prolifération et en voie de différenciation et une population de cellules différenciées, les cellules de Paneth, localisées au fond des cryptes. Le compartiment des villosités est lui, uniquement composé de cellules différenciées post-mitotiques.



### Figure 2. Histologie du tractus digestif

(A) Différences de structure de la paroi du tube digestif selon l'axe proximo-distal : de l'œsophage à l'anus. Les différences histologiques majeures sont observées au niveau de la structure de l'épithélium : il est lisse dans l'œsophage et plus ou moins vallonné dans l'estomac avec des petits orifices correspondant à l'abouchement des glandes gastriques. Dans l'intestin grêle, l'épithélium est constitué de cryptes et de villosités, alors que l'épithélium du côlon est constitué uniquement de cryptes. (Illustration adaptée de Carole Fumat) (B) Structure tridimensionnelle détaillée de la paroi de l'intestin grêle. Les quatre couches de la paroi intestinale sont représentées : la muqueuse, la sous-muqueuse, la musculieuse et la séreuse. (Adapté de Anatomy & Physiology : From Science To Life, Jenkins, 2012)

On y retrouve les cellules absorbantes appelées entérocytes et des cellules sécrétrices comprenant les cellules à mucus et les cellules entéroendocrines (**Figure 3A**) (Cheng et Leblond, 1974). De façon remarquable, l'épithélium intestinal a la capacité de se renouveler intégralement tous les 4 à 5 jours chez les mammifères et ce, tout au long de la vie (Wong et al., 1999). Ce processus est rendu possible par la présence des cellules souches à la base des cryptes qui donnent naissance à des cellules progénitrices qui prolifèrent, puis migrent en direction des villosités et quittent le cycle cellulaire au cours de leur maturation et différenciation terminale. Quand les cellules atteignent l'apex des villosités, elles sont éliminées par apoptose dans la lumière intestinale (**Figure 3B**) (Crosnier et al., 2006). C'est cette structure bien caractéristique de l'épithélium de l'intestin grêle ainsi que la présence de microvillosités à la surface des cellules épithéliales qui permet d'augmenter la surface d'absorption, facilitant ainsi la digestion. L'épithélium fournit à l'intestin des caractéristiques topographiques spécifiques qui varient le long de l'axe antéro-postérieur (A-P) traduisant les rôles différents des segments du tube digestif et leur contribution à la digestion, absorption et sécrétion. L'épithélium du gros intestin et de l'estomac, par exemple, ne possède pas de villosités mais des invaginations apparentées aux cryptes et respectivement appelées glandes intestinales et glandes gastriques (**Figure 2A**).



### Figure 3. Les différents types cellulaires de l'épithélium intestinal

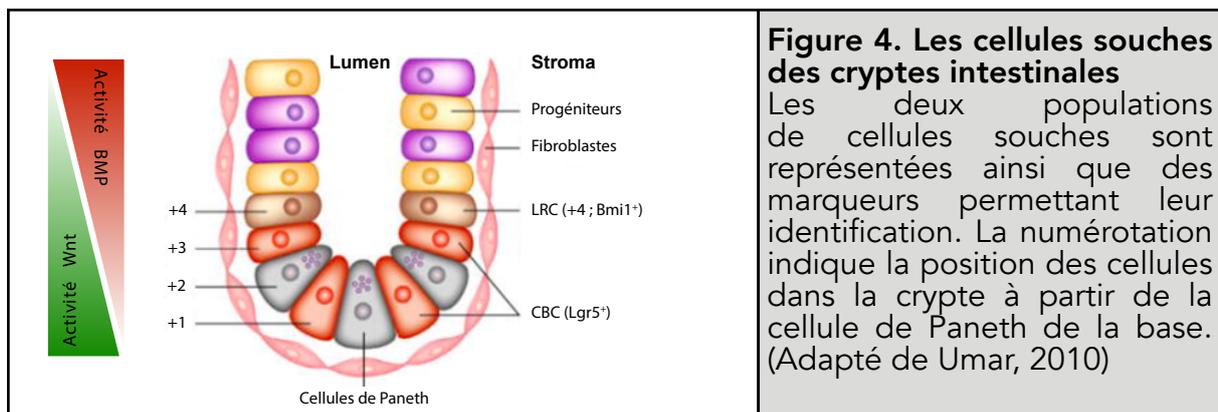
(A) Les types cellulaires de l'épithélium se différencient à partir des cellules souches des cryptes intestinales. (B) Localisation des différentes cellules au sein des deux compartiments structurant l'épithélium de l'intestin grêle. (Adapté de Gerbe et al., 2012)

### 1.3. Les types cellulaires de l'épithélium intestinal

Les cellules épithéliales qui peuplent l'épithélium intestinal sont polaires avec un côté apical dirigé vers la lumière intestinale et une surface basolatérale séparée par des jonctions intercellulaires, critiques dans le maintien de cette polarité. Cette compartimentation permet le transport des ions et des nutriments, ce qui favorise l'absorption et leur octroie la capacité de détecter le contenu luminal (Madara et al., 1990). Les différents types de cellules épithéliales ont des fonctions différentes mais proviennent d'une population de cellules multipotentes communes : les cellules souches. En effet, les cellules souches présentes dans les cryptes intestinales donnent naissance à deux lignages cellulaires distincts : le lignage absorbant et le lignage sécréteur.

#### 1.3.1. Les cellules souches

Deux types de cellules souches ont été identifiés dans les cryptes intestinales (**Figure 4**). Il y a d'une part les cellules souches actives CBC (*Crypt Base Columnar*) localisées entre les cellules de Paneth au fond des cryptes, qui sont des cellules multipotentes capables d'auto-renouvellement et de différenciation en tous les types cellulaires épithéliaux (Ponder et al., 1985). Ces cellules ont été caractérisées par des approches de traçage cellulaire *in vivo* grâce au marqueur *leucine-rich repeat-containing G-protein coupled receptor 5* (*Lgr5*), gène cible de la voie de signalisation Wnt (Barker et al., 2007). C'est pourquoi les cellules CBC sont également appelées cellules  $Lgr5^+$ . D'autre part, il y a les cellules nommées « +4 », de par la position qu'elles occupent dans les cryptes (4<sup>ème</sup> cellule à partir du fond de la crypte) et qui sont qualifiées de *Label-Retaining Cells* (LRC), car elles sont capables de conserver les marqueurs incorporés dans leur ADN pendant la néogenèse des cryptes et le renouvellement cellulaire (Potten et al., 2002). Ces cellules sont maintenues dans un état quiescent par des interactions avec la niche qui régule la fonction et la survie des cellules souches (Barker, 2014). Le rôle de ces cellules souches sera détaillé en **3.1**. de cette partie.



### **1.3.2. Les cellules du lignage absorbant**

#### ***Les entérocytes***

Les entérocytes représentent 90% de la population épithéliale intestinale, ce qui en fait les cellules les plus abondantes. La surface apicale de ces cellules est constituée de microvillosités composées notamment de villine, fimbrine, actine et calmoduline, qui forment une bordure en brosse participant à l'augmentation de la surface d'absorption des nutriments (Mooseker et al., 1984). Ces cellules contribuent à la digestion grâce à la présence d'un tapis glycoprotéique, le glycocalyx, contenant diverses enzymes digestives catalysant les étapes finales de la digestion (Hooton, et al., 2015). Le patron d'expression de ces enzymes varie le long de l'axe A-P et crypte-villosité mais peut également être modulé par le régime alimentaire (Tanaka et al., 1998). La membrane des microvillosités comporte par ailleurs des transporteurs spécifiques assurant la digestion et le passage sélectif des nutriments dans les vaisseaux sanguins et lymphatiques du chorion (Hussain, 2014). Les nutriments dégradés par les enzymes de la bordure en brosse sont absorbés à travers la membrane apicale des entérocytes par des transporteurs sélectifs (Madara et al., 2011 ; Hopfer et al., 1977). Dans le côlon, ces cellules sont appelées les colonocytes.

### **1.3.3. Les cellules du lignage sécréteur**

#### ***Les cellules à mucus***

Les cellules à mucus, encore appelées cellules caliciformes ou « en gobelet », sont disséminées au sein de l'épithélium et produisent un mucus protecteur principalement composé de glycoprotéines (mucines) et dont la principale fonction est de protéger la muqueuse intestinale des enzymes digestives et des micro-organismes pathogènes, et de faciliter le transit du bol alimentaire (Paulus et al., 1993). Les gouttelettes de mucus sont localisées au pôle apical des cellules et sont libérées par exocytose. Ces cellules sont plus abondantes dans la partie distale du tractus digestif où la présence de mucus facilite le passage des fèces dans le côlon (Birchenough et al., 2016).

#### ***Les cellules de Paneth***

Les cellules de Paneth sont les seules cellules différenciées des cryptes intestinales et sont absentes dans le côlon. Contrairement aux autres types cellulaires, elles migrent dans les cryptes, s'intercalant ainsi au sein des cellules souches et sont éliminées par phagocytose. Leur renouvellement est moins fréquent que les autres cellules de l'épithélium et se fait environ tous les deux mois (Cheng et al., 1969). Comme les cellules à mucus, leurs granules de sécrétion sont localisés au pôle apical et elles ont un rôle protecteur de la muqueuse intestinale et plus précisément des cellules

souches, en sécrétant des composés antimicrobiens dans la lumière intestinale, comme les lysozymes, les facteurs *trefoil* (« en feuille de trèfle ») et les défensines (Ganz, 2000). De plus, les peptides antimicrobiens peuvent déterminer la composition du microbiote intestinal (Salzman et al., 2011). Des analyses d'expression de gènes indiquent que les cellules de Paneth produisent EGF, Wnt3 et des ligands Notch, facteurs importants dans la prolifération des cryptes, suggérant un rôle de ces cellules dans la régulation de la niche des cellules souches (Cheng et al., 1969). Enfin, en plus de leur fonction protectrice, ces cellules sont capables de phagocyter les organismes pathogènes qui sont ensuite transférés aux lysosomes pour être dégradés (Erlandsen et al., 1972).

### **Les cellules entéroendocrines**

Les cellules entéroendocrines représentent seulement 1% des cellules épithéliales intestinales et forment pourtant le principal organe endocrine de l'organisme, de par l'importante surface que constitue l'épithélium du tractus digestif. Les cellules entéroendocrines sont des régulateurs clés du statut nutritif et de la signalisation métabolique. En effet, elles sécrètent des hormones qui contrôlent l'homéostasie énergétique en régulant notamment la prise alimentaire, la digestion et l'assimilation du glucose (Posovszky et Wabitsch, 2014). Ces hormones sont contenues dans des vésicules de sécrétion localisées dans la partie basale de la cellule et sont sécrétées en réponse à des stimuli variés (nutriments, microbiote, etc.). La plupart des cellules entéroendocrines possèdent des microvillosités sur leur surface apicale qui leur permettent de sonder le contenu luminal du tube digestif (Gribble et Reimann, 2016). Il a récemment été montré que certaines cellules entéroendocrines possèdent des extensions de leur surface basolatérale, appelées neuropodes, qui pourraient servir d'intermédiaire dans la communication entre ces cellules et le système nerveux (Bohòrquez et al., 2015). Les hormones agissent à distance via la circulation sanguine (action endocrine), localement dans le tissu intestinal (action paracrine) ou encore par interaction avec le système nerveux. Il en existe au moins 12 sous-types différents définis par leur morphologie, leur patron d'expression le long de l'axe proximo-distal de l'intestin et leur principale hormone sécrétée (Furness et al., 2013). Leur répartition le long du tractus digestif dépend du rôle des hormones qu'elles produisent. Leurs fonctions seront détaillées dans la **partie 2** de l'introduction.

#### **1.3.4. Autres types cellulaires**

##### **Les cellules M**

Les cellules M ou cellules à microplis se trouvent au niveau des tissus lymphoïdes associés au tube digestif, notamment dans les plaques de Peyer. Ce sont des cellules

présentant des microvillosités réduites et spécialisées pour la phagocytose et la transcytose de macromolécules présentes dans la lumière intestinale. En effet, elles transportent des antigènes de la lumière aux cellules présentatrices d'antigènes, initiant ainsi une réponse immunitaire (Mabbott et al., 2013).

### **Les cellules « en touffe »**

Aussi appelées *tuft/brush/calveolated cells*, elles représentent une population rare (0,4 à 2%) préférentiellement localisées dans les villosités et à la jonction crypte-villosité. Elles sont reconnaissables à leur aspect fusiforme caractéristique et leurs microvillosités abondantes. Leur fonction est peu décrite mais l'expression de protéines et récepteurs connus pour leur implication dans la gustation suggère un rôle de cellules sensorielles lumineales. Il a également récemment été montré que ces cellules sont capables de déclencher une réaction immunitaire de type 2 médiée par la production d'interleukine-25 menant à l'expulsion des pathogènes (Gerbe et al., 2012, 2016).

### **Les cellules « en coupe »**

Plus connues sous le nom de *cup cells*, ces cellules sont restreintes à l'épithélium de l'iléon et constituent jusqu'à 6% des cellules totales. Elles sont caractérisées par une petite bordure en brosse et une activité phosphatase alcaline réduite (Madara, 1982). Néanmoins, leur rôle précis n'est pas encore connu.

## **2. Morphogenèse intestinale**

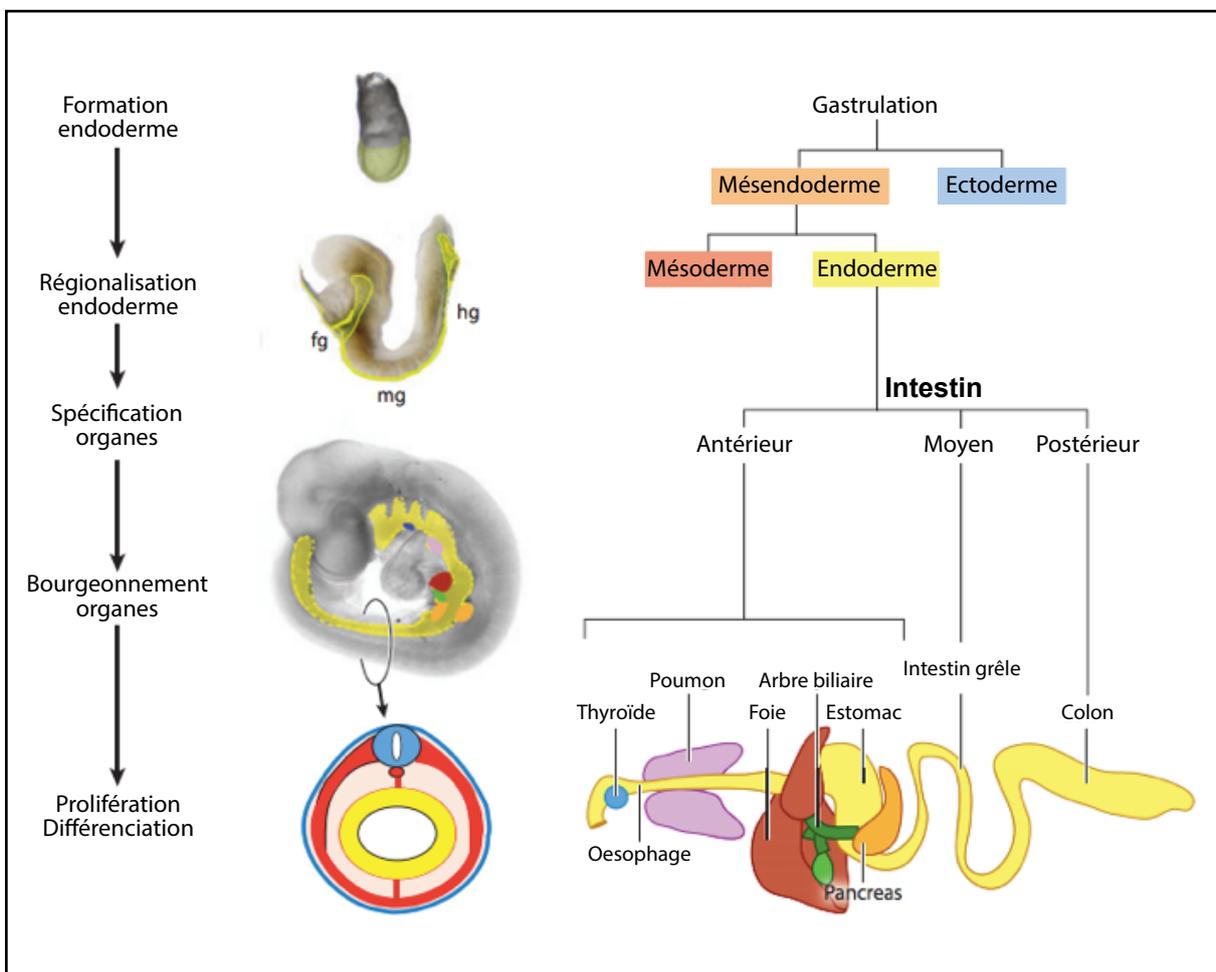
La morphogenèse du tube digestif est un processus relativement conservé entre les différentes espèces de vertébrés classiquement étudiés (souris, xénope, poulet et poisson-zèbre) mais par souci de clarté, les données présentées dans cette partie sont relatives aux études menées chez la souris. La mise en place des trois feuilletts embryonnaires que sont l'endoderme, le mésoderme et l'ectoderme a lieu pendant la gastrulation (**Figure 5**). L'endoderme se développe à partir des invaginations antérieure (*Anterior Intestinal Portal, AIP*) et postérieure (*Caudal Intestinal Portal, CIP*) de l'embryon. Ces invaginations progressent de l'extrémité vers la partie moyenne de l'endoderme, achevant ainsi la fermeture de l'endoderme (Wells et Melton, 1999). L'endoderme et le mésoderme semblent dériver d'une population cellulaire commune, appelée mésendoderme (Kimelman et Griffin, 2000). Suite à la gastrulation, une série d'évènements morphogénétiques permet à l'endoderme naïf de se transformer en tube digestif primitif entouré de mésoderme (**Figure 5**). C'est à ce moment que le tube digestif se régionalise dans les axes dorso-ventral (D-V) et antéro-postérieur (A-P) pour former l'intestin primitif antérieur, postérieur et moyen :

- l'intestin primitif antérieur est notamment à l'origine de la cavité buccale, du pharynx, de l'œsophage, de l'estomac, d'une partie du duodénum, du foie, des voies biliaires et du pancréas;

-l'intestin primitif moyen est notamment à l'origine de la partie terminale du duodénum, du jéjunum, de l'iléon, du caecum, de l'appendice et d'une partie du côlon;

- l'intestin primitif postérieur est à l'origine du reste du côlon, du rectum et de l'anus.

Les bourgeons d'organe sont constitués de l'épithélium endodermique entremêlé avec le mésenchyme environnant, qui prolifèrent et se différencient pour donner naissance aux organes fonctionnels.



**Figure 5. Schéma récapitulant les grandes étapes de la formation des organes d'origine endodermique au cours du développement embryonnaire**

L'ensemble du tractus digestif dérive de l'endoderme qui est l'un des trois feuillets embryonnaires formés au moment de la gastrulation. La formation des différents organes implique consécutivement une régionalisation de l'endoderme, puis une spécification des organes menant à la formation de bourgeons d'organes, dont les cellules vont proliférer et se différencier pour donner naissance aux organes matures. fg: foregut (intestin antérieur); mg: midgut (intestin moyen); hg: hindgut (intestin postérieur) ; bleu: ectoderme; rouge : mésoderme; jaune : endoderme. (Adapté de Zorn et Wells, 2009)

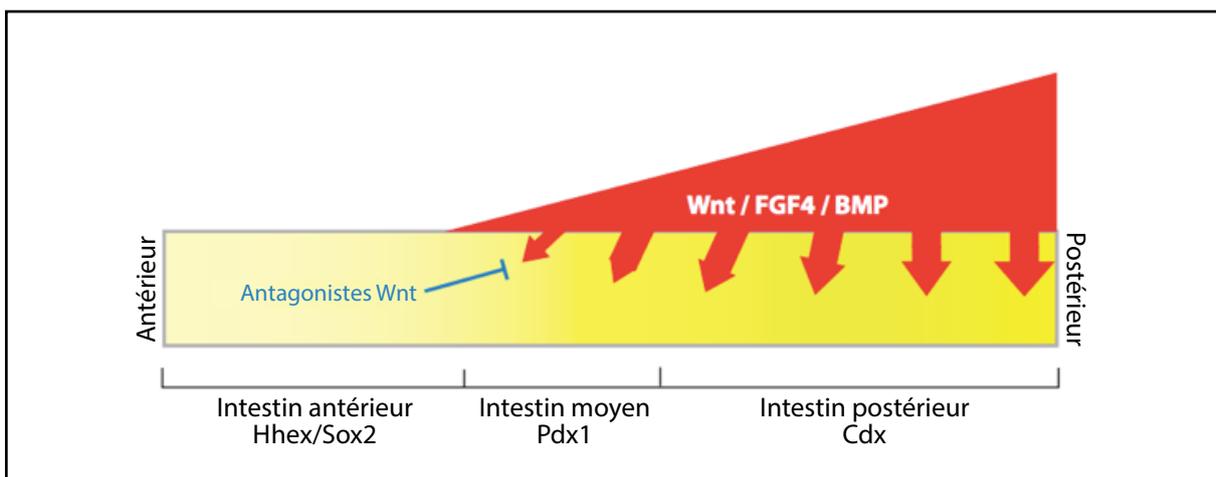
En effet, l'intestin résulte de la fusion de l'endoderme qui donnera naissance à l'épithélium et du mésoderme splanchnique duquel dérive les couches musculaires lisses et les tissus conjonctifs qui viendront s'associer en périphérie de l'endoderme. A l'interface de ces deux feuillets se forme la lame basale composée essentiellement de collagène IV, laminines et protéoglycanes (Aumailley et Verrando, 1993). L'endoderme devient un épithélium pseudostratifié qui s'organise progressivement en épithélium monostratifié et se replie en villosités à la fin du développement embryonnaire. Ces phénomènes sont initiés dans la partie antérieure de l'intestin et se propagent vers la partie postérieure. Peu de temps avant la naissance, les cellules en prolifération se concentrent dans les zones intervillaires qui s'invagineront en cryptes durant la première semaine post-natale (Simon-Assmann et Kedinger, 1993). Ces processus sont contrôlés par plusieurs voies de signalisation et facteurs de croissance incluant FGF, BMP, Wnt, l'acide rétinoïque, Hedgehog, et Notch, qui ont plusieurs rôles spécifiques au stade de développement (Crosnier et al., 2006). Le système nerveux entérique, lui, dérive des cellules des crêtes neurales. Les précurseurs de ces cellules entrent dans le tractus digestif au niveau du pharynx et migrent à travers la couche périphérique du mésenchyme digestif. Durant cette migration, ils prolifèrent et se différencient en neurones et en cellules gliales (Nagy et Goldstein, 2017). Ces cellules constituent les plexus neuronaux des couches de muscles lisses et de la région périphérique de la sous-muqueuse.

## 2.1. Formation de l'endoderme

Chez tous les vertébrés, le développement de l'endoderme et du mésoderme est initié pendant la gastrulation par la voie de signalisation Nodal, qui par la suite est requise pour la régionalisation axiale de l'endoderme. Les ligands de cette voie font partie de la famille des facteurs de croissance TGF- $\beta$ . Il est globalement admis qu'un niveau d'expression élevé de Nodal promeut la formation de l'endoderme via l'expression de Hhex, alors qu'un niveau plus faible induit la formation du mésoderme (Zorn et Wells, 2007 ; D'Amour et al., 2005). Chez la souris, l'endoderme définitif dérive de la ligne primitive antérieure qui est le site d'expression de Nodal. Dans toutes les espèces, les cellules endodermiques se développent à proximité de la source de Nodal, dont l'expression au moment de la gastrulation est stimulée par la voie canonique Wnt, via la  $\beta$ -caténine et les facteurs de transcription TCF (Lickert et al., 2002). La signalisation Wnt est requise pour maintenir l'expression de Nodal, qui est à son tour requis pour maintenir l'expression de Wnt.

## 2.2. Régionalisation de l'endoderme

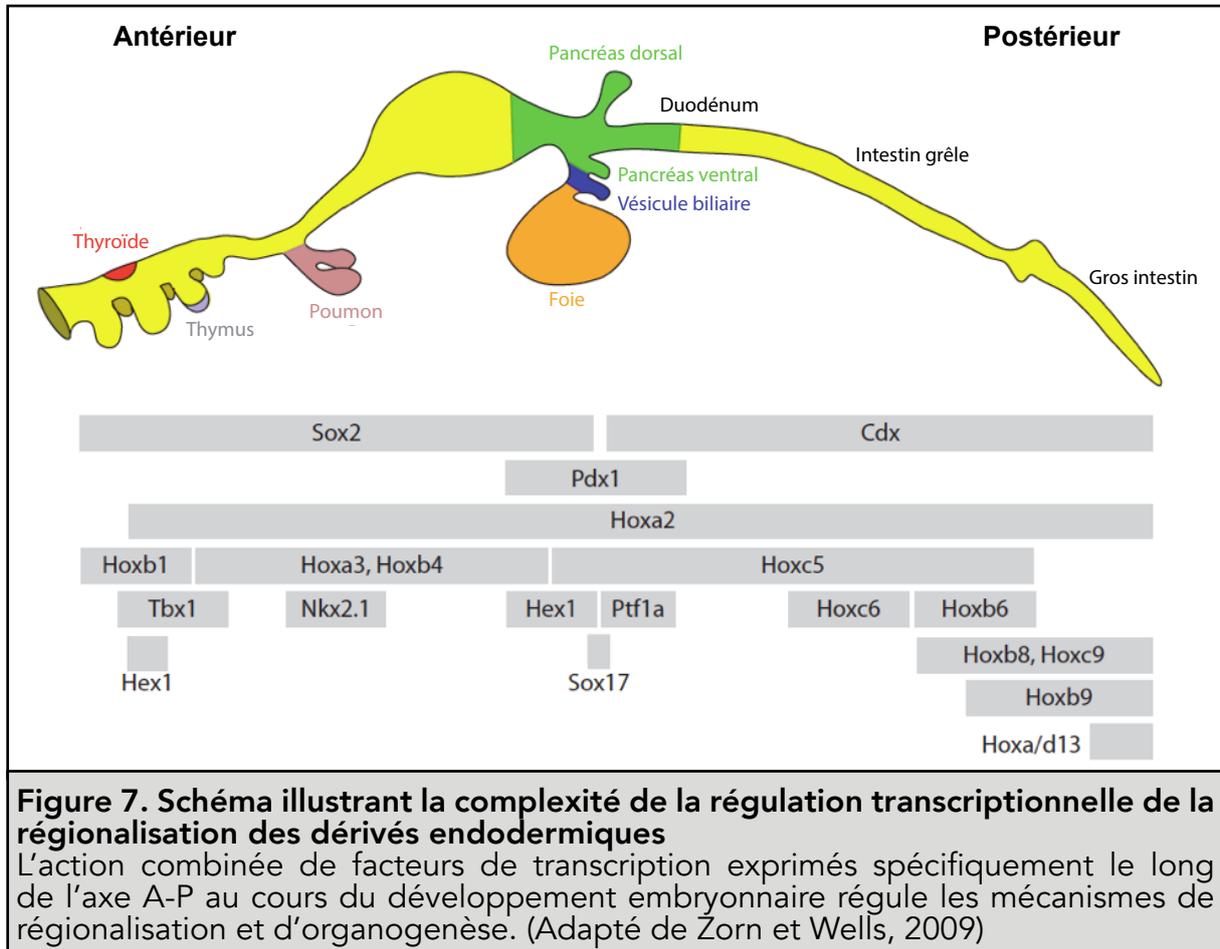
Pendant les deux jours consécutifs à la gastrulation, le feuillet endodermique donne naissance à l'intestin primitif dont émergent les bourgeons d'organes du tractus digestif. Les premiers signes de régionalisation sont détectés dès le stade gastrula lors de la spécification de la couche endodermique. Chez la souris, la forte expression de Nodal au niveau de la ligne primitive promeut d'abord une différenciation en endoderme antérieur. Par la suite, la partie postérieure de la ligne primitive exprime faiblement Nodal, ce qui induit plutôt la formation d'endoderme postérieur. Les facteurs de transcription *Foxa2* et *Mixl1* sont particulièrement impliqués dans la formation d'endoderme antérieur, tandis que *Sox17* est nécessaire pour la formation d'endoderme postérieur (Stainier, 2002). D'autres membres de la famille des facteurs de transcription Sox sont exprimés dans l'endoderme intestinal (particulièrement *Sox2*, *Sox7*, *Sox9*, *Sox17* et *Sox18*). Bien que leur rôle exact ne soit pas connu, leur fonction semble indispensable au développement intestinal (Fu et Shi, 2016). Pendant la gastrulation et les stades somites précoces, une signalisation différentielle le long de l'axe A-P impliquant les facteurs Wnt, FGF4 et BMP, exprimés par le mésoderme, provoque une régionalisation de l'endoderme en intestin primitif antérieur, moyen et postérieur (**Figure 6**). Ces régions expriment respectivement les facteurs de transcription *Hhex* et *Sox2*, *Pdx1* et *Cdx*. L'acide rétinoïque (AR), agent soluble métabolisé dans le mésoderme à partir de la vitamine A et dont les gènes cibles sont localisés dans l'endoderme, semble participer au contrôle de la régionalisation A-P de l'endoderme mais la nécessité de l'AR dans la différenciation des organes diffère selon leur position le long de cet axe (Bayha et al., 2009).



**Figure 6. Facteurs impliqués dans la régionalisation de l'intestin primitif au cours du développement embryonnaire**

Les ligands Wnt, FGF et BMP sont exprimés graduellement par le mésoderme (rouge) signalant à l'endoderme adjacent (jaune) la répression de l'identité antérieure (*Hhex/Sox2*) et promouvant la formation d'intestin postérieur (*Cdx*). La région intermédiaire constitue l'intestin moyen (*Pdx1*). L'identité de l'endoderme antérieur est maintenue par la sécrétion d'inhibiteurs Wnt. (Adapté de Zorn et Wells, 2009)

L'expression mésodermique de gènes *Hox* spécifiques joue un rôle important dans la régionalisation A-P du tube digestif en influençant également la morphologie intestinale et les futures interactions entre l'épithélium et le mésenchyme, nécessaires à la différenciation épithéliale de l'intestin. Ces gènes sont exprimés de manière spécifique dans le temps et l'espace (**Figure 7**) (De Santa Barbara et al., 2003).



### 2.3. Formation des organes

Au cours du stade de formation des somites, des interactions continues entre le mésenchyme et l'intestin primitif provoquent une spécification des différents organes. Ces interactions se font via un panel de signaux cellulaires qui, par leurs actions combinées et localisées, induisent la spécification des différents organes. En effet, il semble qu'il n'y ait pas une correspondance directe entre un facteur donné et un lignage spécifique, mais plutôt des combinaisons complexes de facteurs avec un chevauchement partiel de leur expression le long du tractus gastro-intestinal (**Figure 7**). La spécification de l'estomac est dépendante du facteur de transcription à homéodomaine *Barx1* qui est exprimé par le futur mésenchyme gastrique. Celui-ci inhibe localement mais indirectement la voie de signalisation *Wnt* pour réprimer

l'expression de Cdx dans l'épithélium et ainsi favoriser la différenciation en estomac (Kim et al., 2005). Ainsi, la détermination de la limite estomac/intestin est dépendante d'un antagonisme mutuel entre les voies Barx1/Sox2 (pro-estomac) et Wnt/Cdx2 (pro-intestin) (Stringer et al., 2008). La transition d'un simple épithélium pseudostratifié en une muqueuse intestinale mature avec une structure cryptes-villosités se fait selon l'axe A-P du tractus gastro-intestinal. La spécification cellulaire et la formation des villosités est sous la dépendance des facteurs Wnt, BMP et Cdx2, menant à tous les types cellulaires évoqués précédemment (**voir 1.3.**) (Rubin, 2007). Les mécanismes moléculaires régulant la morphogenèse et la différenciation de l'épithélium intestinal seront détaillées en **3.3.** et **4.** de cette partie.

### **3. Différenciation et renouvellement de l'épithélium intestinal**

#### **3.1. Les cellules souches intestinales : deux modèles décrits**

L'intestin représente le tissu qui se renouvelle le plus rapidement chez les mammifères adultes. En effet, l'épithélium de l'intestin grêle adulte est renouvelé tous les 4 à 5 jours, à l'exception des cellules de Paneth qui sont renouvelées tous les deux mois, grâce aux cellules souches dans le compartiment prolifératif des cryptes intestinales. Ces cellules souches sont à l'origine d'une population de cellules hautement prolifératives qui se divisent rapidement et se différencient en tous les types cellulaires de l'épithélium décrits précédemment (Crosnier et al., 2006). Il est important pour le renouvellement et la maintenance de l'épithélium intestinal qu'un équilibre entre prolifération, différenciation, migration et mort cellulaire soit établi. Cet équilibre est notamment régulé par un dialogue constant entre l'épithélium et le mésenchyme associé. Les cellules souches intestinales sont qualifiées de multipotentes puisqu'elles peuvent donner naissance à tous les types de cellules spécialisées de l'épithélium. Ces cellules souches adultes agissent en tant que système de réparation et de maintenance tout au long de la vie de l'organisme. Elles ont à la fois recours à la division asymétrique et symétrique et sont quiescentes dans un microenvironnement qualifié de niche, qui associe des cellules épithéliales, des myofibroblastes sous-épithéliaux et des substrats extracellulaires qui sont autant de facteurs nécessaires au maintien des cellules souches (Meran et al., 2017).

Deux populations de cellules souches intestinales ont été décrites dans la littérature: la première constituée de cellules souches « actives », responsable du maintien de l'homéostasie épithéliale en condition physiologique ; et la deuxième qui est une population de « réserve », capable de faciliter la régénération de l'épithélium en cas de dommages (Cui et Chang, 2016). La population « active » correspond aux cellules CBC (*Crypt Base Columnar*) entremêlées avec les cellules de Paneth au fond

des cryptes. La population de réserve correspond aux cellules « +4 », aussi appelée LRC (*Label-Retaining Cells*) se trouvant au-dessus des cellules de Paneth à la position +4 à partir de la base des cryptes.

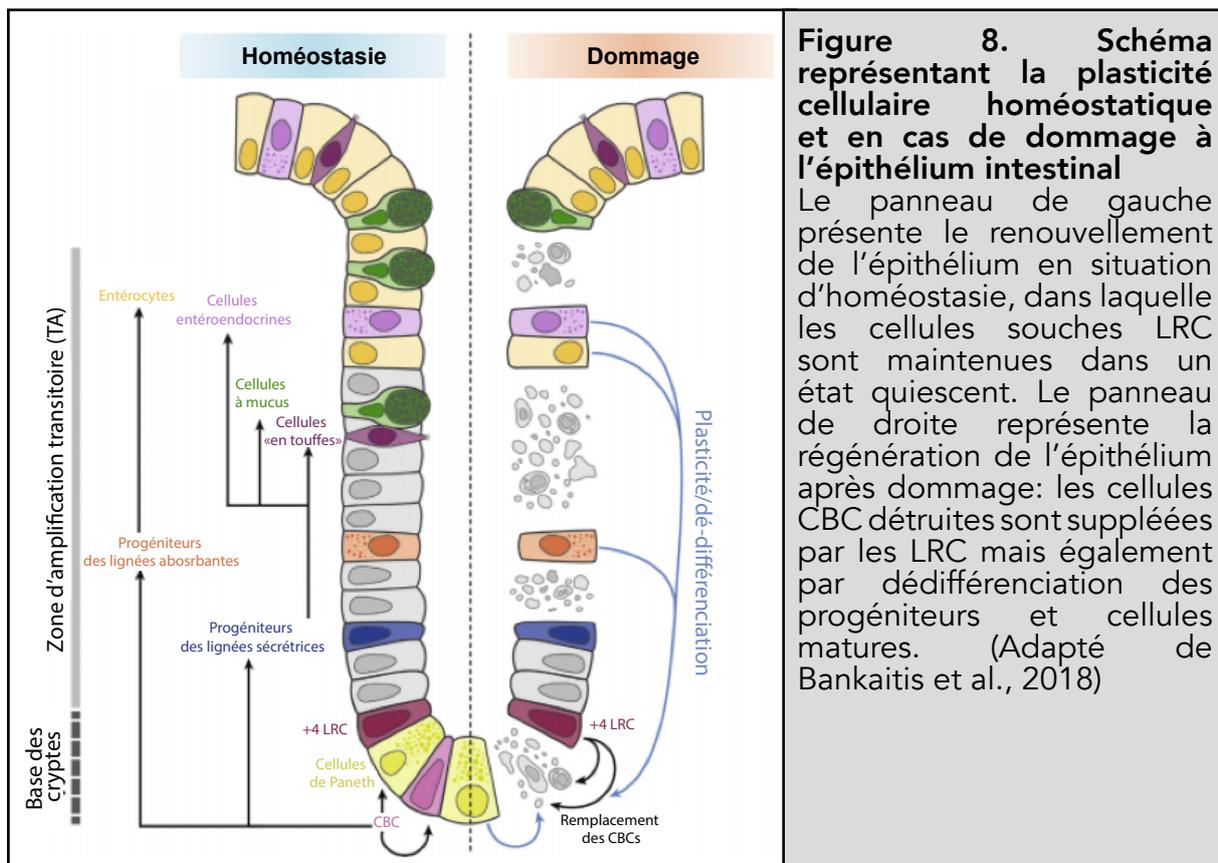
Les cellules LRC sont faiblement prolifératives mais ont la capacité de générer l'ensemble des types cellulaires de l'épithélium. L'expression de *Bmi1* et *Tert* permet l'identification de cette population spécifique de cellules souches, dont la distribution suit un gradient décroissant du duodénum à l'iléon et est absente dans le côlon. Les cellules CBC, présentes dans tous les segments intestinaux, sont identifiées par l'expression du marqueur *Lgr5* et sont capables de générer tous les types de cellules différenciées de l'épithélium. Cette caractéristique a été confirmée lorsqu'une telle cellule isolée a été remise en culture dans des conditions favorables et a permis la formation d'organoïdes intestinaux *in vitro*, récapitulant la morphologie anatomique et fonctionnelle de l'épithélium intestinal (Sato et al., 2009). Une analyse transcriptomique des cellules *Lgr5*<sup>+</sup> a ensuite permis d'identifier d'autres marqueurs de ces cellules comme *Olfm4* et *Ascl2* (Munoz et al., 2012). Les cellules CBC sont mitotiquement actives et sont sensibles aux modulations de la voie canonique Wnt. En effet, *Lgr5* est le récepteur de la R-Spondine et il amplifie la signalisation Wnt/ $\beta$ -caténine (Carmon et al., 2011). Les CBC contribuent à la régénération homéostatique et sont éliminées après irradiation. En revanche, les cellules LRC sont quiescentes et insensibles aux perturbations de la voie Wnt. Elles contribuent très faiblement à la régénération homéostatique mais sont résistantes à des fortes doses d'irradiation (Kim et al., 2017; Yan et al., 2012).

### 3.2. Plasticité de l'épithélium intestinal

Le renouvellement homéostatique de l'intégralité de l'épithélium est assuré par les cellules CBC, cependant une ablation de ces cellules n'affecte pas la capacité de l'épithélium à se renouveler, indiquant l'existence d'une population de cellules souches intestinales de réserve (Tian et al., 2011). En effet, la destruction de ces cellules peut induire la reconversion de cellules engagées dans le lignage sécréteur à un état de cellules souches (van Es et al., 2012).

Ces observations ont remis en question le premier modèle stipulant que les cellules LRC en position +4 sont des cellules souches de réserve qui remplacent les cellules CBC en cas de dommage et suggèrent plutôt que ce sont des progéniteurs du lignage sécréteur qui ont un potentiel de cellule souche en cas de dommage. Il semblerait qu'en réalité la destruction des CBC déclenche le renouvellement épithélial par ces cellules LRC mais également par des progéniteurs des lignages sécréteur et absorbant et même des cellules différenciées (Yan et al., 2017; Tetteh et al. 2016; Van Es et al., 2012). Les LRC, initialement décrites comme cellules souches quiescentes

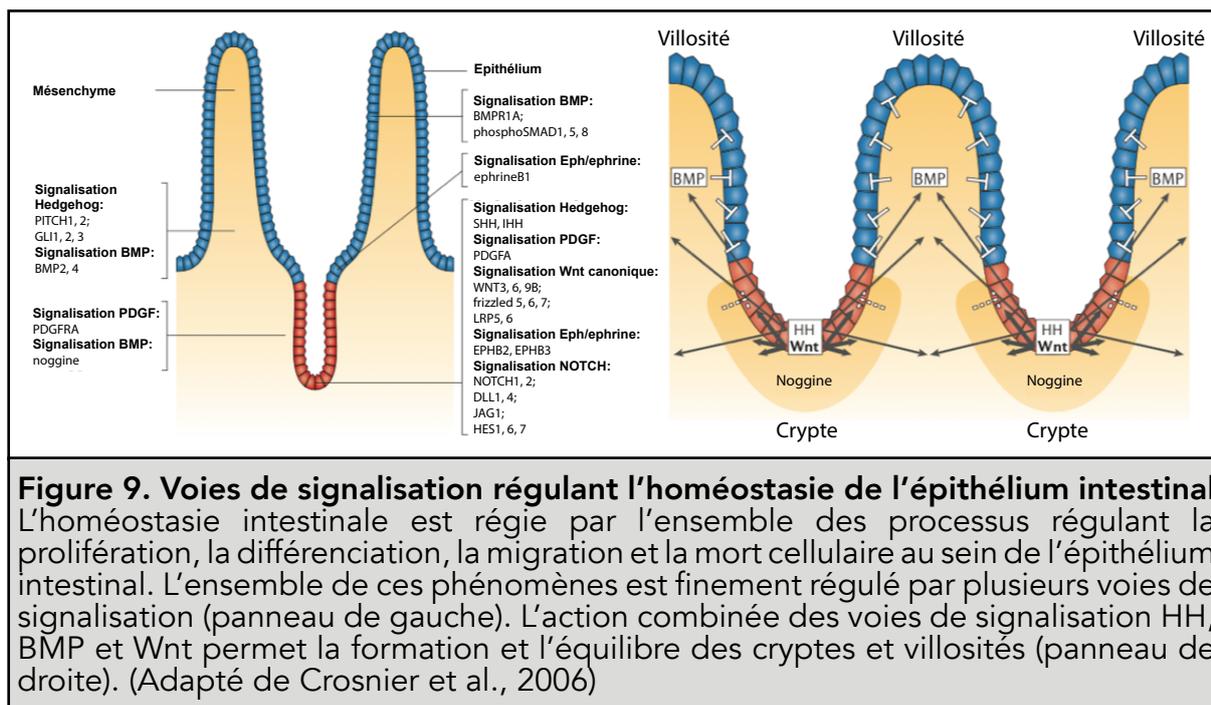
(Potten et al., 1998), sont maintenant reconnues comme des précurseurs des cellules de Paneth et des cellules entéroendocrines contribuant à la réparation de l'épithélium (Jadhav et al., 2017 ; Li et al., 2016 ; Buczacki et al., 2013). Les cellules recrutées en cas de destruction des CBC peuvent migrer dans la crypte et au contact des cellules de Paneth et d'une source de Wnt, se différencier en une population de cellules souches actives en prolifération (**Figure 8**) (Tetteh et al. 2016 ; van Es et al., 2012). La capacité de l'ensemble des lignages cellulaires de l'épithélium à répondre et à s'adapter à des dommages ou changements d'environnement est désignée par le terme de plasticité (Bankaitis et al., 2018).



### 3.3. Maintien de l'homéostasie intestinale

#### 3.3.1. Régulation de l'équilibre crypte/villosité

Au cours de la morphogénèse intestinale, l'épithélium s'organise en deux compartiments morphologiquement et fonctionnellement distincts : les cryptes et les villosités. La mise en place puis le maintien de cette organisation sont finement régulés par des interactions réciproques entre l'épithélium et le mésenchyme sous-jacent, impliquant notamment les signalisations Hedgehog et BMP (*Bone Morphogenetic Protein*) (Crosnier et al., 2006).



Au cours du développement de l'intestin grêle, les ligands *Sonic hedgehog* (Shh) et *Indian hedgehog* (Ihh) sont exprimés spécifiquement par l'épithélium intestinal. Lors de la morphogenèse des villosités, l'expression de Shh et Ihh se concentre progressivement aux zones intervallositaires (**Figure 9**). Similairement, l'expression des récepteurs Ptch1 et Ptch2 et effecteurs Gli1, Gli2 et Gli3 impliqués dans la signalisation Hedgehog se concentre au niveau du mésenchyme sous-jacent (Madison et al., 2005). Dans l'intestin grêle adulte, l'expression de Shh et Ihh persiste à la base des cryptes, site de prolifération cellulaire. Ces ligands diffusent dans le tissu pour agir à distance, promouvant la formation de villosités et inhibant la formation de cryptes à proximité de celles préexistantes, maintenant ainsi l'espacement adéquat entre les cryptes adjacentes. La signalisation BMP permet quant à elle le transfert d'informations du mésenchyme vers l'épithélium, régulant la formation de cryptes en interaction avec la voie Hedgehog. La voie BMP est activée (notamment par BMP2 et 4) au niveau des villosités, ce qui est observable par la présence des phosphoSMAD1,5 et 8 dans les cellules post-mitotiques. Cette signalisation est au contraire inhibée par l'expression de la noggine au voisinage des cryptes (Haramis et al., 2004 ; Karisson et al., 2000). Un dysfonctionnement de la voie BMP ou une surexpression de la noggine induisent une production excessive de cryptes (He et al., 2004). Ainsi, la signalisation BMP est le facteur clé régulant l'action de Hedgehog, empêchant la formation aberrante de cryptes.

### 3.3.2. Régulation de la prolifération et différenciation cellulaires

L'activation de la voie canonique Wnt, notamment par WNT3,6 et 9B, est déterminante pour le maintien de l'état prolifératif des cellules présentes au fond

des cryptes et l'inhibition de leur différenciation (Gregorief et al., 2005). En effet, ces cellules expriment les récepteurs associés (Frizzled5,6,7 et LRP5,6) et présentent une quantité intra-nucléaire importante de  $\beta$ -caténine, signe de l'activation de la voie canonique Wnt. Cette accumulation de  $\beta$ -caténine n'est en revanche pas observée dans les cellules des villosités. Une altération de la voie canonique Wnt peut mener à une perte des cellules en prolifération et une diminution du nombre et de la taille des cryptes (Pinto et al., 2003 ; Korinek et al., 1998).

La signalisation Wnt n'est pas la seule contrôlant la prolifération et la différenciation dans l'épithélium intestinal, en effet, la voie Notch est également déterminante dans ces processus. Notch semble particulièrement impliquée dans le contrôle de la prolifération des cellules transitoires des cryptes. Une perte de fonction de Notch engendre une diminution ou une perte de la prolifération des cellules transitoires, qui sont remplacées par des cellules post-mitotiques, sans que la signalisation Wnt n'en soit affectée (van Es et al., 2005). En plus de leur rôle dans le contrôle de la prolifération, Wnt et Notch sont également impliqués dans les processus de différenciation des cellules de l'épithélium. En effet, Wnt régule la différenciation des cellules de Paneth dans les cryptes et bien que la différenciation des autres cellules s'opère dans des régions où la voie canonique Wnt n'est pas activée, Wnt est impliquée dans la décision du lignage cellulaire. Effectivement, une altération de la signalisation Wnt n'affecte pas la différenciation des cellules absorbantes mais provoque une perte des cellules des lignées sécrétrices. Notch semble avoir un rôle opposé dans le contrôle de la différenciation : une inactivation de Notch provoque une production excessive de cellules entéroendocrines et de cellules à mucus et une perte de cellules absorbantes (van Es et al., 2005 ; Jensen et al., 2000).

Comme cela a été évoqué plus tôt, les cellules de l'épithélium se séparent en deux catégories, présentes dans des régions distinctes : les cellules en prolifération dans les cryptes (sous l'effet de Wnt) et les cellules différenciées dans les villosités. Cette ségrégation s'opère par l'action de Wnt sur l'éphrine B1 et ses récepteurs EphB2 et 3 (**Figure 9**). Les éphrines et leurs récepteurs sont des protéines membranaires impliquées dans les interactions entre les cellules adjacentes. De façon générale, les mécanismes de répulsion cellulaire médiés par les éphrines et leurs récepteurs sont impliqués dans la création de délimitations tissulaires et la définition des voies de migration cellulaire (Xu et al., 1999 ; Winslow et al., 1995). Dans l'intestin, la signalisation Wnt induit l'expression des récepteurs EphB2 et 3 et inhibe l'expression de l'éphrine 1 (van der Wetering et al., 2002 ; Battle et al., 2002). Ainsi, dans l'épithélium mature, l'éphrine B1 est exprimée spécifiquement à la jonction crypte-villosité tandis que EphB2 et 3 sont exprimés dans les cryptes. Ces mécanismes assurent la ségrégation des populations cellulaires épithéliales, notamment en confinant les cellules en prolifération dans les cryptes.

## 4. Description des principales voies de signalisation moléculaire impliquées dans le développement et l'homéostasie intestinale

Au cours du développement embryonnaire puis à l'âge adulte, la morphogenèse intestinale, la différenciation des sous-types cellulaires et l'homéostasie intestinale sont des processus finement régulés par de nombreuses voies de signalisation et dont les interactions complexes ne sont pas encore toutes élucidées. Cette partie vise à présenter les principales voies impliquées dans ces phénomènes et nécessaires à la compréhension de ce manuscrit de thèse.

### **Wnt**

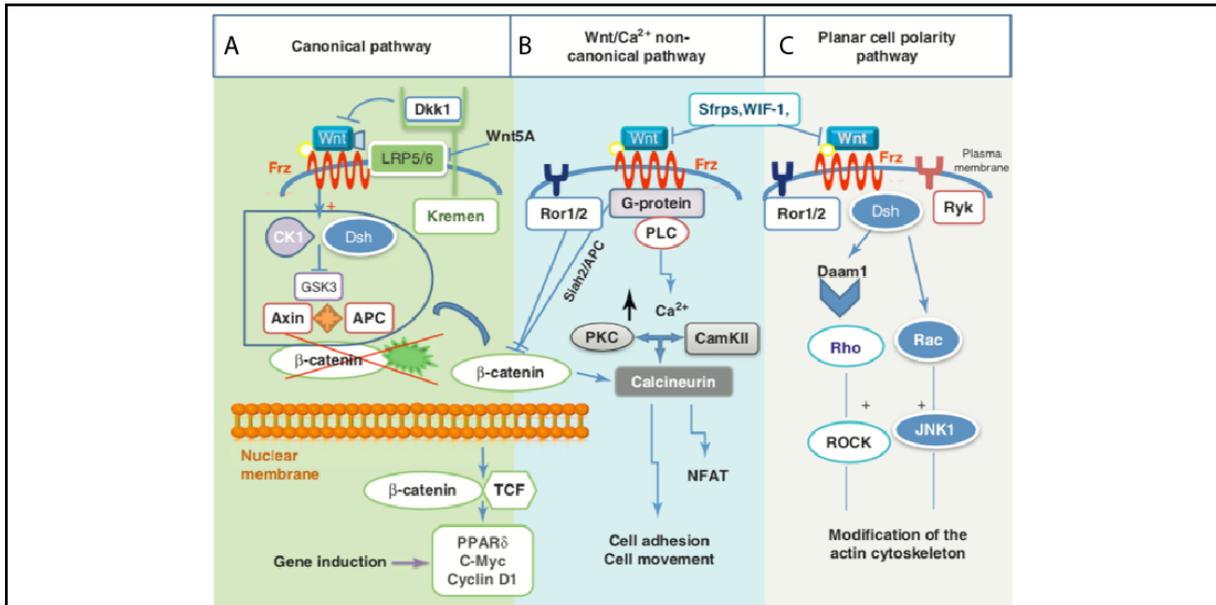
Les voies de signalisation impliquant les ligands glycoprotéiques sécrétés riches en cystéine Wnt régulent de nombreux aspects du développement et en particulier la formation et la régionalisation de l'endoderme, ainsi que l'organogenèse. Les ligands Wnt peuvent intervenir dans différentes cascades de signalisation intracellulaire:

- La voie canonique Wnt/ $\beta$ -caténine (**Figure 10A**) : elle implique une interaction entre le ligand Wnt et un complexe de récepteurs Frizzled (Fz) et LRP5/6 qui provoque la stabilisation et l'accumulation dans le noyau de la  $\beta$ -caténine, effecteur intracellulaire. En l'absence de ligand Wnt, la  $\beta$ -caténine cytoplasmique subit une série de phosphorylations par le complexe de dégradation Axin/APC/GSK3 $\beta$  aboutissant à sa dégradation par la voie du protéasome. Lors de la fixation du facteur sécrété Wnt sur son récepteur Frizzled, GSK3 $\beta$  est inhibée, la  $\beta$ -caténine n'est plus dégradée et peut fixer le *T Cell Factor/Lymphoid-Enhancing Factor* (TCF/LEF), initialement inhibiteur transcriptionnel, pour le convertir en activateur de la transcription des gènes cibles (Sancho et al., 2004);

- La voie non-canonique Wnt/Ca<sup>2+</sup> (**Figure 10B**) : elle implique aussi les récepteurs Fz, qui par l'action de protéines G, mobilisent les stocks de calcium intracellulaire pour activer la Cam kinase. Cette voie est notamment impliquée dans la régulation de l'adhésion et de la migration cellulaire;

- La voie non-canonique Wnt/PCP (*Planar Cell Polarity*) (**Figure 10C**): cette voie implique également les récepteurs Fz qui activent les GTPases Rho et Rac, ce qui aboutit notamment à la modification de la structure du cytosquelette d'actine.

Toutes ces voies de signalisation sont impliquées dans la dynamique du cytosquelette et la morphogenèse cellulaire (la polarité des cellules épithéliales, la migration et l'adhésion cellulaire). La formation de l'endoderme est dépendante de la voie Wnt et son induction au cours de la gastrulation dépend de Wnt3 (Liu et al., 1999).

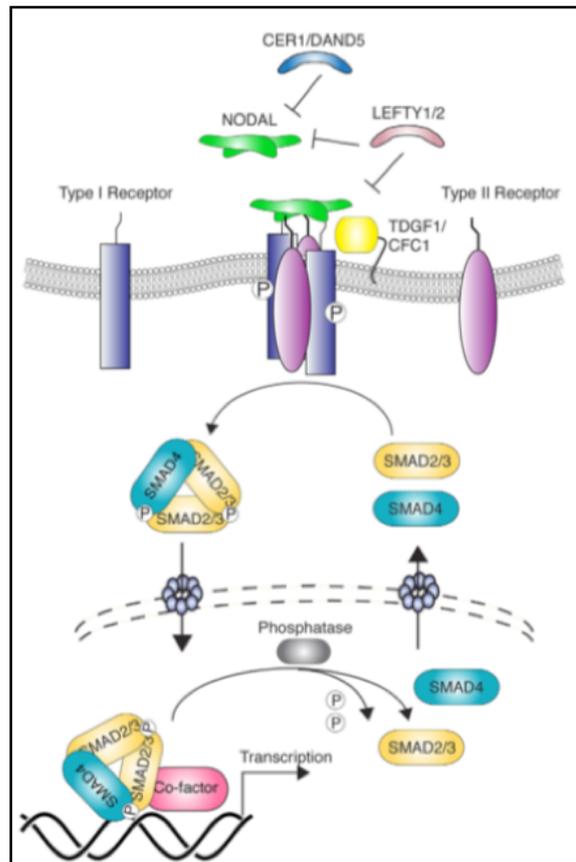


**Figure 10. Voies de signalisation Wnt**

(A) canonique, (B) non-canonique et (C) PCP (*Planar Cell Polarity*). (Marinou et al., 2012)

## Nodal

Nodal est un membre de la famille des TGF- $\beta$  qui joue un rôle primordial dans la formation de l'endoderme et dans la mise en place de l'axe A-P (Tuazon et Mullins, 2015). La signalisation Nodal se fait via des récepteurs transmembranaires à activité sérine/thréonine kinase de type I (Alk4 ou Alk7) ou II (ActRIIA ou ActRIIB). Après activation, Alk4/7 phosphoryle les protéines cytosoliques Smad2 ou Smad3, ce qui provoque la liaison à Smad4 et la translocation dans le noyau cellulaire (**Figure 11**). Les Smads vont alors s'associer à des facteurs de transcription tels que FoxH1, Fast1 et les facteurs de la famille des Mix, qui stimulent la transcription des gènes mésodermiques. La signalisation Nodal promeut également l'expression d'un ensemble de facteurs de transcription dans le lignage endodermique dont Foxa2, Sox17, éomésodermine et Gata4-6. L'action conjointe de ces facteurs est responsable

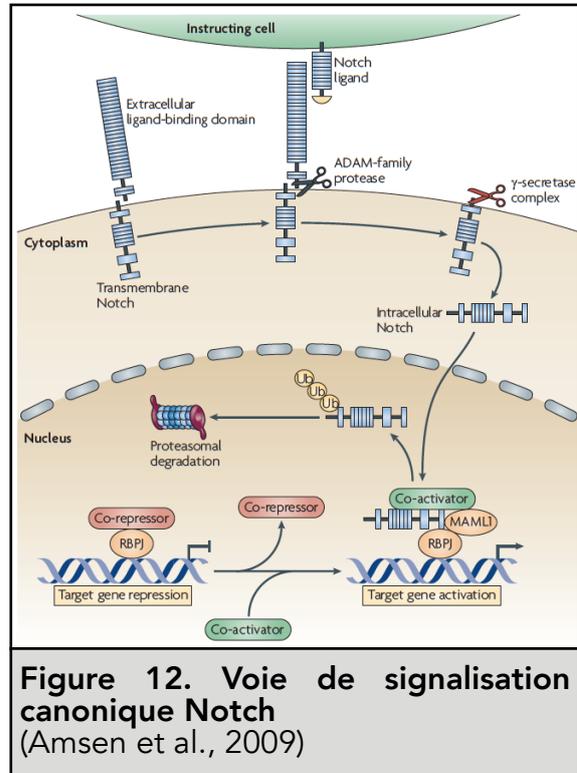


**Figure 11. Voie de signalisation Nodal** (Hill et al., 2017)

de la ségrégation des lignages mésodermique et endodermique, de l'engagement des cellules dans une destinée endodermique et dans les premiers événements de régionalisation de l'endoderme. L'expression de Nodal est notamment contrôlée par la voie canonique Wnt via la  $\beta$ -caténine ainsi que le facteur de transcription TCF.

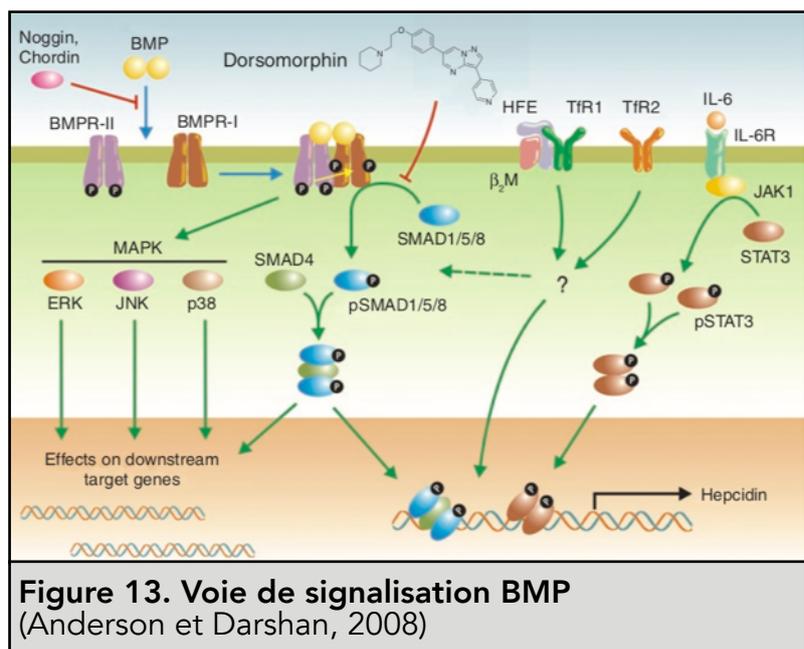
### Notch

Notch est exprimé à la surface cellulaire et consiste en un domaine de liaison au ligand extracellulaire et un polypeptide transmembranaire. La signalisation est initiée par l'interaction de Notch et un ligand *Delta-like* (DLL) ou *Jagged*. S'ensuit un clivage dans la région transmembranaire de Notch qui provoque la libération de son domaine intracellulaire et sa translocation dans le noyau où il active la transcription de ses gènes cibles (**Figure 12**) (Amsen et al., 2009). La signalisation NOTCH contrôle à la fois la prolifération et la spécification des lignages de l'épithélium intestinal.



### BMP

Les BMP sont des membres de la famille des TGF- $\beta$  et sont produits dans le mésenchyme. Au cours de la gastrulation, ils inhibent la formation de l'endoderme en constituant un hétérodimère inactif avec Nodal et favorisent la formation du mésoderme. La signalisation BMP est relativement complexe mais de façon générale, celle-ci aboutit à l'activation de la phosphorylation de Smad1, 5 et 8 (**Figure 13**). Les protéines Smad ainsi activées sont transloquées dans le noyau où elles activent la transcription de leurs gènes cibles (Le Guen et al., 2015 ; Whitman, 1998).



## Partie 2. Les cellules entéroendocrines

### 1. Fonctions physiologiques et diversité des cellules entéroendocrines

Bien que les cellules entéroendocrines soient une population rare de cellules, ne représentant que 1% des cellules au sein de l'épithélium intestinal, elles font de l'intestin le système endocrine le plus important de l'organisme. Ce système endocrine a la particularité de ne pas être organisé en structure glandulaire, mais au contraire, en cellules isolées disséminées de manière hétérogène le long du tractus gastro-intestinal, de l'estomac au rectum. Son rôle est de détecter les composants de la lumière intestinale, monitorer le statut énergétique et provoquer les réponses physiologiques adaptées pour contrôler l'homéostasie métabolique postprandiale en réponse à la nourriture ingérée (**Figure 14**) (Amsen et al., 2009).

Fonction majeure	Exemples d'effets	Hormones impliquées
<b>Ingestion</b>	Appétit, satiété, rejet des éléments toxiques ou nocifs	CCK, GLP1, PYY (satiété); ghréline (appétit); 5-HT (nausée)
<b>Digestion</b>	Libération d'agents digestifs (acides, hormones), motilité, induction de transporteurs	Gastrine, histamine, CCK (sécrétion d'agents digestifs); motiline, 5-HT (motilité);
<b>Métabolisme</b>	Absorption et stockage du glucose, sécrétion d'hormones de croissance, mobilisation des lipides, métabolisme osseux	GIP, GLP1, ghréline, 5-HT

**Figure 14. Principales fonctions des cellules entéroendocrines et hormones impliquées** (Adapté de Amsen et al., 2009)

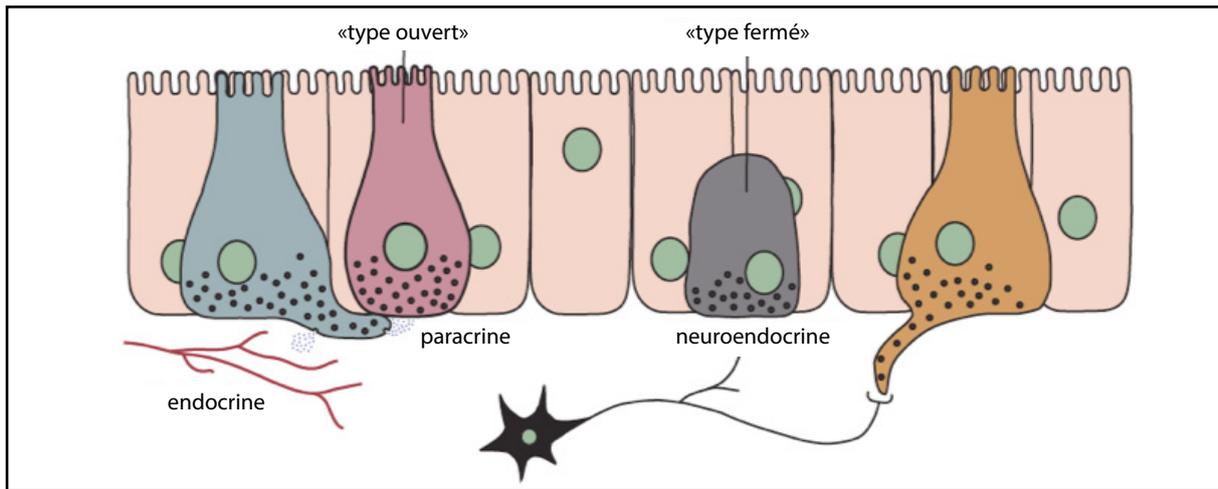
Pour ce faire, ces cellules produisent plus de 20 hormones différentes qui agissent par des mécanismes paracrine, endocrine et neuroendocrine (**Figure 15**) (Adriaenssens et al., 2018) :

-L'action paracrine est liée à la diffusion locale des peptides et amines dans le chorion de la muqueuse. Le produit de sécrétion exerce son action via les cellules avoisinantes (cellules épithéliales, fibres nerveuses, cellules conjonctives et musculaires);

-L'action endocrine résulte du passage des polypeptides dans le sang circulant. Par cette voie, la cellule exerce son action à distance sur des organes cibles;

-L'action neuroendocrine est rendue possible par la présence de fibres nerveuses qui projettent à la muqueuse intestinale. De plus, certaines cellules entéroendocrines forment des synapses avec les terminaisons nerveuses grâce à des structures appelées neuropodes.

Il existe des cellules de type « ouvert » (*open-type*), caractérisées par la présence de microvillosités au pôle apical, participant ainsi à la bordure en brosse, alors que d'autres sont qualifiées de « fermées » (*closed-type*), car elles sont englobées dans la couche épithéliale sans contact avec la lumière du tube digestif (**Figure 15**).

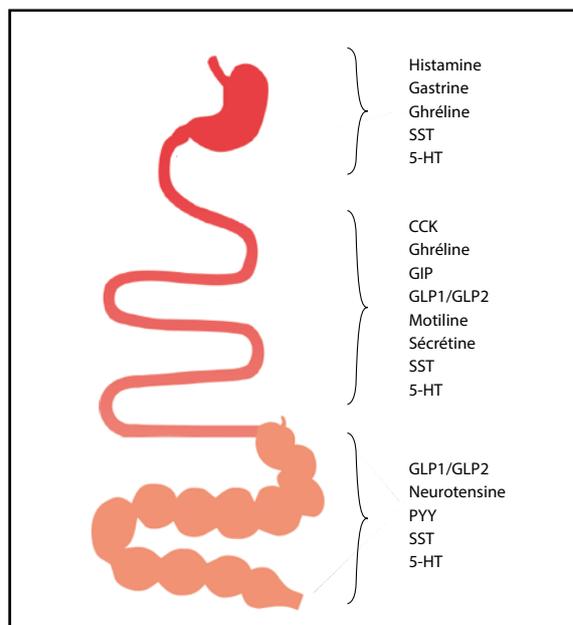


**Figure 15. Morphologies et modes de signalisation adoptés par les cellules entéroendocrines**

Il existe des cellules dites de type «ouvert» (rouge) et d'autres de type «fermé» (gris). Certaines possèdent des extensions à la surface basolatérale facilitant la signalisation paracrine avec les cellules voisines (bleu) et d'autres possèdent des extensions similaires appelées neuropodes qui permettent la communication neuroendocrine (orange). Les capillaires sanguins de la *lamina propria* permettent la signalisation endocrine des hormones sécrétées. (Adapté de Adriaenssens et al., 2018)

Les hormones entéroendocrines jouent un rôle primordial en coordonnant les fonctions digestives des différents organes et échangent des informations avec le système nerveux central. Chacun de ces organes remplit une fonction précise en produisant notamment des sucs, enzymes et hormones nécessaires au bon déroulement de la digestion. Le pancréas, par exemple, sécrète les sucs pancréatiques et abrite quatre populations de cellules endocrines. Les cellules entéroendocrines sont définies par leur morphologie, leur patron d'expression le long de l'axe proximo-distal de l'intestin et leur principale hormone sécrétée (Furness et al., 2013). En effet, elles sont réparties le long de l'axe proximo-distal du tube digestif selon la fonction physiologique que remplissent les hormones qu'elles sécrètent (**Figures 16 et 17**).

Chaque hormone participe à un des trois rôles majeurs que sont l'ingestion de la nourriture, la digestion et les régulations métaboliques qui en découlent (**Figure 14**). En effet, la ghréline est sécrétée avant les repas et stimule la sensation de faim (Cummings et al., 2001).



**Figure 16. Patron d'expression des hormones entéroendocrines le long de l'axe proximo-distal chez la souris** (Adapté de Adriaenssens et al., 2018)

Type cellulaire	Hormone intestinale	Fonctions	Stimuli principaux
Cellules I	Cholecystokinine (CCK)	Stimulation de la contractilité de la vésicule biliaire Sécrétion d'enzymes pancréatiques et intestinales Réduction de la production d'acide gastrique Baisse de la motilité gastrointestinale Baisse de l'appétit	Lipides Protéines
Cellules G	Gastrine (GAST)	Augmentation de la production d'acide gastrique Augmentation de la motilité gastrointestinale	Protéines Distension gastrique
Cellules X/A-like	Ghréline (GHRL)	Augmentation de l'appétit Inhibition de la libération d'insuline Augmentation de la motilité gastrointestinale Augmentation de la production d'acide gastrique	Faim
Cellules K	<i>Glucose-dependent insulinotropic polypeptide</i> (GIP)	Baisse de la motilité gastrointestinale Stimulation de la libération d'insuline	Lipides Monosaccharides
Cellules L	<i>Glucagon-like peptide-1</i> (GLP1)	Baisse de la motilité gastrointestinale Stimulation de la libération d'insuline Baisse de l'appétit	Lipides Monosaccharides Protéines Acides biliaires Métabolites microbiens
Cellules ECL	Histamine	Augmentation de la production d'acide gastrique	Gastrine
ND	<i>Insulin-like peptide 5</i> (InsI5)	Augmentation de l'appétit Homéostasie du glucose	Restriction énergétique
Cellules M	Motiline (MLN)	Augmentation de la motilité gastrointestinale Augmentation de l'appétit	Lipides Diminution du pH luminal
Cellules N	Neurotensine (NTS)	Baisse de la motilité gastrointestinale Réduction de la production d'acide gastrique Sécrétion d'enzyme pancréatiques Baisse de l'appétit	Lipides
Cellules L	Peptide YY (PYY)	Baisse de la motilité gastrointestinale Réduction de la production d'acide gastrique Inhibition de la sécrétion d'enzymes pancréatiques Inhibition de la contractilité de la vésicule biliaire Baisse de l'appétit Homéostasie des fluides	Monosaccharides Lipides Protéines Acides biliaires Métabolites microbiens
Cellules D	Somatostatine (SST)	Baisse de la motilité gastrointestinale Réduction de la production d'acide gastrique	Protéines Diminution du pH luminal
Cellules S	Secrétine (SCT)	Baisse de la motilité gastrointestinale Réduction de la production d'acide gastrique Libération d'acide biliaire Sécrétion de bicarbonate Baisse de l'appétit	Diminution du pH luminal
Cellules EC	Sérotonine (5-HT)	Augmentation de la motilité gastrointestinale Nausée et hypersensibilité viscérale	Irritants alimentaires Métabolites microbiens

**Figure 17. Fonctions des hormones entéroendocrines sécrétées en réponse à divers stimuli**

(Adapté de Adriaenssens et al., 2018)

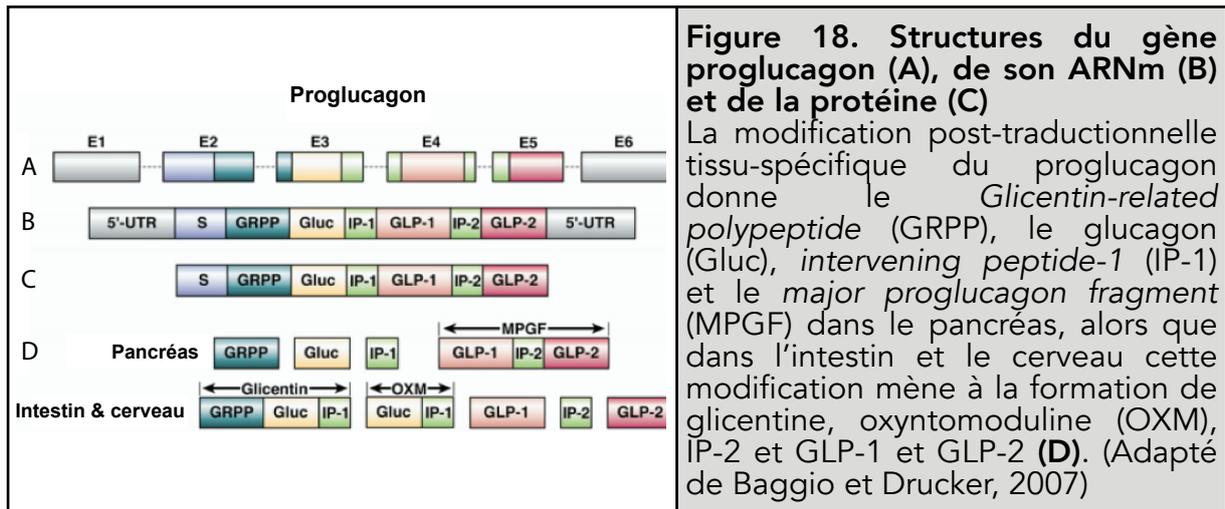
Par la suite, les aliments ingérés stimulent un certain nombre d'hormones de satiété comme par exemple la cholécystokinine (CCK) (Lieverse et al., 1994). Plusieurs hormones sont impliquées dans le contrôle de la dégradation des aliments par l'acide gastrique, la dégradation enzymatique et le péristaltisme (Gribble et al., 2018). Certaines hormones intestinales régulent le métabolisme et le stockage des nutriments. GLP1 et GIP, hormones dites incrétines, favorisent la sécrétion d'insuline glucose-dépendante, protégeant ainsi l'organisme du pic postprandial de glucose sanguin (Verdich et al., 2001), ce qui en fait des cibles d'intérêt thérapeutique dans le cadre du diabète de type 2. Ces deux hormones seront décrites plus en détail dans la partie suivante. D'autres rôles des hormones gastro-intestinales ont été identifiés plus récemment, comme la stimulation de la réparation et de la croissance de la muqueuse intestinale par GLP2 (Burrin et al., 2007) mais aussi l'effet anti-dépresseur et anxyolitique de la ghréline (Mani et Zigman, 2017).

### 1.1. GLP1 / GIP : l'effet incrétine

Dans l'intestin, le GIP est principalement sécrété par les cellules K de l'intestin grêle proximal et sa sécrétion est stimulée par l'ingestion de nutriments. Son action principale est la stimulation de la sécrétion d'insuline en réponse au glucose (*Glucose-Stimulated Insulin Secretion, GSIS*) (Drucker, 2005). De plus, il a été montré que le GIP favorise la biosynthèse d'insuline et active des voies anti-apoptotiques, ce qui augmente la survie des cellules  $\beta$  sécrétrices d'insuline du pancréas. Les souris déficientes pour le récepteur au Gip (*GipR*) présentent des défauts dans la sécrétion d'insuline après administration orale de glucose mais ont moins de graisses corporelles, car elles utilisent les lipides comme source préférentielle d'énergie et sont plus sensibles à l'insuline (Miyawaki et al., 1999). L'action de GIP améliore donc la sécrétion d'insuline mais promeut le stockage énergétique et réduit l'action de l'insuline. Par conséquent, la stratégie à adopter pour générer des thérapies contre le diabète via cette signalisation hormonale n'est pas claire.

Le GLP1 et le glucagon sont synthétisés à partir d'un précurseur commun, le préproglucagon, qui est ensuite transformé en proglucagon et va subir une maturation par l'action de protéases qui donnera le glucacagon, le GLP1 et le GLP2 (**Figure 18**) (Baggio et Drucker, 2007). Le GLP1 est sécrété par les cellules L de l'intestin en réponse à l'ingestion de nutriments, à des signaux neuronaux ou à la stimulation par le GIP. Les souris déficientes pour le récepteur au Glp1 (*Glp1R*) sont intolérantes au glucose et présentent un défaut de GSIS (Scrocchi et al., 1996).

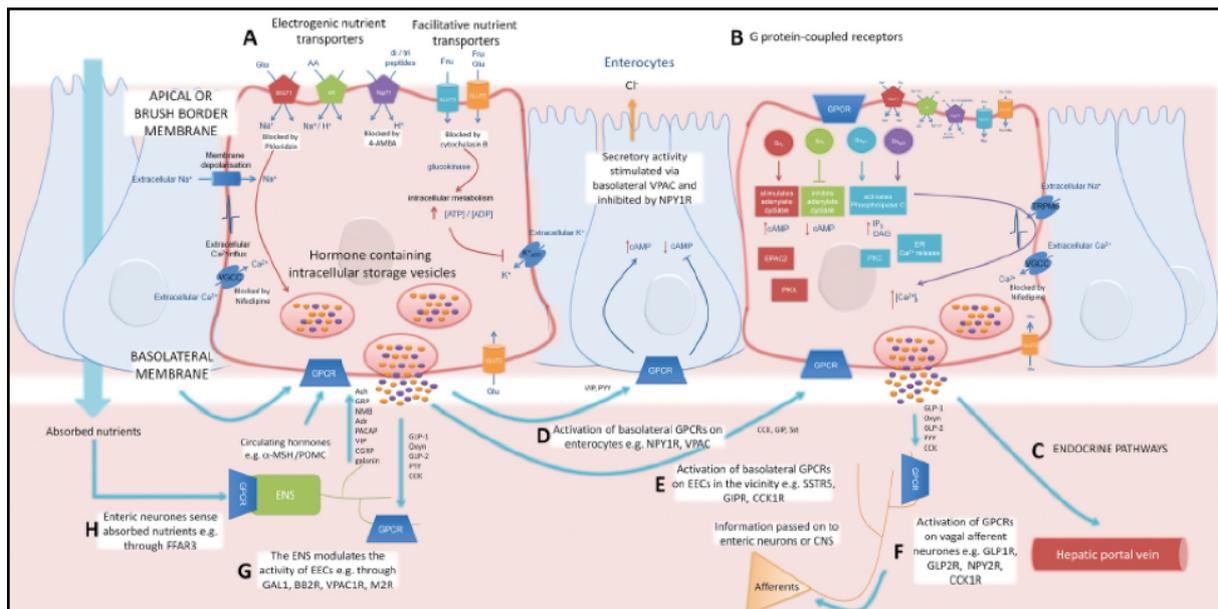
Le GLP1 augmente donc la synthèse et la sécrétion d'insuline et la survie et prolifération des cellules  $\beta$  tout en réduisant la sécrétion de glucagon. Ce sont ces propriétés du GLP1 qui ont été utilisées pour établir des thérapies contre le diabète



(Lee, 2016). En effet, des analogues du GLP1 ou des agonistes des récepteurs du GLP1 ont été développés et constituent une option thérapeutique prometteuse, car ils améliorent le contrôle glycémique, diminuent le poids et offrent l'espoir d'une stabilisation ou d'une amélioration de la fonction des cellules  $\beta$  en favorisant leur prolifération et en inhibant leur apoptose.

## 2. Mécanismes de sécrétion des hormones

Les nutriments sont détectés par les cellules entéroendocrines grâce à un large nombre de protéines dont des transporteurs et récepteurs couplés aux protéines G (GPCRs, Mace et Marshall, 2013). D'autres composants du chyme alimentaire comme les cytokines inflammatoires, les acides biliaires et des neurotransmetteurs sont aussi détectés par les cellules entéroendocrines. De plus, les hormones sécrétées par ces cellules sont également régulées par d'autres cellules entéroendocrines, comme par exemple la sécrétion de somatostatine par les cellules D qui inhibe la sécrétion de GLP1 par les cellules L via le récepteur à somatostatine SSTR5 (Moss et al., 2012). La sécrétion des hormones par les cellules entéroendocrines est régulée par une cascade de réactions et d'évènements ayant lieu au sein des cellules (**Figure 19**). En effet, la détection des nutriments au niveau de la bordure en brosse (membrane apicale) provoque la dépolarisation de la membrane plasmique, qui régule l'ouverture des canaux  $\text{Ca}^{2+}$  voltage-dépendant. Les transporteurs recrutés dépendent de la nature des nutriments : le transport du glucose se fait via le transporteur de glucose couplé au sodium (SGLT1), le transport peptidique via le transporteur de peptide couplé au proton (PepT1) et les acides aminés par leur transporteur électrogénique. Les lipides sont quant à eux détectés par des GPCRs de type FFA (*Free Fatty Acid*) (Adriaenssens et al., 2018). L'ouverture des canaux calciques provoque l'entrée de  $\text{Ca}^{2+}$  extracellulaire et ainsi l'augmentation du niveau de  $\text{Ca}^{2+}$  intracellulaire, qui stimule la sécrétion des hormones par les cellules entéroendocrines.



**Figure 19. Sécrétion des hormones par les cellules entéroendocrines**

Représentation schématique des mécanismes de sécrétion des hormones suite à l'absorption des nutriments par transports électrogéniques, transports facilités ou par les récepteurs couplés au protéines G (GPCRs). (Mace et al., 2015)

Le glucose est également détecté par le transporteur GLUT2 et entre dans le métabolisme intracellulaire du glucose via la glucokinase. Ceci induit la fermeture des canaux K<sup>+</sup> sensible à l'ATP qui provoque également la dépolarisation de la membrane, l'entrée du Ca<sup>2+</sup> et la sécrétion hormonale subséquente. Les nutriments peuvent également être détectés par les GPCRs et provoquer la sécrétion d'hormones via l'activation ou l'inhibition de l'activité adénylate-cyclase ou l'activation de la phospholipase C qui module les niveaux de Ca<sup>2+</sup> intracellulaire (Mace et al., 2015).

### 3. Révision du dogme un sous-type de cellules entéroendocrines / un peptide produit

Historiquement, les cellules entéroendocrines sont classées selon un code qui associe une lettre à une hormone unique qui est supposée être la seule hormone produite par le sous-type entéroendocrine en question. Dans certains cas, la lettre affiliée au type cellulaire correspond à la première lettre de l'hormone produite par ces cellules ou fait référence à des caractéristiques morphologiques des cellules. En effet, les cellules L par exemple, possèdent de grandes vésicules de sécrétion (*Large* = L) alors que celles des cellules S sont plus petites (*Small* = S) (Fothergill et Furness, 2018). Les cellules L représentent d'ores et déjà une exception à la classification, puisqu'il a été démontré que ces cellules contenaient à la fois les GLP et PYY (Ali Rachedi et al., 1984). Cette classification est maintenant en révision suite à différentes études menées chez la souris et ayant montré que les sous-types entéroendocrines sont des populations de cellules mixtes exprimant plusieurs hormones (Habib et al.,

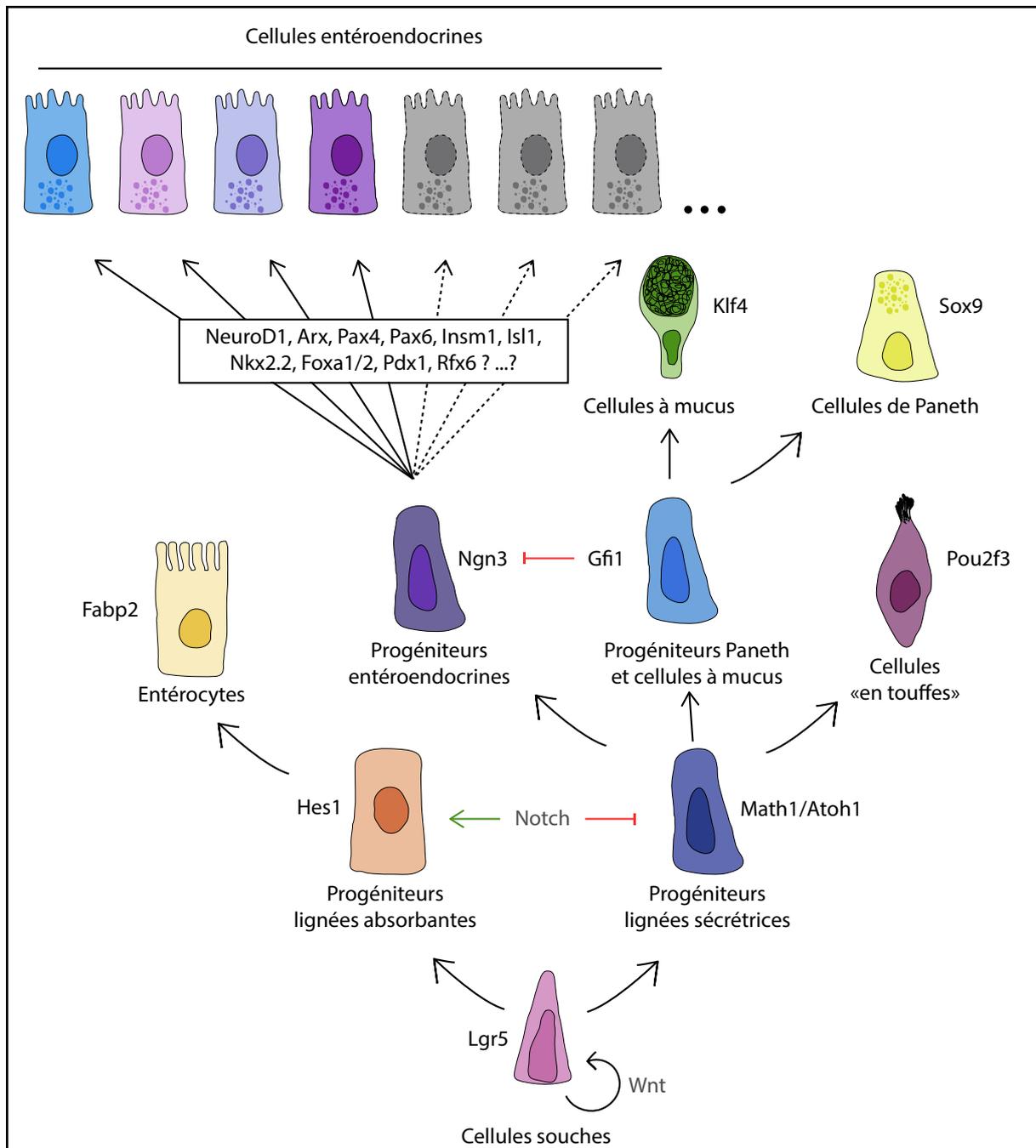
2012 ; Egerod et al., 2012). Au sein d'une même cellule, les hormones peuvent être concentrées ensemble ou dans des vésicules de sécrétion séparées. L'avènement récent des technologies de transcriptomique sur cellules uniques a notamment permis de confirmer l'expression conjointe de plusieurs hormones au sein d'une cellule entéroendocrine. Si l'on considère l'ensemble de la population de cellules entéroendocrines (en différenciation et matures), l'hétérogénéité résulterait à la fois de la variété des différents sous-types entéroendocrines mais également, pour un sous-type donné, de l'hétérogénéité liée à la maturation progressive des cellules progénitrices (Haber et al., 2017; Petersen et al., 2017; Wang et al., 2016; Grün et al., 2015).

Il est évident que la classification de ces cellules sera adaptée et mise à jour, bien que ceci soit compliqué par le fait que la répartition des hormones et leur co-expression diffèrent entre les espèces et les régions du tube digestif (Fothergill et Furness, 2018). La différenciation des sous-types entéroendocrines est régulée par une cascade de facteurs de transcription qui permet aux cellules de s'orienter vers une spécification d'un sous-type plutôt qu'un autre. Bien que l'ensemble de ces mécanismes ne soient pas encore bien compris, plusieurs facteurs de transcription ont d'ores et déjà été identifiés.

#### **4. Régulation transcriptionnelle de la différenciation entéroendocrine**

Les cellules souches intestinales donnent naissance à tous les types de cellules épithéliales, notamment par l'intermédiaire de progéniteurs du lignage sécréteur et du lignage absorbant. En effet, les progéniteurs exprimant *Hes1* se différencient en entérocytes (Ueo et al., 2012 ; Jensen et al., 2000) et les progéniteurs communs aux lignées sécrétrices exprimant *Math1/Atoh1* se différencient en cellules à mucus, cellules de Paneth ou cellules entéroendocrines (**Figure 20**) (Tomic et al., 2018 ; Yang et al., 2001). Cependant, *Math1* ne semble plus être le seul régulateur de la différenciation du lignage sécréteur puisqu'il a récemment été montré que la spécification des cellules « en touffes » et des cellules entéroendocrines peut être induite par l'expression de *Sox4*, indépendamment de *Math1*. En effet, *Sox4* permet la différenciation en cellules entéroendocrines et « en touffes », même en l'absence de *Math1*. (Gracz et al., 2018). Il y a également un lien développemental entre les différents sous-types de cellules entéroendocrines avec un précurseur commun multipotent capable de donner naissance à chacun des sous-types entéroendocrines. En effet, des travaux de l'équipe ont montré que ce précurseur commun est caractérisé par l'expression du facteur de transcription neurogénine 3 (*Ngn3*) (Jenny et al., 2002), comme c'est le cas du précurseur commun aux quatre lignées endocrines dans le pancréas (Gradwohl et

al., 2000). Bien que certains facteurs de transcription exprimés en aval de Ngn3 aient été identifiées, la cascade de régulation transcriptionnelle impliquée dans le lignage des différents sous-types est peu connue, en particulier chez l'homme.



**Figure 20. Principaux facteurs de transcription impliqués dans la différenciation des sous-types entéroendocrines sécréteurs d'hormones**

Tous les sous-types cellulaires de l'épithélium intestinal dérivent des cellules souches. Une cascade complexe de facteurs de transcription, dont certains sont représentés sur le schéma, régie la spécification des différents lignages cellulaires. Tous les facteurs de transcription contrôlant la spécification entéroendocrine ne sont pas définis et la position exacte des facteurs de transcription décrits n'est pas clairement établie. (Schéma réalisé à partir d'illustrations issues de Worthington et al., 2018 ; Ye et Kaestner, 2009)

### **Ngn3**

Ngn3 est un facteur de transcription à domaine basique Hélice-Boucle-Hélice (*basic Helix-Loop-Helix*; bHLH) dont l'expression a été initialement décrite chez l'embryon de souris dans des régions restreintes du système nerveux central (moelle épinière ventrale et hypothalamus) et au niveau du pancréas endocrine (Sommer et al., 1996). Des études ultérieures ont démontré le rôle essentiel de ce facteur de transcription dans le développement des cellules endocrines du tractus digestif (Mellitzer et al., 2010 ; Jenny et al., 2002 ; Gradwohl et al., 2000). En effet, l'inactivation de ce gène chez la souris empêche la formation des cellules des îlots pancréatiques, dont les cellules  $\beta$  sécrétrices d'insuline, et provoque un diabète sévère induisant la mort des souris (Gradwohl et al., 2000). De la même façon, aucune cellule entéroendocrine ne se forme en absence de Ngn3 (Jenny et al., 2002). Ces travaux démontrent qu'au cours de l'embryogenèse, Ngn3 a une fonction pro-endocrine en déterminant la destinée endocrine des progéniteurs multipotents. Par contre, dans l'estomac, l'inactivation de Ngn3 semble altérer la différenciation des cellules G sécrétrices de gastrine et D sécrétrices de somatostatine, alors que les cellules entérochromaffines (EC) sécrétrices de sérotonine, entérochromaffines-like (ECL) sécrétrices d'histamine et les cellules X/A sécrétrices de ghréline ne semble pas affectées. Ceci suggère une différenciation endocrine à la fois dépendante et non-dépendante de Ngn3 dans l'estomac chez la souris (Jenny et al., 2002 ; Lee et al., 2002). Une étude a montré qu'une population de cellules progénitrices se différenciant en cellules non-endocrines, à savoir en cellules à mucus et cellules de Paneth, exprime aussi Ngn3 (Schonhoff et al., 2004). Plus récemment, les travaux de l'équipe ont démontré que Ngn3 contrôle également le programme de différenciation entéroendocrine dans l'intestin adulte. L'inactivation de *Ngn3* spécifiquement dans l'intestin provoque une mortalité de 50% des souris durant la première semaine de vie et les souris survivantes ne présentent aucune cellule entéroendocrine ainsi qu'une absorption lipidique et un métabolisme glucidique altérés (Mellitzer et al., 2010). De plus, il a été montré que Ngn3 est important pour la différenciation entéroendocrine chez l'homme grâce à des études de perte de fonction à l'aide de lentivirus exprimant des shRNAs ciblant *NGN3* dans un modèle d'organoïdes humain *in vitro* (Spence et al., 2011). Néanmoins, ce modèle ne présente pas une inactivation totale de *NGN3* puisqu'une diminution de 63% du niveau de l'ARNm de *NGN3* a été induite. De plus, seule l'absence du marqueur générique de cellules entéroendocrines Chromogranine A (CHGA) a été observée. Ainsi, pour mieux caractériser le rôle de *NGN3* dans la différenciation entéroendocrine, il est indispensable d'établir un modèle de perte de fonction totale permettant d'étudier plus en détails les sous-types entéroendocrines et les cascades de régulations transcriptionnelles affectées.

### **Arx et Pax4**

Arx et Pax4 sont des facteurs de transcription à homéodomaines qui sont également nécessaires au contrôle de la spécification des sous-types entéroendocrines (Beucher et al., 2012 ; Du et al., 2012). En effet, l'équipe a montré que chez la souris déficiente pour Arx, la différenciation des sous-types est altérée, menant à l'absence des lignages entéroendocrines sécréteurs de Glp1, Gip, Cck, gastrine, ghréline et sécrétine. Une redirection vers une différenciation en cellules D, concomitante avec une augmentation de somatostatine a été observée. A l'inverse, chez la souris déficiente pour Pax4, la différenciation des cellules à sérotonine et à somatostatine est empêchée, aussi bien que celle des cellules sécrétrices de gastrine et Gip. De plus, une augmentation de Arx associée à l'augmentation du nombre de cellules L productrices de Glp1 est observée. Arx et Pax4 sont respectivement nécessaires à la différenciation des cellules L et D et ont une action antagoniste dans la spécification de ces sous-types. La différenciation de certains lignages entéroendocrines dépend donc de Arx et Pax4 alors que d'autres dépendent de l'un ou de l'autre (Beucher et al., 2012). Des fonctions similaires de ARX ont été mises en évidence chez l'homme grâce à des études de perte de fonction à l'aide de lentivirus exprimant des shRNAs ciblant ARX dans un modèle d'organoïdes humain *in vitro* (Du et al., 2012). Effectivement, l'absence de ARX provoque une diminution de l'expression des gènes codant la sécrétine, CCK et le préproglucagon dont le GLP1 dérive. Il semblerait que le rôle de Arx soit conservé entre la souris et l'homme, mais des analyses complémentaires doivent être entreprises puisque le modèle n'implique pas une inactivation totale de ARX mais une diminution de 50% du niveau de son ARNm. Ainsi, il est nécessaire d'établir un modèle de perte de fonction totale de ARX dont l'analyse phénotypique détaillée permettra une meilleure compréhension de l'implication de ce facteur de transcription dans la régulation de la différenciation entéroendocrine chez l'homme.

### **Rfx6**

Rfx6 est un facteur de transcription à domaine *winged helix* qui est exprimé durant le développement dans l'endoderme de l'intestin primitif postérieur et devient peu à peu restreint dans l'estomac, l'endoderme intestinal et le pancréas. Dans le pancréas en développement, il est exprimé dans le lignage endocrine où il co-localise avec une population de cellules Ngn3<sup>+</sup>, puis son expression est maintenue dans les cellules endocrines adulte (Piccand et al., 2014 ; Smith et al., 2010 ; Soyer et al., 2010). La fonction de Rfx6 est de mieux en mieux caractérisée dans le pancréas mais à l'heure actuelle très peu de données de la littérature décrivent sa fonction dans l'intestin. Des travaux en voie de publication réalisés par l'équipe semblent toutefois indiquer un rôle dans la différenciation entéroendocrine, puisque les souris déficientes pour Rfx6

présentent une baisse drastique de l'expression des hormones, excepté la sérotonine dont l'expression semble augmentée. De plus, sa délétion spécifique dans l'intestin provoque des diarrhées et une malabsorption des lipides associées à la perte des cellules entéroendocrines (Piccand et al., manuscrit en préparation).

### **NeuroD1**

NeuroD1 ou Beta2 est également un facteur de transcription à bHLH, initialement découvert comme un facteur important dans l'activation de la transcription du gène codant l'insuline (Naya et al., 1995) et dans la différenciation neuronale (Lee et al., 1995). Il a été montré que Ngn3 se lie au promoteur du gène codant NeuroD1 pour en augmenter sa transcription (Huang et al., 2000). NeuroD1 semble réguler la transcription de certains gènes codant des hormones puisque les souris déficientes pour *NeuroD1* ne développent pas de cellules productrices de sécrétine et de Cck (Naya et al., 1997). En effet, NeuroD1 régule la différenciation des cellules entéroendocrines à sécrétine en coordonnant la transcription du gène codant la sécrétine et l'arrêt du cycle cellulaire (Ratineau et al., 2002 ; Mutoh et al., 1998).

### **FoxA1 et FoxA2**

Ce sont des facteurs de transcription de la famille des *winged helix* qui contrôlent la différenciation des cellules à mucus et des cellules entéroendocrines L et D chez la souris. Ils interviennent donc dans la différenciation entéroendocrine et plus généralement dans la spécification du lignage sécrétoire (Ye et Kaestner, 2009).

### **Autres facteurs de transcription**

Un nombre important d'autres facteurs de transcription semblent impliqués dans la spécification des sous-types entéroendocrines mais ne sont pas forcément bien décrits. De manière générale, les facteurs impliqués dans la différenciation endocrine du pancréas participent également à la régulation de la différenciation entéroendocrine. Il y a par exemple **Pax6**, dont l'inactivation chez la souris provoque une importante baisse du nombre de cellules à Gip. Dans l'estomac, son absence mène à une diminution des cellules sécrétrices de somatostatine et de gastrine (Larsson et al., 1998). L'inactivation de **Pdx1** chez la souris provoque la mort périnatale des souris, qui ont des malformations digestives dont l'absence de villosités. Elles présentent également une diminution des cellules à sérotonine, sécrétine, Cck, Gip, somatostatine et neurotensine dans le duodénum (Offield et al., 1996 ; Larsson et al., 1996). Dans l'estomac, il semblerait que Pdx1 soit important pour la maturation des cellules G sécrétrices de gastrine puisque ces cellules sont quasiment perdues chez ces souris, alors qu'une augmentation des cellules à sérotonine est observée (Larsson et al., 1996). **Nkx2.2** est également impliqué dans cette cascade de régulation

puisque son inactivation chez la souris provoque la diminution du nombre de cellules entéroendocrines et en particulier les cellules à Cck, sérotonine, Gip, gastrine, Glp1, somatostatine et neurotensine (Desai et al., 2008). **Insm1** est également exprimé dans l'intestin et son absence cause une perte des cellules sécrétrices de neurotensine et de substance P et une diminution des cellules à Cck, Pyy et sérotonine (Gierl et al., 2006). De plus, **Isl1** est nécessaire à la différenciation des cellules L, K (Gip), I (Cck) et D mais inhibe la spécification des cellules EC (5-HT) (Terry et al., 2014).

En résumé, ces données suggèrent que la différenciation entéroendocrine et l'expression hormonale dépendent d'un réseau complexe de facteurs de transcription. Cependant, les mécanismes qui déclenchent la spécification des progéniteurs  $\text{Ngn3}^+$  et les raisons pour lesquelles certaines cellules se différencient en un sous-type entéroendocrine plutôt qu'un autre, restent à élucider.

La majorité des données concernant le rôle des facteurs de transcription dans la différenciation entéroendocrine a été obtenue dans des modèles murins d'inactivation de ces gènes. Un tableau non exhaustif récapitulant les principales caractéristiques phénotypiques de modèles d'étude de nos gènes d'intérêt (*NGN3*, *RFX6*, *ARX* et *PAX4*) décrits dans la littérature est disponible en **annexe 1**. Ces travaux ont été majoritairement réalisés chez la souris au cours du développement embryonnaire et leur fonction dans la différenciation entéroendocrine n'est pas forcément transposable à l'adulte. Bien que la souris et l'homme aient une forte homologie de séquence génomique, la différenciation entéroendocrine n'est pas forcément régie par les mêmes mécanismes, d'où la nécessité de développer des modèles d'études humains qui seront décrits dans une partie suivante. La compréhension de ces mécanismes est d'autant plus importante que des défauts dans ces processus, et notamment des mutations dans les facteurs de transcription impliqués, mènent à des pathologies chez l'homme.

## 5. Pathologies humaines associées à des défauts dans la régulation transcriptionnelle de la différenciation entéroendocrine

### 5.1. Mutation de *NGN3* : diarrhée congénitale malabsorptive

En 2006, Wang et collaborateurs ont rapporté pour la première fois deux cas de patients atteints de diarrhée congénitale associée à un diabète néonatal, provoquées par des mutations ponctuelles dans le gène *NGN3*. Ils ont alors nommé cette pathologie « anendocrinose entérique », bien que le terme de « diarrhée malabsorptive congénitale » soit le plus fréquemment utilisé. Depuis, environ une dizaine de cas de diarrhées néonatales associées à des mutations de *NGN3* ont été rapportées (Hancili et al., 2018 ; Germán-Díaz et al., 2017 ; Sayar et al., 2013 ; Pinney et al., 2011 ; Rubio-Cabezas et al., 2011 ; Cortina et al., 2007). Les mutations décrites affectent la structure du domaine de liaison à l'ADN ou provoquent la synthèse d'une protéine tronquée, ce qui aboutit dans les deux cas à une perte de fonction de *NGN3*. Tous les patients présentent un phénotype clinique comparable avec principalement une diarrhée provoquée par une malabsorption de tous les nutriments (glucides, lipides et protéines) dès la naissance. La diarrhée est absente pendant les périodes de jeûne et l'absorption de l'eau est normale. Hormis ces symptômes, certains patients développent un diabète de type 1 au cours de leurs premières années de vie (entre 1 et 10 ans). Ces patients sont contraints à être nourris par voie parentérale et leur diabète est pris en charge par des traitements à base d'insuline. Les analyses histopathologiques ont révélé que la structure globale du tractus digestif de ces patients n'est pas altérée. En revanche une absence totale de cellules entéroendocrines (révélées par le marquage de CHGA) est observée dans l'intestin grêle et le côlon de tous les patients. Les autres types cellulaires de l'épithélium intestinal ne sont pas affectés par la mutation.

Bien que les cas d'anendocrinose entérique soient extrêmement rares et que les mécanismes de cette pathologie soient encore mal compris, il semble évident que les troubles de l'absorption intestinale provoquant la diarrhée soient causés par la perte des hormones intestinales régulant l'ensemble des fonctions physiologiques du tractus digestif. En effet, les hormones entéroendocrines sont notamment des médiateurs critiques de l'absorption des lipides. Une fois les triglycérides ingérés, CCK et SCT provoquent la sécrétion des enzymes pancréatiques et la sécrétion de la bile par la vésicule biliaire (Chang et al., 2001). De plus, une implication de NTS et GLP2 dans l'absorption des acides gras et la production des chylomicrons a également été mise en évidence (Li et al., 2016 ; Dash et al., 2014). Les modèles murins déficients pour *Ngn3* spécifiquement dans l'intestin et présentant une perte totale de cellules

entéroendocrines développent également des troubles de l'homéostasie énergétique dont une malabsorption des lipides (Mellitzer et Gradwohl, 2011 ; Mellitzer et al., 2010).

## 5.2. Mutation de *RFX6* : syndrome de Mitchell-Riley

Récemment, des analyses génétiques par séquençage réalisées sur des enfants souffrant de diabète néonatal ont révélé la présence de mutations dans le gène *RFX6*. Le phénotype clinique de ces patients atteints du syndrome de Mitchell-Riley est caractérisé par un diabète néonatal, une atrésie intestinale, une hypoplasie de la vésicule biliaire ainsi que des diarrhées (Khan et al., 2016 ; Skopkova et al., 2016 ; Sansbury et al., 2015 ; Concepcion et al., 2013 ; Spiegel et al., 2011; Smith et al., 2010 ; Mitchell et al., 2004). Ces symptômes sont associés à une altération du pancréas endocrine, avec l'absence de cellules exprimant l'insuline, le glucagon et la somatostatine (Mitchell et al., 2004), bien que des cellules exprimant la CHGA soient présentes et que le pancréas exocrine soit normal. La plupart des patients sont nourris par voie parentérale et leur diabète est traité à l'insuline. Le pronostic vital des patients est très différent selon le type de mutation présent : dans les formes les plus sévères de la maladie, le plus souvent associées à des mutations homozygotes provoquant une perte de fonction totale de *RFX6*, les patients décèdent durant les premiers mois postnataux. Certains patients avec une forme moins sévère de la pathologie survivent plus longtemps, en effet plusieurs cas de patients âgés d'une dizaine d'année ont été rapportés.

## 5.3. Mutation de *ARX*

Un nombre important de pathologies neurologiques graves sont associées à des mutations du facteur de transcription *ARX*, traduisant son rôle crucial dans le développement du système nerveux. En effet, des mutations de *ARX* sont retrouvées chez des patients atteints d'encéphalopathie épileptique, d'hydrocéphalie, de lissencéphalie, de retard mental, ou encore des syndromes de Partington et de Proud (Guerrini et al., 2007 ; Bienvenu et al., 2002 ; Bonneau et al., 2002 ; Kitamura et al., 2002 ; Stromme et al., 2002). De façon intéressante, environ la moitié des patients présentant des mutations faux-sens et non-sens, provoquant une perte de fonction de *ARX*, souffrent de diarrhée congénitale. En 2015, Terry et al. ont rapporté le cas d'un patient présentant une expansion polyalanine dans *ARX* (mutation la plus fréquemment retrouvée chez les patients) et souffrant, en plus de graves troubles du système nerveux, d'une pseudo-obstruction intestinale chronique. Celle-ci se caractérise, chez ce patient, par une intolérance à l'alimentation, révélée notamment par des vomissements et des douleurs abdominales. Après plusieurs chirurgies abdominales,

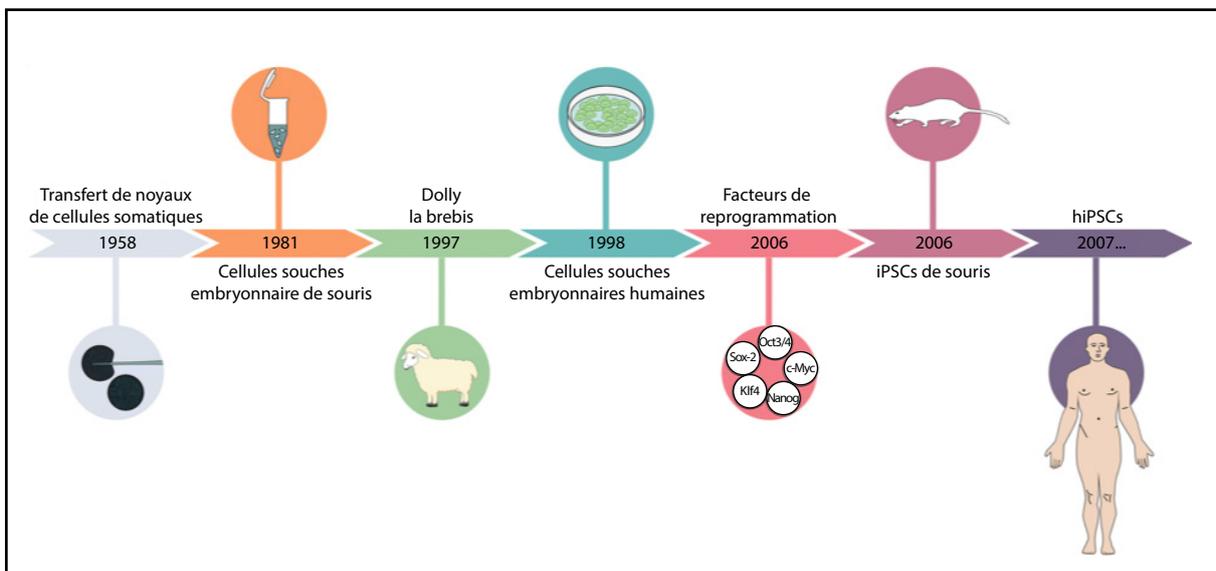
le patient a dû être nourri par voie parentérale. D'un point de vue histopathologique, ce patient ne présentait aucune altération apparente de la structure du tractus digestif et pas de diminution du nombre de cellules entéroendocrines (marquées par CHGA). En revanche, une diminution drastique du nombre de cellules productrices de CCK, SST et GLP1 est observée. De façon intéressante, les modèles murins avec une perte de fonction de Arx présentent quant à eux une augmentation du nombre de cellules à Sst (Beucher et al., 2012 ; Du et al., 2012).

Ainsi, une meilleure compréhension des mécanismes qui contrôlent la différenciation et la fonction des cellules entéroendocrines représente un enjeu majeur. Des traitements tirant profit de l'effet incrétine du GLP1 sont d'ores et déjà utilisés contre le diabète et l'obésité et des stratégies pouvant augmenter la sécrétion d'hormone endogène représentent une piste prometteuse dans le développement de nouvelles thérapies. De plus, ces cellules semblent posséder une certaine plasticité, illustrée par les changements d'environnement nutritionnel et métabolique qui peuvent modifier le nombre, l'activité et la sensibilité de ces cellules (Richards et al., 2016 ; Aranas et al., 2015). Il serait alors intéressant de développer des molécules permettant de modifier la distribution des sous-types entéroendocrines.

De nombreuses études du développement et de la fonction intestinale ont été réalisées sur des modèles animaux *in vivo* et malgré la conservation entre les vertébrés, tous les mécanismes cellulaires et moléculaires ne sont pas forcément transposables à l'homme (Nguyen et al., 2015; Gabdoulline et al., 2015; Schnerch et al., 2010). Un effort a donc été accompli ces dernières années pour développer des modèles humains *in vitro*. La diversité de ces modèles utilisés est exposée dans la partie ci-après.

## Partie 3. Modèles d'étude du développement et de la différenciation intestinale

Historiquement les modèles animaux ont été les plus utilisés pour étudier l'organogenèse, et notamment celle du tractus gastro-intestinal, mais de nouvelles approches plus récentes offrent la possibilité d'explorer les mécanismes moléculaires régulant le développement embryonnaire humain. C'est le cas notamment d'un certain nombre d'approches *in vitro* développées depuis l'avènement de la culture de cellules souches pluripotentes et qui visent à mimer le développement embryonnaire (**Figure 21**). Ces modèles sont en constante progression et deviennent de plus en plus complexes, dans un souci de reproduire un maximum des caractéristiques retrouvées *in vivo*. Ces nouvelles approches permettent, en récapitulant les grandes étapes du développement, de caractériser des événements morphologiques et moléculaires encore mal décrits. Ainsi, la mise en commun des données obtenues par le biais des différents types de modèles devrait améliorer nos connaissances sur les aspects moléculaires et physiologiques du développement.



**Figure 21. Chronologie des découvertes majeures dans le domaine des cellules souches**

La reprogrammation génétique a vu le jour dès 1958 avec le premier clonage par transfert nucléaire de cellules somatiques. Les cellules souches embryonnaires dérivées de la souris (mESCs) ont été générées en 1981. Le premier clonage d'un organisme animal entier en 1997, la célèbre brebis Dolly, a été une avancée majeure notamment aux yeux du grand public. Peu de temps après, en 1998, les premières ESCs dérivées de l'homme (hESCs) ont été générées. Ces cellules étaient alors les seules cellules pluripotentes humaines disponibles jusqu'à la découverte des facteurs de reprogrammation en 2006 par Shinya Yamanaka, lui ayant valu le Prix Nobel de médecine en 2012 et qui a permis la génération des iPSCs de souris et humaine (hiPSCs). (Adapté de Abou-Saleh et al., 2018)

## 1. Modèles *in vivo*

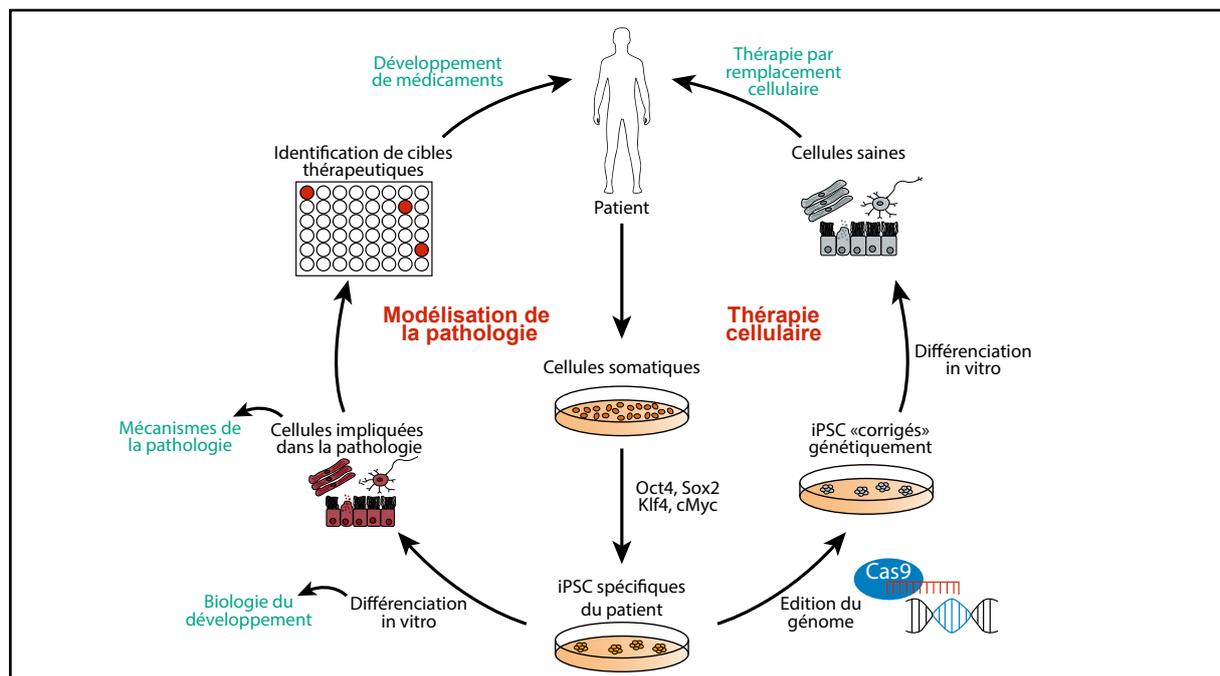
L'approche la plus classiquement utilisée pour explorer les mécanismes du développement consiste à étudier des espèces animales modèles en prélevant des embryons à différents stades de développement ou en post-natal afin de les analyser, notamment au niveau de l'expression des gènes et des protéines. Au cours des dernières décennies, ces approches ont permis d'élucider de nombreux mécanismes du développement des vertébrés. Les modèles d'étude les plus utilisés sont le poisson-zèbre, la drosophile, le poulet, le xénope ou encore la souris. En effet, ces organismes modèles se sont tous avérés particulièrement précieux dans l'étude du développement de par leurs caractéristiques propres. De façon générale leur reproduction est aisée en condition de laboratoire et leur développement est rapide. Le xénope a par exemple été utilisé pour la relative simplicité technique à réaliser des fécondations *in vitro*. Le poisson zèbre possède quant à lui la caractéristique d'avoir des embryons totalement transparents, permettant une visualisation des tissus d'intérêt suite à des marquages cellulaires facilitée. La plupart des études sur le développement intestinal des vertébrés ont été menées sur le poulet et la souris. Les embryons de poulet se développent rapidement et sont facilement manipulables dans le cadre des expériences. Cependant, les outils de modification génétique sont bien plus limités que chez la souris et il y a tout de même des différences significatives dans le développement intestinal entre le poulet et les mammifères, ce qui limite les comparaisons inter-espèces (Walton et al., 2016).

C'est dans ce contexte que les études sur la souris paraissent plus pertinentes, d'autant plus qu'il existe de nombreux outils de modification génétique tissu-spécifique. Ainsi, plusieurs souris génétiquement modifiées ont été générées et notamment des souris déficientes pour certains gènes impliqués dans le développement intestinal, afin d'en déterminer le rôle précis. C'est le cas par exemple des souris déficientes pour *Rfx6*, dont la mutation a été induite dans tous les tissus (*Rfx6*<sup>-/-</sup>) mais également de manière tissu-spécifique dans l'intestin (*Rfx6*<sup>ΔInt</sup>) ou dans les cellules endocrines (*Rfx6*<sup>ΔEndo</sup>) et également dans l'intestin adulte de façon inductible au tamoxifène (*Rfx6*<sup>ΔAdInt</sup>) (Piccand et al., manuscrit en préparation). Cependant, la longue durée de reproduction nécessaire à l'obtention des lignées génétiquement modifiées et la difficulté d'accès au tissu embryonnaire sont autant de limites à l'utilisation des modèles murins. De plus, malgré la forte similarité génétique (Waterston et al., 2002) et la relative conservation des mécanismes développementaux, toutes les découvertes faites chez la souris ne sont pas applicables à l'homme et ce, de façon imprévisible, surtout compte tenu de nos connaissances limitées du développement intestinal humain. Il est vrai que de plus en plus de données indiquent une corrélation limitée entre l'homme et la souris, due aux différences physiologiques inter-espèces

(Nguyen et al., 2015 ; Gabdoulline et al., 2015 ; Schnerch et al., 2010). L'accès au tissu foetal humain ainsi que les nouvelles techniques de culture de tissu *in vitro* à partir de cellules souches pluripotentes permettent maintenant d'étudier le développement de l'intestin dans un modèle humain.

## 2. Modèles *in vitro*

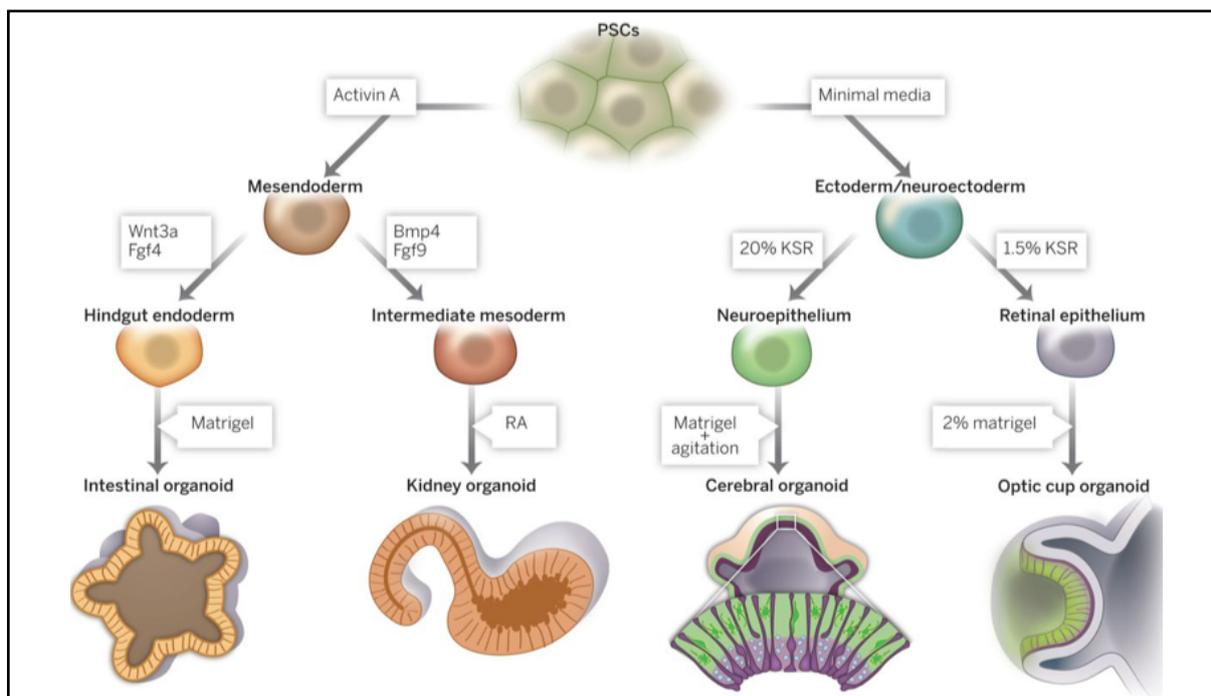
L'établissement de lignées de cellules souches embryonnaires (hESCs) à partir de blastocystes humains (Thomson et al., 1998) ou plus récemment de cellules souches pluripotentes humaines, dites induites (hiPSCs) (**Figure 21**) (Takahashi et al., 2007a ; Yu et al., 2007), a ouvert la possibilité de générer tous les tissus de l'organisme humain, *in vitro*, en récapitulant les programmes de différenciation opérant *in vivo*. Comme les hESCs, les hiPSCs sont des cellules pluripotentes pouvant se différencier en plusieurs types cellulaires. Cependant, les hiPSCs sont issues de la reprogrammation et la dédifférenciation de cellules somatiques humaines adultes grâce à un cocktail de quatre facteurs de transcription : OCT4, SOX2, KLF4, cMYC (Takahashi et al., 2007b) ou OCT4, SOX2, NANOG et LIN28 (Yu et al., 2007).



**Figure 22. Potentiel des hiPSCs : approches thérapeutiques et modélisation**

Les applications de cette technologie sont découpées en deux domaines principaux: la thérapie cellulaire et la modélisation de pathologies. Les cellules somatiques sont reprogrammées en hiPSCs grâce à un cocktail de facteurs de transcription. Ces hiPSCs peuvent d'une part être génétiquement modifiées pour « corriger » la mutation responsable de la pathologie. Ces hiPSCs corrigées sont différenciées en tissu d'intérêt et introduites dans le tissu affecté. D'autre part, elles peuvent être directement différenciées pour étudier le développement du tissu d'intérêt, les mécanismes de la pathologie et ainsi permettre l'identification de cibles thérapeutiques. (Adapté de Stadtfeld et Hochedlinger, 2010)

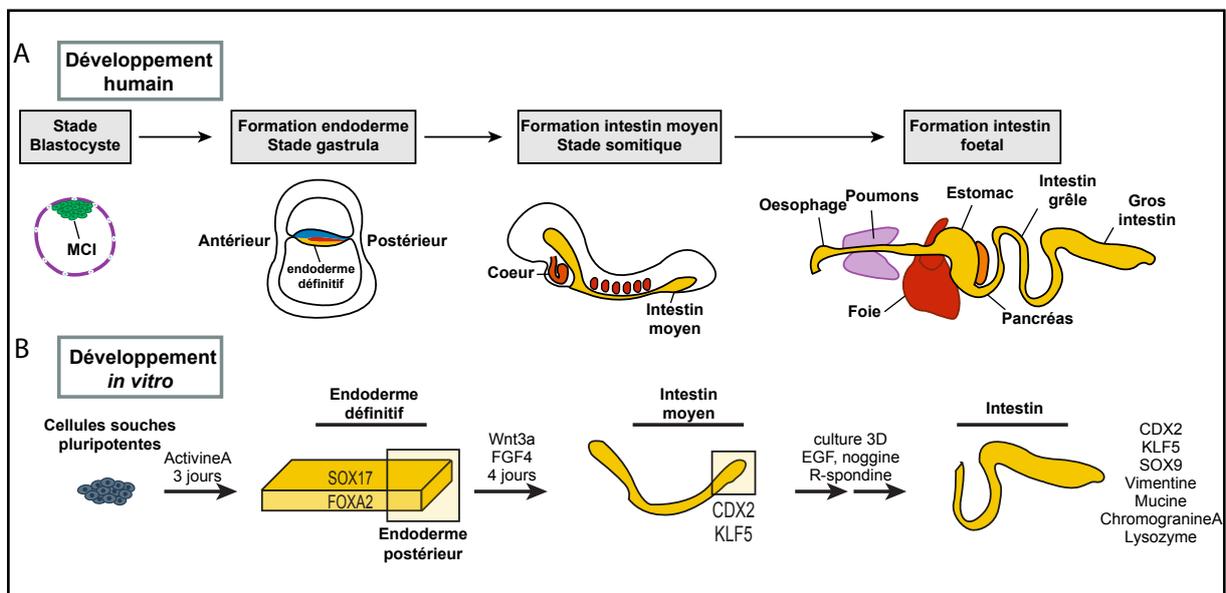
Ces nouveaux outils cellulaires sont utiles pour développer des modèles de pathologies mais également pour étudier la biologie développementale humaine, tout en outrepassant les problèmes éthiques que posaient l'utilisation des hESCs. Ils ouvrent également des perspectives à plus long terme quant à la thérapie cellulaire. En effet, les hiPSCs peuvent être reprogrammées à partir des cellules somatiques de patients puis génétiquement modifiées, par la technologie CRISPR/Cas9 par exemple, pour corriger la mutation responsable de la pathologie et les réintroduire dans les tissus des patients après différenciation *in vitro*. Ces hiPSCs peuvent également être utilisées dans le but de modéliser des pathologies ou de reproduire des mutations afin d'en étudier les mécanismes et/ou d'identifier des cibles thérapeutiques pour le développement de nouveaux médicaments (**Figure 22**). La découverte de la reprogrammation cellulaire ainsi que les connaissances acquises au cours d'années de recherche sur le développement ont permis l'élaboration de protocoles de différenciation visant à reproduire le développement embryonnaire de différents tissus *in vitro* (**Figure 23**). Une avancée importante dans le domaine de l'étude du développement intestinal chez l'homme a été l'établissement par l'équipe de J. Wells (Cincinnati Children's Hospital Medical Center) d'un protocole robuste permettant de générer des tissus intestinaux tridimensionnels à partir de cultures de hESCs et de



**Figure 23. Différenciation *in vitro* dirigée des cellules souches pluripotentes en divers tissus**

La différenciation des cellules souches pluripotentes en tissus spécifiques nécessite des conditions de culture et des facteurs de croissance précis, déterminés grâce aux nombreuses études du développement embryonnaire menées sur les modèles animaux. Ce schéma représente une vue générale des tissus qu'il est possible de reproduire mais n'est pas exhaustif, de même que les facteurs de croissance nécessaires à la différenciation de ces tissus ne sont pas tous indiqués. (Lancaster et Knoblich, 2014)

hiPSCs *in vitro*, grâce à des facteurs de croissance et petites molécules mimant les grandes étapes du développement embryonnaire intestinal et un système de culture favorisant la croissance, la différenciation et la morphogenèse intestinale (Spence et al., 2011; McCracken et al., 2011) (**Figure 24**). Les étapes de ce protocole seront largement détaillées dans la partie 1 des résultats ainsi que dans le matériel et méthodes. Tout d'abord, les hiPSCs sont différenciées en endoderme définitif, qui est en réalité un mélange d'endoderme et de mésoderme (**Figure 23**), mimant l'étape de gastrulation par l'ajout d'activine A dans le milieu de culture pendant 3 jours. L'activine A, membre de la famille des TGF- $\beta$ , se lie au même récepteur que Nodal, induisant une signalisation intracellulaire similaire. Nodal, également membre de la famille des TGF- $\beta$ , est essentiel pour la formation de l'endoderme pendant la gastrulation chez la souris. En effet, une forte expression de ce dernier induit la formation de l'endoderme définitif tandis que des niveaux d'expression plus faibles induisent la formation du mésoderme (Kubo et al., 2004 ; D'amour et al., 2005). Ainsi, l'activine A est utilisée *in vitro* pour reproduire la fonction de Nodal pendant la gastrulation (D'Amour et al., 2005). Ensuite, les évènements de postériorisation sont récapitulés en traitant les cellules avec les facteurs de croissance Wnt3a et FGF4 pendant 4 jours pour former l'intestin postérieur. FGF4 est exprimé par la ligne primitive et induit l'expression de marqueurs postérieurs et la répression de marqueurs antérieurs de l'endoderme de manière dose-dépendante (Dessimoz et al., 2006).



**Figure 24. La différenciation *in vitro* des cellules souches pluripotentes en organoïdes intestinaux humains (HIOs) mime le développement embryonnaire humain**

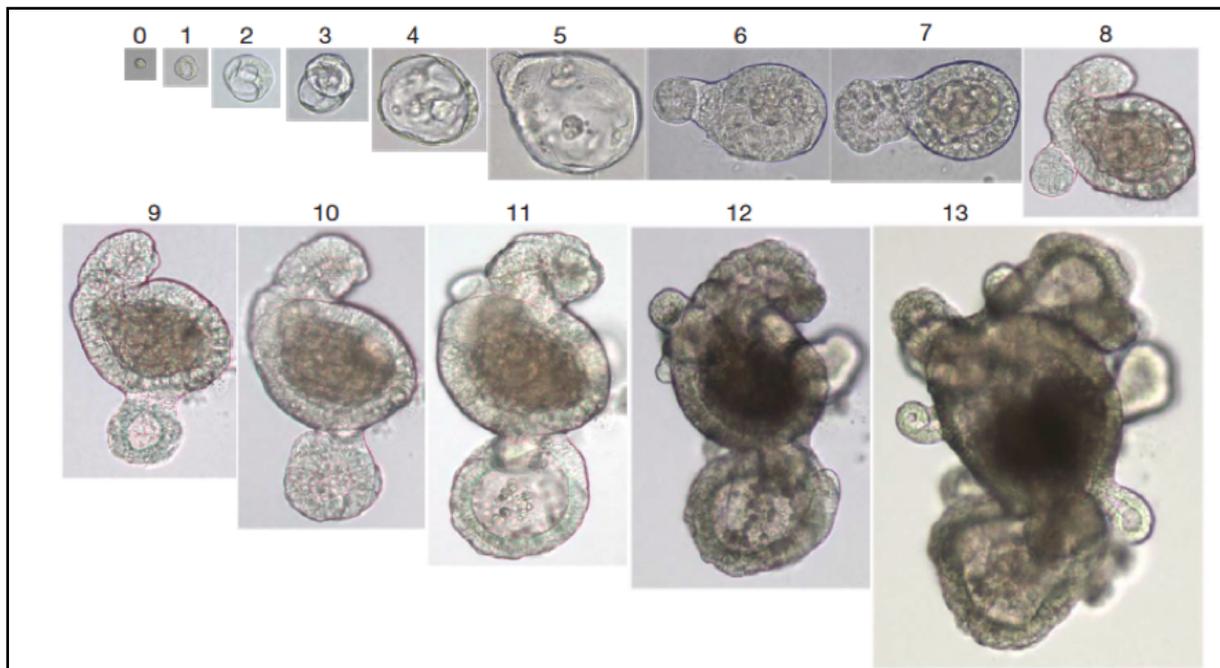
(A) Représentation schématique du développement intestinal humain. MCI: Masse Cellulaire Interne (B) Représentation schématique de la différenciation *in vitro* dirigée des cellules souches pluripotentes humaines en organoïdes intestinaux. (Adapté de Spence et al., 2011)

Wnt3a active la voie canonique WNT/ $\beta$ -caténine qui, en agissant en synergie avec FGF4, promeut la postériorisation de l'endoderme. Enfin, le milieu de culture favorable à la morphogenèse intestinale composé de EGF, noggine et R-spondine et la matrice permettant la croissance tridimensionnelle est ajouté pour former le tissu intestinal en 3D. L'agoniste de WNT, R-spondine1, tout comme l'EGF, jouent un rôle essentiel dans la survie et la prolifération cellulaire (Sato et al., 2011 ; Abud et al., 2005). Noggine est un polypeptide inhibant la transduction du signal TGF- $\beta$  et dont la fonction est de protéger les cryptes de l'action des BMPs, permettant ainsi aux cellules des cryptes de continuer à proliférer (Crosnier et al., 2006). L'intestin ainsi généré, appelé organoïde intestinal (*Human Intestinal Organoids* - HIOs), possède des caractéristiques majeures de l'épithélium intestinal humain, dont l'expression de divers marqueurs spécifiques à ce dernier. D'autres tissus peuvent être générés *in vitro* à partir de l'endoderme définitif formé. En effet, l'inhibition spécifique de la voie de signalisation BMP par l'ajout de l'antagoniste Noggine provoque la répression du marqueur postérieur CDX2, l'activation du marqueur d'intestin antérieur SOX2 et la formation de sphéroïdes de tissu intestinal proximal. La postériorisation des sphéroïdes formées par l'intestin primitif antérieur est accomplie par l'ajout d'acide rétinoïque à la fin du traitement WNT/FGF/Noggine et mène ainsi à la formation de sphéroïdes caractérisées par l'expression de SOX2 et HNF1 $\beta$ . La combinaison des conditions de culture tridimensionnelle et de l'acide rétinoïque permet la formation d'épithélium SOX2/PDX1<sup>+</sup>, spécifique de l'antra gastrique (*Human Gastric Organoids* - HGOs, McCracken et al., 2014). Des organoïdes de côlon peuvent également être formés à partir des sphéroïdes intestinales générées lors de la différenciation *in vitro* en HIOs, par l'ajout d'un milieu contenant du BMP qui permet la formation d'un épithélium exprimant le marqueur distal SATB2 (*Human Colonic Organoids* - HCOs, Mùnera et al., 2017).

### 3. Modèles ex vivo

Les modèles *ex vivo* désignent les expériences menées sur des tissus provenant d'un organisme mais réalisées dans un environnement extérieur avec un minimum d'altérations des conditions naturelles. C'est donc un moyen de conduire des recherches sur des tissus dans des conditions plus contrôlées qu'avec des expériences *in vivo*, tout en risquant néanmoins de modifier l'environnement naturel. En utilisant le principe des modèles *in vitro*, des cellules extraites des organes d'intérêt peuvent être cultivées dans des boîtes de culture afin de déterminer les conditions de culture optimales impliquées dans le renouvellement, la prolifération et la différenciation cellulaire. Les expériences menées sur la souris ont pris une nouvelle dimension avec la mise au point d'un système permettant de cultiver des échantillons de tissus

intestinaux néonataux (Ootani et al., 2009). Les conditions de culture ont permis de maintenir la prolifération et la différenciation au sein de l'épithélium intestinal jusqu'à 350 jours de culture. Cependant, au-delà, les zones prolifératives et la viabilité du tissu commencent à se détériorer. C'est ainsi que les recherches se sont orientées vers un système tirant profit de la manipulation des cellules souches en maintenant des zones prolifératives fournissant de nouvelles cellules continuellement afin d'établir un système de culture viable sur le long terme. L'identification des cellules CBC exprimant Lgr5 par des études de traçage cellulaire *in vivo* ont permis de mettre en évidence les propriétés de cellules souches de ces cellules, c'est-à-dire leur auto-renouvellement et leur caractère multipotent au regard de la différenciation cellulaire (Barker et al., 2007). L'établissement d'un système de culture primaire d'intestin grêle a pu être établi à partir de ces cellules souches intestinales. En effet, en isolant des cellules Lgr5<sup>+</sup> et en les cultivant de manière appropriée, un système de culture dans lequel les cellules se renouvèlent continuellement et la différenciation des cellules épithéliales est contrôlée, a pu être développé (Sato et al., 2009). Ces cellules souches isolées sont capables de former des structures d'organoïdes *in vitro*, appelés entéroïdes ou «mini-intestins», dans un environnement contrôlé et sous l'influence des facteurs de croissance nécessaires à la régulation du renouvellement et de la différenciation cellulaire (Sato et Clevers, 2013) (**Figure 25**).



**Figure 25. Croissance d'une cellule souche Lgr5<sup>+</sup> en entéroïde *ex vivo***

La chronologie de la croissance d'une cellule Lgr5<sup>+</sup> est illustrée du jour 0 (mise en culture) au jour 13 (entéroïde) par une série d'images de microscopie optique à lumière transmise. (Sato et al., 2009)

Ces structures formées, tout comme les modèles *in vitro* décrits, présentent des caractéristiques histologiques similaires à l'épithélium intestinal *in vivo*, incluant les différents types de cellules épithéliales différenciées (notamment des cellules entéroendocrines) ainsi que les cellules souches et prolifératives. Les entéroïdes peuvent ensuite être dissociés et remis en culture pendant plus de 8 mois. Ce système représente donc le protocole optimal pour établir des cultures de cellules intestinales à long terme et est particulièrement utile pour l'étude de la régulation de l'auto-renouvellement des cellules souches ainsi que la différenciation de l'épithélium intestinal.

Jusqu'alors, les études sur la différenciation épithéliale se faisaient principalement sur des lignées cellulaires épithéliales ou *in vivo* sur les modèles animaux. Cependant ces lignées généralement tumorales ne sont pas représentatives des cellules *in vivo* et il y a un vrai intérêt à l'établissement d'un modèle *ex vivo* qui mimerait le développement intestinal *in vivo* tout en ayant la flexibilité et praticité de la culture cellulaire *in vitro*. Ainsi, depuis le développement du système murin *ex vivo*, des efforts ont été fournis quant à l'application aux cellules souches intestinales humaines. En effet, les conditions de culture ont été adaptées et des protocoles de culture d'épithélium de côlon humain ont notamment été mis au point (Sato et al., 2011). Les entéroïdes représentent un modèle préclinique de choix et ces outils pourraient servir dans la biologie régénérative à travers l'expansion de l'épithélium intestinal *ex vivo* (Foulke-Abel et al., 2014).

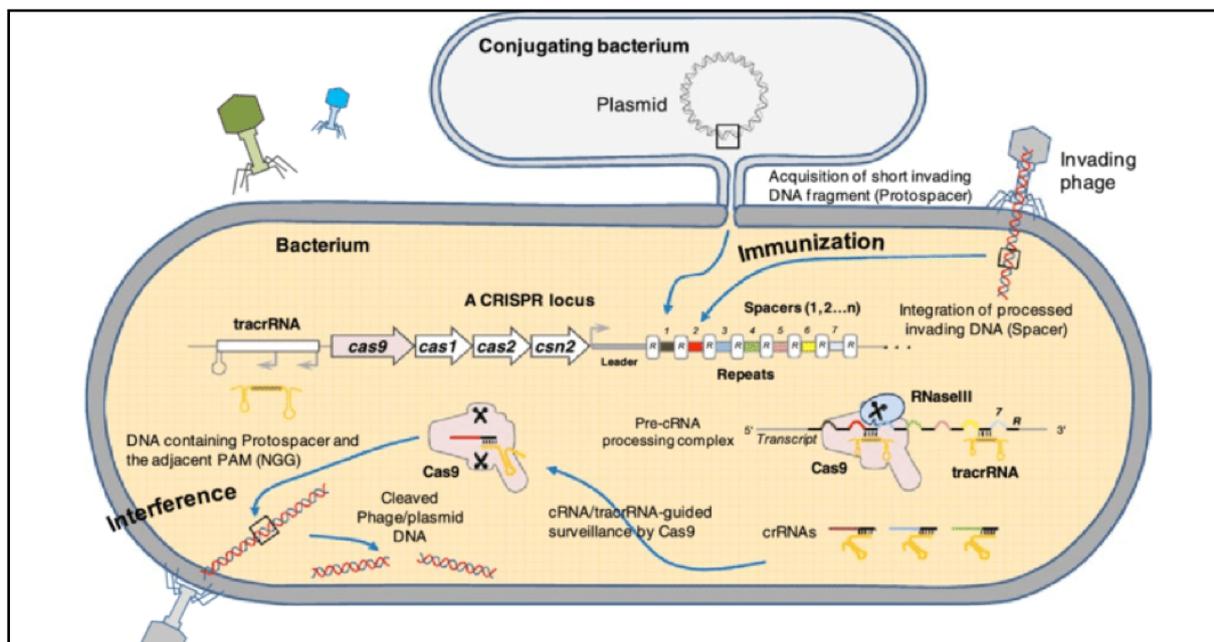
Bien que des cellules entéroendocrines soient présentes dans ces modèles, celles-ci sont peu nombreuses, comme c'est le cas *in vivo*, ce qui constitue une limite à l'étude de ces cellules et de leur différenciation. Pour pallier à ces limitations, des méthodes ont récemment été développées pour augmenter la production de cellules entéroendocrines (Sinagoga et al., 2018 ; Petersen et al., 2018 ; 2015 ; 2014 ; Basak et al., 2017). Ces avancées techniques seront développées en discussion.

Les progrès considérables réalisés quant à l'élaboration de modèles *ex vivo* et *in vitro* sont également rendus possibles par des avancées techniques majeures comme l'étude du transcriptome des cellules épithéliales sous forme de cellules uniques (*single cell*) ou encore le développement d'outils performants permettant l'édition du génome (CRISPR/Cas9). En effet la technologie CRISPR/Cas9 peut être appliquée aux différents types d'organoïdes (entéroïdes et HIOs), afin d'inactiver des gènes d'intérêt ou de corriger des mutations associées à des pathologies (Driehuis et Clevers, 2017). Ces modifications peuvent être induites dans les cellules souches pluripotentes avant différenciation en organoïde ou même directement sur les organoïdes différenciés (Fuji et al., 2015).

## Partie 4. CRISPR/Cas9, outil de modification du génome

### 1. Historique et rôle dans l'immunité adaptative chez les procaryotes

Les techniques de modification du génome ont permis des avancées considérables en biologie au cours des 60 dernières années. Les premières nucléases permettant la modification du génome, les ZFNs (*Zinc-Finger Nucleases*) et les TALENs (*Transcription Activator-Like Effector Nucleases*), ont prouvé leur efficacité dans beaucoup de modèles d'étude mais leur utilisation reste limitée par la nécessité de créer des protéines spécifiques pour chaque site cible (Shen et al., 2014). Une nouvelle technique de modification du génome permettant de contourner cette contrainte a récemment été mise au point : la technologie CRISPR/Cas9. Les CRISPRs (*Clustered Regularly Interspaced Short Palindromic Repeats*) et les protéines associées (Cas) constituent un système naturel décrit dans beaucoup de procaryotes, incluant la majorité des archées et un grand nombre de bactéries (Grissa et al., 2007). Cette technologie repose sur un mécanisme jouant un rôle dans le système immunitaire en tant que défense contre les infections virales (Jinek et al., 2012). Six types de systèmes CRISPR/Cas9 ont été identifiés et sont divisés en deux classes selon la complexité de leur mécanisme (Makarova et al., 2015). Sa principale caractéristique, partagée par tous les types de système, est un locus contenant des répétitions de petites séquences palindromiques (*repeats*) espacés par des séquences appelés *spacers* à intervalles réguliers, qui partagent l'homologie de séquence avec les éléments étrangers acquis lors d'infections par des bactériophages ou plasmides (Marraffini et al., 2010). C'est l'expression de ce locus CRISPR qui a été identifiée comme étant la composante clé de cette immunité adaptative. L'immunité médiée par CRISPR se déroule en trois étapes distinctes : l'acquisition de l'ADN étranger, la biogenèse de l'ARN CRISPR (crRNA) et l'interférence de la cible (**Figure 26**). Effectivement, durant la première phase dite d'adaptation, les acides nucléiques étrangers sont épissés en petits fragments qui forment de nouveaux *spacers* qui sont intégrés au niveau des loci CRISPR notamment grâce aux protéines Cas1 et Cas2. Ces loci constituent ainsi une mémoire génétique des «agressions» passées (Barrangou et al., 2007). L'analyse des séquences *protospacers*, qui sont les séquences homologues aux *spacers* au sein des acides nucléiques étrangers, a révélé la présence d'une petite séquence nucléotidique conservée adjacente au *protospacer*, appelée motif PAM (*Protospacer Adjacent Motif*).



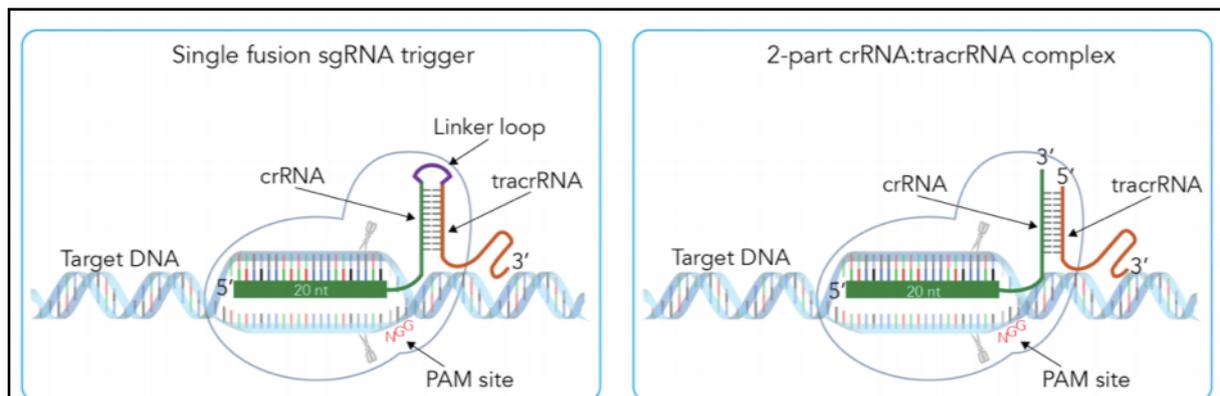
**Figure 26. Mécanisme de l'immunité adaptative médiée par le système CRISPR/Cas de classe II chez les procaryotes**

Le mécanisme d'immunité se déroule en 3 étapes : l'acquisition ou immunisation, la biogenèse des crRNAs et l'interférence de l'ADN étranger. Lors de la première étape, l'ADN étranger est pris en charge et de petits fragments sont intégrés au locus CRISPR en tant que *spacers*. Le complexe Cas9-RNase III-tracrRNA permet la transcription et la maturation des crRNAs. Ces crRNAs forment avec les tracrRNAs un duplex associé à la Cas9 qui va reconnaître les séquences complémentaires aux *spacers* (*protospacers*) au sein de l'ADN étranger. Enfin, la Cas9 induit la coupure double-brin de l'ADN trois nucléotides en amont du motif PAM adjacent au *protospacer*. (Fichtner et al., 2014)

La deuxième étape consiste en la biogenèse des crRNAs par la transcription du locus CRISPR en précurseur de l'ARN CRISPR, appelé pré-crRNA, suivie de la maturation en crRNAs qui possèdent chacun un *spacer* spécifique flanqué de petites séquences d'ARN (Brouns et al., 2008). Dans le système CRISPR/Cas de classe II, par exemple, la maturation des pré-crRNAs en crRNAs nécessite la présence de la nucléase Cas9, la RNase III et le tracrRNA (*trans-activating CRISPR RNA*). Le duplex crRNA-tracrRNA est associé à la Cas9 pour former un complexe qui conduit à la troisième étape, appelée interférence. Pendant cette ultime phase, le crRNA du complexe ribonucléoprotéine Cas9-crRNA-tracrRNA (crRNP) se fixe par appariement de bases complémentaires sur le *protospacer* correspondant et recrute la Cas9 pour la reconnaissance et la destruction de la séquence complémentaire par coupure double-brin de l'ADN. La Cas9, par ses deux domaines endonucléases distincts HNH et RuvC, induit la coupure du *protospacer* trois nucléotides en amont du motif PAM. Le domaine HNH clive le brin d'ADN complémentaire alors que le domaine RuvC clive le brin non-complémentaire (Jinek et al., 2012). Le système CRISPR/Cas9 de type II et sa Cas9 associée issu de la bactérie *Streptococcus pyogenes* est le mieux caractérisé et le premier utilisé en tant qu'outil de biologie moléculaire, car du fait de sa simplicité et de sa taille, il présente le plus d'avantages pour être détourné au profit de l'édition génomique.

## 2. Applications dans la modification de génome

Ce système confère une technologie fiable et efficace pour modifier le génome des cellules à des sites spécifiques de manière précise. Sa première application a été le développement d'un seul ARN guide chimérique (sgRNA) à partir de la fusion du tracrRNA et du crRNA qui maintient la capacité de guider la protéine Cas9 du système CRISPR/Cas de classe II à un site cible. En effet, le sgRNA conserve les éléments critiques au fonctionnement de ce système mais ses deux fragments sont liés par une boucle pour n'en former qu'un unique (**Figure 27**). Ce sgRNA est une séquence de 20 nucléotides (nt) complémentaire à la séquence d'ADN cible et possède une structure tertiaire en double brin à l'extrémité 3' qui lui permet de se lier à la Cas9. Tout comme la forme originale retrouvée chez les bactéries, elle permet la reconnaissance de la séquence cible et la coupure double-brin subséquente par la Cas9 (Tsai et Joung, 2016 ; Ran et al., 2013a).



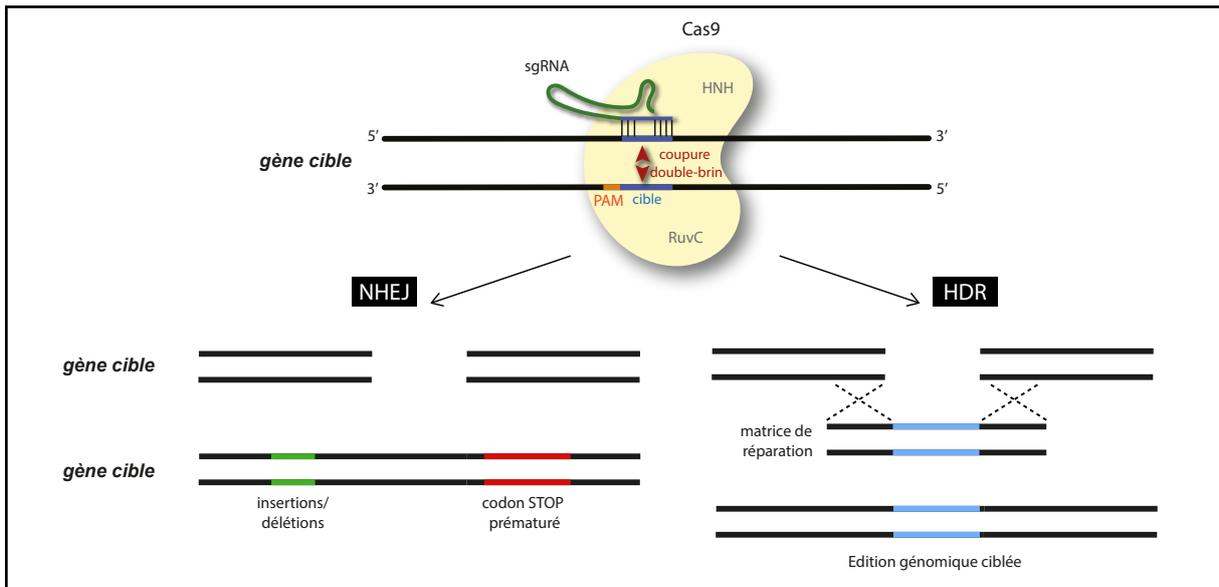
**Figure 27. ARN guide chimérique (sgRNA) synthétisé pour remplir les fonctions du duplex crRNA-tracrRNA**

Le panneau de gauche présente le sgRNA synthétique mimant la fonction du duplex crRNA-tracrRNA par la présence de ces deux éléments liés par une boucle. Le panneau de droite présente le mécanisme naturel tel qu'il est présent chez les procaryotes. (source: idtdna.com)

Cette coupure double-brin générée peut recruter des acteurs des deux voies de réparation de l'ADN possible : le mécanisme sujet à erreur NHEJ (*Non-Homologous End Joining*) ou le mécanisme fidèle HDR (*Homology Directed Repair*) (Tsai et Joung, 2016) (**Figure 28**). Les deux voies peuvent être utilisées à des fins différentes selon la modification génétique souhaitée.

En effet, en l'absence d'une matrice de réparation, c'est-à-dire d'une séquence d'ADN flanquée de sites homologues à la séquence cible, la coupure sera réparée par NHEJ, qui induira des mutations par le biais d'insertions ou délétions (*indel*). Il est possible de tirer profit de cette voie de réparation pour induire l'inactivation de gènes cibles puisque les *indels* générées dans les séquences codantes des gènes peuvent provoquer un décalage du cadre de lecture et/ou l'apparition d'un codon

STOP prématuré. A l'inverse, le mécanisme par HDR peut entraîner des modifications précises aux sites cibles grâce à une matrice d'ADN exogène. Cette matrice peut être sous forme d'ADN double-brin ou d'oligonucléotides simple-brin. Cela représente un système intéressant pour générer des modèles de pathologies dans lesquelles des mutations ponctuelles sont spécifiquement retrouvées dans certains gènes. Il représente également un challenge à plus long terme, pouvant permettre le développement de moyens pour corriger des mutations associées à des pathologies dans les cellules des patients.

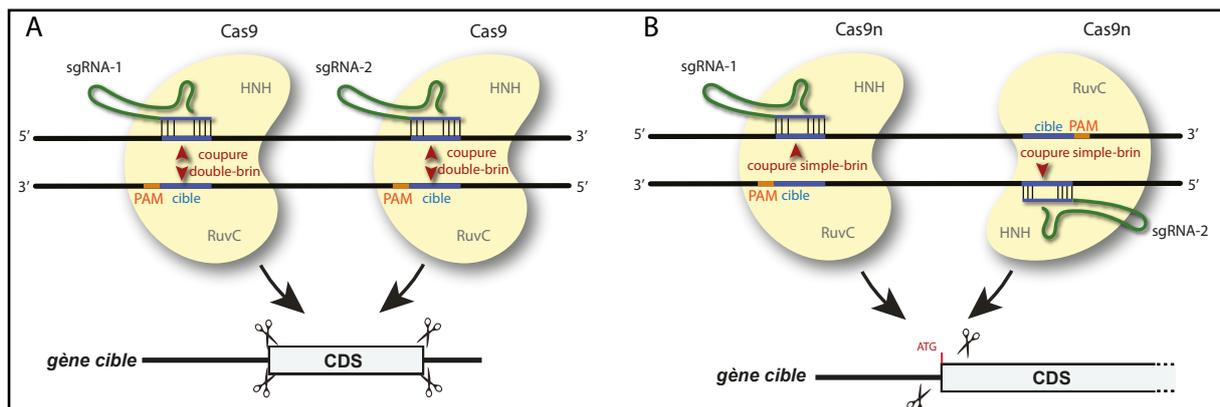


**Figure 28. Mécanismes de réparation de la coupure double-brin générée par la Cas9**

La Cas9 induit une coupure double-brin qui pourra être réparée par deux mécanismes différents : par recombinaison homologue (HDR) si une matrice de réparation composée d'une séquence d'ADN flanquée de sites complémentaires à la cible est disponible, permettant ainsi la modification précise du site, ou dans le cas contraire, par jonctions d'extrémités non-homologues (NHEJ) qui induira des mutations de type insertions ou délétions et possiblement l'apparition d'un codon STOP prématuré.

Une version « humanisée » de la Cas9 du système de *S.pyogenes* permet aujourd'hui de modifier le génome dans des cellules humaines, notamment dans les cellules hiPSCs (Mali et al., 2013). Effectivement, une version optimisée de la Cas9 rendant possible son fonctionnement dans les cellules humaines a été synthétisée et clonée dans un système contenant le signal de localisation nucléaire (NLS) de l'antigène T du virus simien 40 (SV40) permettant son expression dans le noyau des cellules de mammifères. Les sgRNAs sont exprimés dans un vecteur sous contrôle du promoteur U6 de la polymérase III humaine. L'une des principales utilisations de l'outil CRISPR/Cas9 consiste en l'inactivation de gènes. C'est dans ce but que l'approche CRISPR a été utilisée dans le cadre de ce projet de thèse. A cette fin, deux stratégies différentes ont été employées (**Figure 29**), l'une consistant en la délétion de domaines entiers du

gène cible et l'autre en la délétion du codon initiateur ATG. La première stratégie est rendue possible par l'utilisation de deux sgRNA qui permettent l'induction de deux coupures double brin (par la Cas9) de part et d'autre de la séquence cible (**Figure 29A**). La deuxième stratégie utilise une version modifiée de la Cas9. Comme mentionné précédemment, la Cas9 est composée de deux domaines nucléases HNH et RuvC qui permettent le clivage spécifique de chacun des deux brins d'ADN. Des mutations volontairement induites dans ces domaines promettent des fonctions additionnelles dans la modification génétique. Une mutation du dixième acide aminé aspartate en alanine (D10A) du domaine catalytique RuvC et une mutation de l'acide aminé 840 histamine en alanine (H840A) du domaine HNH rendent respectivement l'un des deux domaines nucléases inactif, menant ainsi au clivage d'un seul brin d'ADN au lieu de la coupure double-brin initialement provoquée. Cette nouvelle activité *nickase* de la Cas9, appelée Cas9n, entraîne une coupure simple-brin qui sera réparée par recombinaison homologue grâce au brin complémentaire qui sert alors de matrice de réparation. Ce mécanisme étant plus fidèle, il limite l'introduction involontaire de mutations *indels* dans l'ADN (Tsai et Joung, 2016). Néanmoins, l'utilisation d'une paire de sgRNAs comme guides de la Cas9n, stratégie appelée *double-nickase*, autorise le clivage simultané des deux brins d'ADN mimant ainsi une coupure double-brin, à condition que les deux sgRNAs ciblent des séquences suffisamment proches. De plus, cette stratégie augmente la spécificité de reconnaissance des cibles, de par la nécessité de fixation simultanée des deux sgRNAs sur leurs séquences cibles respectives (**Figure 29B**).



**Figure 29. Inactivation de gène : deux stratégies CRISPR/Cas9 possibles**  
**(A)** Stratégie permettant la délétion de larges domaines géniques par l'utilisation de deux sgRNAs et de la Cas9, induisant deux coupures double-brin. **(B)** Stratégie de double-*nickase* permettant l'excision du codon d'initiation, utilisant deux sgRNAs et la Cas9n mutée D10A, induisant deux coupures simple-brin.

### 3. Les limites de l'outil CRISPR/Cas9

CRISPR/Cas9 semble être l'outil d'édition du génome parfait mais des préoccupations majeures concernant sa spécificité d'action ont récemment vu le jour. En effet, il peut mener à des coupures à des endroits non-ciblés, qualifiés d'effets *off-target*, car le sgRNA peut se fixer sur des séquences pas entièrement complémentaires à sa propre séquence et ce, malgré la nécessité du motif PAM attendant à la séquence cible (Tsai et Joung, 2016). Des outils *in silico*, notamment développés par le laboratoire de Zhang (MIT, USA), permettent de prévenir ces effets en déterminant des sgRNAs à haute spécificité de reconnaissance de la séquence cible, en leur attribuant un score de fidélité. D'autres précautions peuvent être prises pour contrecarrer ce problème et augmenter la spécificité, comme l'utilisation de la double-*nickase* (Ran et al., 2013b) ou encore l'utilisation de dCas9 (sans activité endonucléase), fusionnée à une nucléase FokI (fCas9) (Guilinger et al., 2014).

### Objectif du projet de thèse

Dans ce contexte, l'objectif de mon projet de thèse a été d'étudier les conséquences de l'inactivation totale des facteurs de transcription NGN3, ARX, PAX4 et RFX6 sur la différenciation entéroendocrine chez l'homme. Pour ce faire, ces gènes ont été inactivés par la technique CRISPR/Cas9 dans des cellules souches pluripotentes humaines induites (hiPSCs), qui ont par la suite été différenciées *in vitro* en organoïdes intestinaux humains (HIOs). Les HIOs ainsi générés ont été étudiés par des approches transcriptomiques (RT-qPCR et RNA-seq) et d'imagerie cellulaire (immunofluorescences).

# **MATÉRIEL & MÉTHODES**

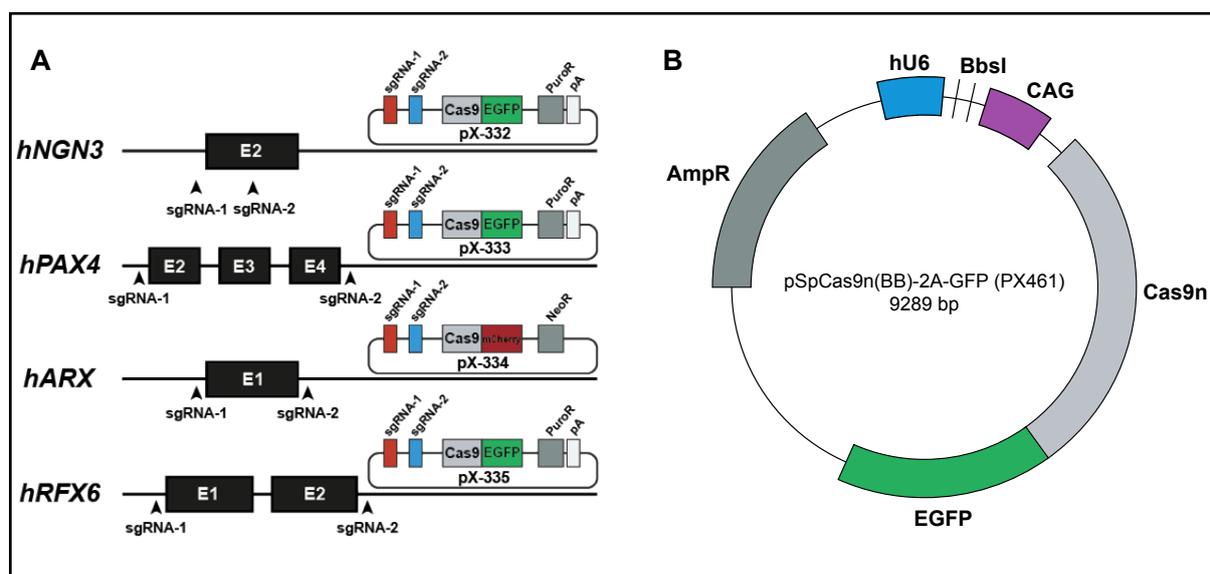
# Matériel & méthodes

## 1. CRISPR/Cas9

L'ensemble de la procédure expérimentale utilisée pour inactiver les gènes par la stratégie CRISPR/Cas9 est largement inspiré des travaux du laboratoire de Zhang publiés dans *Nature Protocol* en 2013 (Ran et al., 2013a).

### 1.1. Plasmides utilisés

Les plasmides utilisés dans le cadre de la stratégie de délétion d'exons entiers ont été clonés par l'équipe de B. Reina San Martin à l'IGBMC (**Figure 30A**). Le plasmide pX461 (**Figure 30B**) utilisé pour la stratégie de double-*nickase* provient de Addgene (#48140) et les sgRNAs #22 et #17 ciblant le codon d'initiation de *NGN3*, ont été clonés entre les sites de restriction *BbsI* d'après le protocole décrit en **1.2**.



**Figure 30. Plasmides CRISPR/Cas9**

(A) Plasmides utilisés dans la stratégie visant à la délétion d'exons entiers des gènes cibles et clonés par l'équipe de B. Reina San Martin (IGBMC). pX332 : les sgRNA-1 et sgRNA-2 ciblent la séquence codante codée par l'exon 2 de *NGN3*, pX333 : les sgRNA-1 et sgRNA-2 ciblent les exons 2, 3 et 4 de *PAX4*, pX334 : les sgRNA-1 et sgRNA-2 ciblent l'exon 1 de *ARX*, pX336 : les sgRNA-1 et sgRNA-2 ciblent les exons 1 et 2 de *RFX6*. Ils contiennent les séquences des deux sgRNAs, la séquence codant la Cas9 fusionnée à un marqueur fluorescent eGFP (*enhanced GFP*) ou mCherry, ainsi qu'une cassette de résistance à l'antibiotique puromycine ou néomycine. (B) Plasmide pX461 utilisé dans la stratégie de double-*nickase* dans lequel les séquences des sgRNA#17 ou #22, ciblant le codon d'initiation de *NGN3*, ont été clonées entre les sites de restriction *BbsI*. Il contient la séquence codant la Cas9 *nickase* (Cas9n), le marqueur fluorescent eGFP et la cassette de résistance à l'ampicilline.

## 1.2. Clonage des sgRNAs dans les plasmides pX461

Le protocole utilisé a été développé par le laboratoire de Zhang et est disponible en version détaillée sur la plateforme Addgene (<https://www.addgene.org/crispr/zhang>). Les oligonucléotides sgRNAs #22 et #17 (**Figure 31**) sens et antisens ont été phosphorylés et hybridés afin de former des duplex d'ARN. Une étape unique de digestion par les sites de restriction *BbsI* et de ligation a été réalisée pour cloner les duplex d'oligonucléotides sgRNAs dans le plasmide pX461 contenant la Cas9n et la GFP (Addgene) (**Figure 30B**). Des bactéries chimiocompétentes *Escherichia coli* TOP10 ont ensuite été transformées avec les plasmides contenant les sgRNAs (pX461-sgRNA#22 et pX461-sgRNA#17). L'ADN plasmidique des colonies a été purifié en utilisant le kit NucleoBond® Xtra Maxi (Macherey-Nagel) et analysé par séquençage (GATC Biotech) à l'aide des amorces indiquées par Addgene : Seq sgRNA F (sens) et Seq sgRNA R (antisens) pour valider la présence des sgRNAs dans les plasmides (**Figure 32**).

Gène	Cible	sgRNA	Séquence	PAM	Brin	Score Off-target	Chromosome
NGN3	5' E2	hNGN3-1	GCGCAGTCCGCGATTCCG	AGG	-	95	Chr 10
	3' E2	hNGN3-2	TTTGAGGGCCGCCCGCT	CGG	-	96	
	5' ATG	sgRNA #22	GTTGAGGCGTCATCCTACGG	CGG	-	95	
	3' ATG	sgRNA #17	TGTCCAAGTGACCCGTGAGA	CGG	+	86	
PAX4	5' E2	hPAX4-1	GCCAACTCTCCTGATCTA	AGG	-	82	Chr 7
	3' E4	hPAX4-2	CACCACCTAGTTGATGGC	AGG	+	82	
ARX	5' E1	hARX-1	TACCGCTTGCTGAGCG	CGG	+	86	Chr X
	3' E1	hARX-2	GCGTTAAACATCTAGCCA	GGG	+	83	
RFX6	5' E1	hRFX6-1	AGAAAGGACGCCGCGCAC	AGG	-	92	Chr 6
	3' E2	hRFX6-2	GGCTGTAAACGTCAAC	AGG	-	86	

**Figure 31. Tableau récapitulatif des sgRNAs utilisés pour cibler les gènes d'intérêt**  
Les régions géniques cibles des sgRNAs et leurs séquences sont indiquées. Le score *Off-target* représente la spécificité des sgRNAs : plus le score est élevé, plus la fidélité des sgRNAs est élevée.

## 1.3. Surveyor®: détection de mutation

L'efficacité de clivage de la séquence cible par la Cas9 via les différents sgRNAs a été testée grâce au kit Surveyor® Mutation Detection kit (IDT), dont le principe est décrit dans la partie 2 des résultats. Après transfection des cellules HEK293T avec les plasmides pX461-sgRNA#22 et pX461-sgRNA#17, l'ADN des cellules transfectées a été extrait et la séquence cible de *NGN3* a été amplifiée par PCR à l'aide des amorces sens F3 et antisens R3 (**Figure 32**). Une PCR contrôle a été réalisée sur des échantillons d'ADN plasmidique G et C fournis dans le kit. L'ADN a été dénaturé puis

hybridé pour former des duplex d'ADN sauvage et muté. L'ADN a ensuite été digéré par la nucléase Surveyor (nucléase S) qui reconnaît les mutations et clive l'ADN à l'endroit des mésappariements en présence de l'enhancer S et 15 mM de MgCl<sub>2</sub>. Le produit de digestion a été analysé par électrophorèse sur gel d'agarose 2% (p/v).

## 2. Matériel biologique

### 2.1. Cellules HEK293T

#### **Origine et culture**

Les cellules HEK293T (*Human Embryonic Kidney 293T*) dérivent de cellules embryonnaires humaines de rein et expriment l'antigène T du virus simien 40 (SV40) de manière stable (Graham et al., 1977). Ces cellules ont été utilisées dans le cadre des tests de fonctionnalité des sgRNAs et de la Cas9 (Surveyor). Les cellules HEK293T ont été cultivées dans des flasques T75 (Falcon) dans du milieu DMEM (*Dulbecco's Modified Eagle's Medium* ; Invitrogen) à 1 g/L de glucose, supplémenté de 10% (v/v) de FCS (*Fetal Calf Serum* ; Invitrogen), 100 mg/mL de pénicilline et streptomycine (Invitrogen) et incubées à 37 °C en présence de 5% de CO<sub>2</sub>. Les cellules ont été passées deux fois par semaine en les décollant à l'aide de trypsine diluée, puis en les diluant au 1/15<sup>ème</sup> dans du DMEM.

#### **Transfection**

Les cellules ont été transfectées avec l'ADN plasmidique grâce à la Lipofectamine™ 2000 (Invitrogen), qui permet l'endocytose de l'ADN par formation de liposomes cationiques. La veille de la transfection, les cellules ont été comptées grâce à une cellule de numération Bürker et 6,5x10<sup>4</sup> cellules ont été réparties dans du milieu DMEM sans antibiotiques. Elles ont ensuite été incubées 24h à 37°C en présence de 5% de CO<sub>2</sub>. La transfection a été effectuée selon les indications du fournisseur et lorsque les cellules ont atteint une confluence d'environ 70%. La lipofectamine et l'ADN plasmidique dilués dans du milieu Opti-MEM (Invitrogen) ont été ajoutés à un ratio 1:1. Les cellules ont ensuite été incubées pendant 24h et l'efficacité de transfection a pu être estimée grâce à l'expression de la GFP dont le gène est contenu dans le pX461.

## 2.2. hiPSCs

### **Origine et culture**

La lignée de cellules hiPSCs utilisée nous a été fournie par l'équipe de S.Viville à l'IGBMC. C'est une lignée issue de fibroblastes dermiques primaires, appelée CPRE2, provenant de la biopsie cutanée d'une femme âgée de 38 ans souffrant d'obésité. Ces cellules ont été reprogrammées en cellules pluripotentes par infection virale à l'aide de vecteurs lentiviraux portant une cassette monocistronique exprimant les gènes ONSL pour *OCT4*, *NANOG*, *SOX2*, *LIN28* (Jung et al., 2014). Les cellules hiPSCs, appelées LV37, ont été cultivées dans des boîtes de culture P35 (Falcon) recouvertes de Matrigel hESC-qualified (BD Biosciences) en présence de mTeSR1 (Invitrogen) et incubées à 37°C en présence de 5% de CO<sub>2</sub>. Le milieu a été changé tous les jours et le passage s'est fait tous les 6 jours environ. Lors des passages, les colonies sujettes à de la différenciation spontanée ont été éliminées en les grattant avec un cône de pipette. De la dispase à 1 mg/mL (Invitrogen) a été ajoutée pour décoller les cellules qui ont été grattées à l'aide d'un scraper (Dutscher). Les colonies ont été broyées en faisant des allers-retours avec une pipette et les particules formées, réensemencées de manière homogène sur du Matrigel en présence de mTeSR1. Les produits notamment nécessaires à la culture des hiPSCs sont répertoriés en **figure 34**.

### **Transfection**

La transfection des hiPSCs a été réalisée grâce au kit Amaxa® Human Stem Cell Nucleofector® Starter Kit (Lonza). 2-3h avant la transfection, le milieu des cellules a été remplacé par du mTeSR1 additionné de Rock inhibitor Y27632 à 1/1000 (Invitrogen). Les cellules ont été rincées avec du PBS 1X stérile puis détachées avec de l'acutase (Invitrogen) et incubées 6 min à 37°C afin d'obtenir des cellules isolées. L'acutase a été inactivée à l'aide de milieu F15 (4,5 g/L DMEM-F12, 15 % sérum ES, pénicilline et streptomycine 100X, 0,1 mM β-mercaptoéthanol). Après centrifugation pendant 4 min à 140g, les culots ont été repris dans du mTeSR1 additionné de Y27632 (1/1000) et le nombre de cellules a été compté à l'aide du Z1 Coulter Counter (Beckman Coulter). 800 000 cellules ont été utilisées par transfection, et centrifugées 4 min à 140g. Le culot a été repris dans 100 µL de solution de nucléofection 2 supplémentée. 5 µg d'ADN total ont été ajoutés aux cellules qui ont été électroporées avec le programme B016 du Amaxa® Nucleofector®. Le contenu des cuvettes a été transféré dans du mTeSR1 additionné de Y27632 (1/1000) etensemencé dans des boîtes de culture P60 recouvertes de Matrigel. Les cellules GFP<sup>+</sup> ont été triées au FACS le lendemain et réensemencées.

### Génotypage des clones

Les hiPSCs ont été récoltées dans du tampon de lyse (100 mM Tris HCl pH 8,5, 200 mM NaCl, 5 mM EDTA, 0,2 % SDS) après avoir été rincées au PBS 1X. L'extraction au phénol-chloroforme de l'ADN génomique a été réalisée puis la concentration d'ADN a été mesurée dans du tampon TE 1X (10 mM Tris HCl pH 8, 1 mM EDTA) au NanoDrop™ 2000 (Thermo Scientific). 100 ng d'ADN ont été utilisés par réaction d'amplification PCR réalisée avec la polymérase Expand High Fidelity (Roche) et 5% de DMSO. Les différentes amorces ciblant les gènes d'intérêt sont listées dans le tableau de la **figure 32**. Les stratégies de génotypage seront présentées pour chaque gène dans la partie des résultats dédiée. Les produits PCR ont ensuite été analysés par électrophorèse sur gel d'agarose 1% (p/v) et purifiés grâce au kit NucleoSpin Gel and PCR Clean-up (Macherey-Nagel) avant séquençage de l'ADN (GATC Biotech).

	Nom de l'amorce	Séquence 5' → 3'	Fonction
NGN3	F1	AGAACGCTGGGATACTATCTACAA	Génotypage clones hiPSCs
	F2	CTAGTCAGGACAATGAGCTCGT	
	F3	CTTTGTCCGGAATCCAGCTG	
	R1	CATAGGAGAGAAAGAGAGATGCAAG	
	R2	CTCCAGAAAGAGTGTCATCCCT	
	R3	TACAAGCTGTGGTCCGCTAT	
	R4	TCAGTGCCAACTCGCTCTTAG	
ARX	F1	CTTCGGCCAACACTAAAGGG	
	F2	CCAACACACACCCATCCAT	
	R1	TGAGGATAACAGGTTCCGGGG	
RFX6	F1	GTGCCCAATAGTGTAACCTGAAAG	
	F2	CTGGGGCTCCGCTGGGGAAAC	
	R1	CACAGTTCATCTTATGGAGTATACAC	
	R2	TTAACTTCAAAGGAGTAGCCTACTGG	
	R3	CCAGGTACACTGTTTCTTCC	
PAX4	F1	AGATCCAGTCCTCAGCCAAC	
	F2	TCTGGGACTTGGGTCCTTC	
	F3	GTCTTATAGCCTCCTCTGTCC	
	R1	AGGACTCTCTGACCCTCCAT	
	R2	ACCTCAATTTTCTTCTGCTGGG	
pX461	Seq sgRNA F	GAGGGCCTATTTCCCATGAT	
	Seq sgRNA R	GGGCGTACTTGGCATATGAT	

**Figure 32. Amorces utilisées pour les amplifications PCR du génotypage des clones hiPSCs et pour le séquençage du plasmide pX461**  
F (*Forward*) = amorce sens; R (*Reverse*) = amorce antisens

## 2.3. Génération et maintien des HIOs

### 2.3.1. Différenciation intestinale *in vitro*

Une version extensive du protocole de différenciation *in vitro*, rédigée à partir du protocole publié par Spence et al. (2011), est disponible en **annexe 2**. En résumé, les hiPSCs ont été passées lorsqu'elles ont atteint une confluence de 70–80%. Les cellules ont été passées avec de l'accutase de manière à obtenir des cellules isolées. La suspension cellulaire a été placée dans un Falcon de 15 mL et centrifugée à 200g pendant 4 min. Le surnageant a été retiré et le culot resuspendu dans du mTeSR1 supplémenté de Y-27632 (1/1000). Les cellules ont été homogénéisées et un aliquot a été récupéré afin de compter les cellules à l'aide d'une cellule de Neubauer. 420 000 cellules ont étéensemencées par P35 avec du mTeSR1 supplémenté de Y-27632 (1/1000). Les cellules ont été réparties de manière homogène dans les boîtes de culture et cultivées pendant 2 à 3 jours pour atteindre une confluence de 90-95%. La différenciation a ensuite été initiée. Les étapes de la différenciation et les milieux utilisés sont détaillés sur la **figure 33**. L'ensemble des produits utilisés dans le cadre de la culture et la différenciation des hiPSCs est listé en **figure 34**.

### 2.3.2. Passage des HIOs

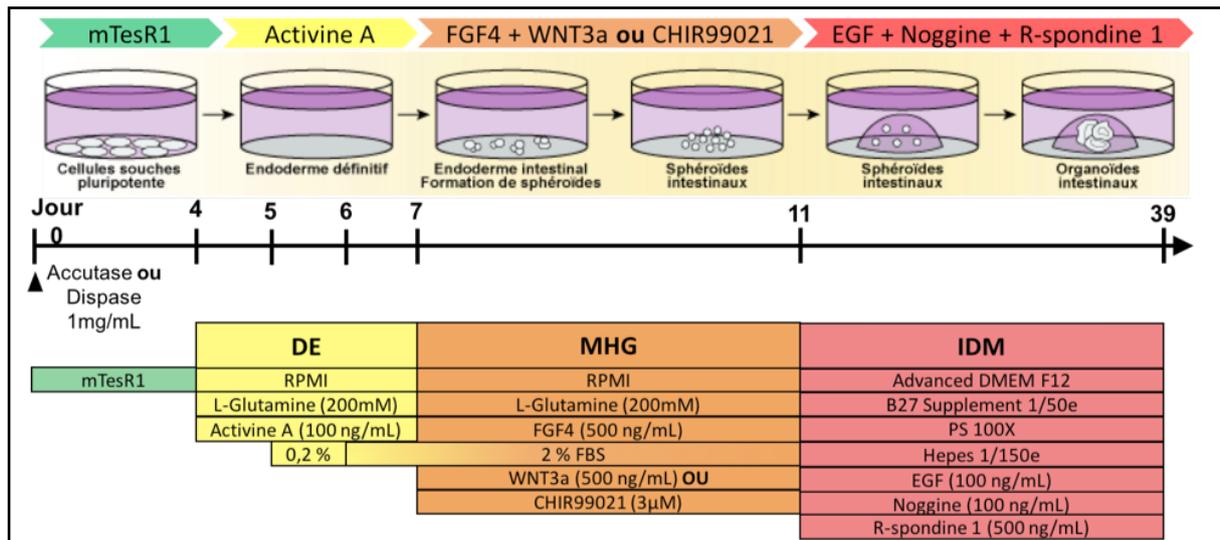
Les HIOs ont été passés après environ 14 jours de culture pour permettre leur expansion. Ils ont été mécaniquement dissociés du Matrigel, rincés dans du milieu Advanced DMEM-F12 et coupés en deux à l'aide d'un scalpel stérile. Ils ont ensuite été réensemencés dans un nouveau dôme de Matrigel avec du milieu IDM dans de nouvelles plaques 4 puits (nunc) à raison de 8 HIOs par puits, ou récoltés dans le cadre des analyses.

## 3. Analyses transcriptomiques

### 3.1. Extraction d'ARN

#### *HIOs*

Les HIOs ont été mécaniquement dissociés du Matrigel, récoltés puis rincés au PBS 1X. L'extraction d'ARN a été réalisée avec le kit RNeasy Mini (Qiagen). Les HIOs ont été placés dans un tube eppendorf contenant 350µL de tampon de lyse RLT et ont été broyés et homogénéisés à l'aide d'une aiguille et d'une seringue. Les échantillons ont ensuite été traités selon le protocole du manuel d'utilisation du kit. Un échantillon correspond à l'ARN extrait d'une quinzaine de HIOs.



**Figure 33. Protocole de différenciation *in vitro* des hiPSCs en HIOs**

L'échelle temporelle indique la chronologie des différents stades de différenciation, J0 étant ici l'ensemencement des hiPSCs. Les milieux utilisés et leur composition sont détaillés. DE : *Definitive Endoderm*, MHG : *Mid-Hind Gut*, IDM : *Intestinal Differentiation Medium*.

Produit	Fournisseur	Référence
Noggin	R&D	6057-NG-100
EGF	R&D	236-EG-01M
R-Spondine 1	Peptotech	120-38
Wnt-3a	R&D	5036-WN-010
Activine A	R&D	338-AC-010
FGF-4	R&D	235-F4-025
CHIR99021	Tocris	4423
B27	Thermo Fisher	17504044
Y27632	Miltenyi Biotec	130-095-563
Matrigel hESC-qualified matrix	BD	354277
Matrigel BMM GFR Phenol Red-free	Corning	356231
HEPES	service culture/Invitrogen	15630-080
Dispase	Invitrogen	17105-041
DMEM-F12	Invitrogen	11320-074
Advanced DMEM	Invitrogen	12634-010
RPMI 1640	Invitrogen	61870-010
dFBS	Thermo Fisher	10270098
L-Glutamine	Sigma	G7513
mTeSR1 basal medium	Stemcell Technologies	85851
mTeSR1 5X supplement	Stemcell Technologies	85852
p/s	Invitrogen	15140-122
plaques 4 puits Nunc	Thermo Fisher	176740
scraper	Dutscher	199010

**Figure 34. Liste des produits utilisés pour la culture et la différenciation *in vitro* des hiPSCs en HIOs**

La qualité et la concentration des ARNs ont ensuite été déterminées par la plateforme GenomEast à l'IGBMC.

### hiPSCs

Les hiPSCs ont été récoltées dans du trizol (TRI Reagent®, Life Technologies) et l'ARN a été extrait selon le manuel d'utilisation fourni par le fabricant. L'ARN a ensuite été resuspendu dans de l'eau sans RNAses.

### 3.2. RT-qPCR

Les ARNs extraits ont été traités à la DNase I (Roche) afin d'éviter les contaminations avec de l'ADN génomique avant l'étape de rétrotranscription. Les sondes et amorces utilisées pour cibler les gènes d'intérêt sont des sondes UPL (*Universal Probe Library*, Roche) (**Figure 35**). Les réactions ont été effectuées avec le kit LightCycler® 480 Probes Master (Roche) et suivant un programme réalisé dans le LightCycler 480 II (Roche). Les cycles seuils (Ct) des gènes d'intérêt ont été normalisés par rapport à ceux du gène de référence *RPLP0* ( $\Delta Ct : Ct_{\text{gène}} - Ct_{\text{RPLP0}}$ ) et le niveau d'expression relatif a été calculé selon la formule  $2^{-\Delta Ct}$ .

### 3.3. RNA-Seq

Les ARNs extraits ont été confiés à la plateforme GenomEast à l'IGBMC qui a entrepris le séquençage par la technologie Illumina HiSeq 4000 Sequencing. Les données brutes ont ensuite été traitées et analysées par la bioinformaticienne de l'équipe, C. Vagne qui s'est également chargée de la rédaction de cette section du matériel et méthodes.

Pour les deux jeux de données de RNASeq (S18202 - NGN3<sup>-/-</sup> et S18103 - RFX6<sup>-/-</sup>), la procédure d'analyse a été la suivante. Les fichiers fastq ont été filtrés : les adaptateurs et les séquences de mauvaise qualité ont été supprimés à l'aide du logiciel cutadapt

Gène	Séquence 5' → 3'	RefSeq
CDX2	ATCACCATCCGGAGGAAAG TGCGGTTCTGAAACCAGATT	NM_001265.4
CHGA	GCGGTTTTGAAGATGAACTCTC GCTCTCCACCGCCTCTT	NM_001275.3
FABP2	CAATCTAGCAGACGGAAGTCAA CCGTTTGAATTTTCCAATAAGTTT	NM_000134.3
LYZ	CCGCTACTGGTGAATGATGG CATCAGCGATGTTATCTTGACAG	NM_000239.2
MUC2	GCTGCTATGTCGAGGACACC GGGAGGAGTTGGTACACACG	NM_002457.3
NGN3	TTTTGCGCCGGTAGAAAG GGGCAGGTCACTTCGTCTT	NM_020999.2
RFX6	CTTGAGCAACACAGGAGCTG TGGGTAGTTTATGTTGCTCCA	NM_173560.1
ARX	GCACCACGTTACCAGCTA CAGCCTCATGGCCAGTTC	NM_139058.2
LMX1A	TGGAGGAGAAGTCCAAAGC AGCCCTCGCAGACAGACTT	NM_177398.3
NKX2.2	CGAGGGCCTCAGTACTCC GGGGACTTGGAGCTTGAGT	NM_002509.3
TPH1	TTGGAAGATGTCTCCAACCTTTTA AAACCTGATAAGAAATCTCTGGTG	NM_004179.2
SST	ACCCAGACTCCGTCAGTTT ACAGCAGCTCTGCCAAGAAG	NM_001048.3
OCT4	CAATTTGCCAAGCTCCTGA AGATGGTCGTTGGCTGAAT	
NANOG	AGATGCCTCACACGGAGACT TTTGCGACTCTTCTCTGC	NM_024865.2
SOX2	GGGGGAATGGACCTTGATAG GCAAAGCTCTACCGTACCA	NM_003106.3
LIN28A	GAAGCGCAGATCAAAGGAG GCTGATGCTCTGGCAGAAGT	NM_024674.4
RPLP0	TCTACAACCCTGAAGTGCTTGAT CAATCTGCAGACAGACTGG	NM_001002.3

**Figure 35. Amorces UPL utilisées pour les expériences de RT-qPCR**

(Martin, 2011). Les *reads* s'alignant sur les séquences d'ARNr (identifiés avec Bowtie2 (Langmead et Salzberg, 2012)) ont également été écartés de l'analyse. Après ce prétraitement, les *reads* ont été alignés sur le génome humain hg38 à l'aide du logiciel STAR (Dobin et al., 2013). La quantification du nombre de *reads* par gène a été réalisé avec HTSeq (Anders et al., 2015) et la base de données d'annotation Ensembl version 90 (Aken et al., 2017). La normalisation des données de comptage a été effectuée à l'aide de la méthode implémentée dans le *package* R: DESeq2 (Love et al., 2014). Pour représenter les données, une analyse en composante principale (ACP/PCA) et une *heatmap* fondée sur le coefficient SERE (Schulze et al., 2012) (qui permet de mesurer de manière fiable la différence entre deux échantillons de RNA-Seq) ont été réalisées.

L'analyse de l'expression différentielle a également été effectuée avec DESeq2 (Love et al., 2014). Elle a permis d'identifier les gènes différentiellement exprimés entre les HIOs RFX6<sup>-/-</sup> et CTRL d'une part (S18202), ainsi qu'entre les HIOs NGN3<sup>-/-</sup> et CTRL d'autre part (S18103).

Concernant le projet S18103 (RFX6<sup>-/-</sup> vs CTRL), l'effet *batch* a été pris en compte dans le modèle de DESeq2, car les deux échantillons provenant d'une expérience différente (LNMR102 et LNMR103) présentaient un profil différent, comme le suggéraient l'ACP et le coefficient SERE. L'effet *batch* n'a en revanche pas été pris en compte pour le projet S18202 (NGN3<sup>-/-</sup> vs CTRL). En effet, l'ACP et le coefficient SERE indiquait qu'il n'existait pas d'effet *batch* affectant les données. Une analyse différentielle complémentaire, toujours réalisée avec DESeq2, a montré qu'aucun gène n'est significativement différentiellement exprimé entre les deux *batches*.

L'analyse fonctionnelle a été réalisée avec le *package* R GAGE (Luo et al., 2009), en prenant comme données d'entrées les log fold-change (logFC) calculés par DESeq2 (après une opération de « *shrinkage* »). Cette méthode permet de déterminer si un ensemble de gènes ayant une fonction commune est sur-représenté parmi les gènes différentiellement exprimés. Les ensembles de gènes testés sont ceux répertoriés dans la base de données MSigDB version 6.2. Les gènes sur- et sous-exprimés ont été analysés de manière indépendante.

## 4. Histologie

### 4.1. Immunofluorescences sur HIOs entiers/*in toto* (*whole-mount*)

Les HIOs ont été mécaniquement dissociés du Matrigel, récoltés puis rincés au PBS 1X. Ils ont ensuite été fixés dans du PFA 4% 2h à 4°C, puis rincés 3 fois 10 min dans du PBS + 0,2% Triton. Les HIOs ont été incubés dans du Tris-HCl pendant

5 min à température ambiante, puis 15 min à 70°C pour le démasquage des sites antigéniques. Ils ont à nouveau été lavés 3 fois 10 min dans du PBS + 0,2% Triton. L'étape de blocage a été réalisée dans du PBS + 0,2% Triton + 5% de sérum d'âne pendant 2h à température ambiante. L'anticorps primaire (**Figure 36A**) a été dilué dans la solution de blocage et incubé pendant 1 à 3 jours à 4°C. Des lavages ont ensuite été réalisés dans la solution de blocage 3 fois 45 min à température ambiante, puis dans du PBS + 0,2% Triton 3 fois 20 min et enfin 2h à nouveau dans la solution de blocage. L'anticorps secondaire (**Figure 36B**) dilué dans la solution de blocage a été ajouté pendant 2 à 4 jours à 4°C. Les HIOs ont ensuite été lavés dans du PBS + 0,2% Triton 3 fois 30 min et montés dans une boîte de Pétri spécifiquement conçue par la plateforme d'imagerie de l'IGBMC pour imager les HIOs montés dans un dôme de UltraPure™ Low Melting Point Agarose (Invitrogen) 0,8%.

<b>A</b>				
Antigène cible	Origine	Référence	Concentration	
CDX2	souris	<i>BioGenex AM392-5M</i>	1/500	
CHGA	lapin	<i>DakoCytomation A0430</i>	1/500	
NGN3	mouton	<i>R&amp;D AF3444</i>	1/200	
GLP1	chèvre	<i>Santa Cruz sc-7782</i>	1/200	
GIP	chèvre	<i>Santa Cruz sc-23554</i>	1/200	
5-HT	chèvre	<i>Abcam ab66047</i>	1/2000	
SST	chèvre	<i>Santa Cruz sc-7819</i>	1/200	
E-CADH	lapin	<i>Abcam ab53033</i>	1/300	
FOXA2	souris	<i>Abcam ab60721</i>	1/500	
SOX9	lapin	<i>AbCys SA VPA 5535</i>	1/500	
SOX17	chèvre	<i>R&amp;D AF1924</i>	1/500	

<b>B</b>				
Espèce cible	Fluorochrome	Origine	Référence	Concentration
souris	DyLight 488	âne	<i>Jackson 715-485-150</i>	1/500
	DyLight 549		<i>Jackson 715-506-150</i>	1/500
chèvre	DyLight 488		<i>Jackson 705-485-147</i>	1/500
	DyLight 549		<i>Jackson 705-505-147</i>	1/500
lapin	DyLight 488		<i>Jackson 711-546-152</i>	1/500
	DyLight 549		<i>Jackson 711-505-152</i>	1/500
mouton	Cy3		<i>Jackson 713-165-147</i>	1/500

**Figure 36. Anticorps utilisés pour les immunofluorescences et analyses au FACS (A) Anticorps primaires (B) Anticorps secondaires**

## 4.2. Transparisation

La transparisation a été réalisée après l'immunofluorescence et le montage des HIOs. Elle consiste en des bains successifs de fructose à concentration croissante sous agitation : 45% (p/v) pendant 4h, 70% pendant 4h puis 90% sur la nuit. Un bain supplémentaire à 110% de fructose peut être réalisé pour les échantillons plus denses. Les solutions de fructose ont été additionnées de 0,5% de  $\alpha$ -thioglycérol. Les HIOs ont été imagés dans la dernière solution de fructose utilisée.

## 4.3. Imagerie

Les images de culture cellulaire et de différenciation ont été prises avec un microscope optique à lumière transmise inversé (Leica). Les immunofluorescences ont été imagées avec un microscope optique à épifluorescence inversé (DMIRE2, Leica) ou un macroconfocal (LSI, Leica) ou encore un microscope confocal droit (SP5, Leica). Les images d'essais de transparisation des HIOs ont été prises avec un microscope confocal multiphoton droit (SP8MP, Leica).

## 5. Analyse au FACS

Les cellules ont été rincées avec du PBS 1x et dissociées avec de la dispase dans un tampon contenant du HBSS -Ca, -Mg, FCS et EDTA 0,5M. La suspension cellulaire a été passée dans un filtre 70  $\mu$ m et les cellules ont ensuite été fixées dans du PFA 4% 20 min à 4°C. Les cellules fixées ont été perméabilisées dans du PBS contenant 5% de DS et 0,2% de Triton X100 pendant 30 min à 4°C. Les cellules ont été incubées avec l'anticorps primaire (**Figure 36A**) dans du PBS contenant 5% de DS et 0,1% de Triton X100 sur la nuit à 4°C. Le lendemain, les cellules ont été lavées avec du PBS puis incubées avec l'anticorps secondaire (**Figure 36B**) dans du PBS contenant 0,1% de Triton X100 pendant 45 min à température ambiante. Les cellules ont été analysées au FACS BD LSRFortessa.

## 6. Analyses statistiques

L'analyse statistique des données de RT-qPCR a été faite à l'aide du logiciel Graphpad Prism, en utilisant le test non-paramétrique U de Mann-Whitney, en raison des faibles effectifs étudiés. Les différences sont jugées significatives si  $p < 0,05$ . Les données sont représentées par la moyenne  $\pm$  erreur type.

# RÉSULTATS

# Résultats

La première partie de cette section dédiée aux résultats concernera la mise au point et l'optimisation du protocole permettant l'obtention des HIOs à partir des hiPSCs. La deuxième présentera les étapes de la génération des différents clones hiPSCs déficients pour les gènes codant les facteurs de transcription d'intérêt (NGN3, RFX6, ARX et PAX4). Les troisième et quatrième parties décriront les résultats obtenus par l'analyse phénotypique des HIOs dérivés des clones NGN3<sup>-/-</sup> et RFX6<sup>-/-</sup> respectivement.

## Partie 1. Optimisation de la production d'organoïdes intestinaux humains (HIOs)

Durant les dernières décennies, la recherche en biologie du développement a permis d'élucider de nombreux mécanismes du développement des vertébrés, notamment grâce à des modèles d'étude comme le poisson-zèbre, le poulet, le xénope ou encore la souris. Ces découvertes ont pu être exploitées *in vitro* suite à l'émergence des cultures de cellules souches embryonnaires humaines (Thomson et al., 1998) et la reprogrammation des cellules telles que les fibroblastes, en cellules souches pluripotentes humaines induites (Takahashi et al., 2007). En effet, nos connaissances des voies de signalisation et des cascades de régulation transcriptionnelle impliquées dans le développement *in vivo* nous permettent maintenant de mimer le développement *in vitro* et ainsi, de générer des tissus spécifiques afin d'étudier le développement chez l'homme et certains processus physiologiques et pathologiques.

C'est dans ce contexte qu'un protocole robuste permettant de diriger la différenciation de cellules souches pluripotentes humaines (hESCs et hiPSCs) en tissu intestinal *in vitro* a été publié en 2011 par l'équipe de J.Wells (Cincinnati Children's Hospital; Spence et al., 2011). Il offre la possibilité de récapituler le développement intestinal embryonnaire par l'utilisation consécutive de facteurs de croissance agissant sur les voies de signalisation connues pour leur implication dans le développement intestinal chez les vertébrés (Zorn et Wells, 2009) et un système de culture favorisant la croissance tridimensionnelle et la morphogenèse intestinale (Spence et al., 2011; McCracken et al., 2011; Sato et al., 2009).

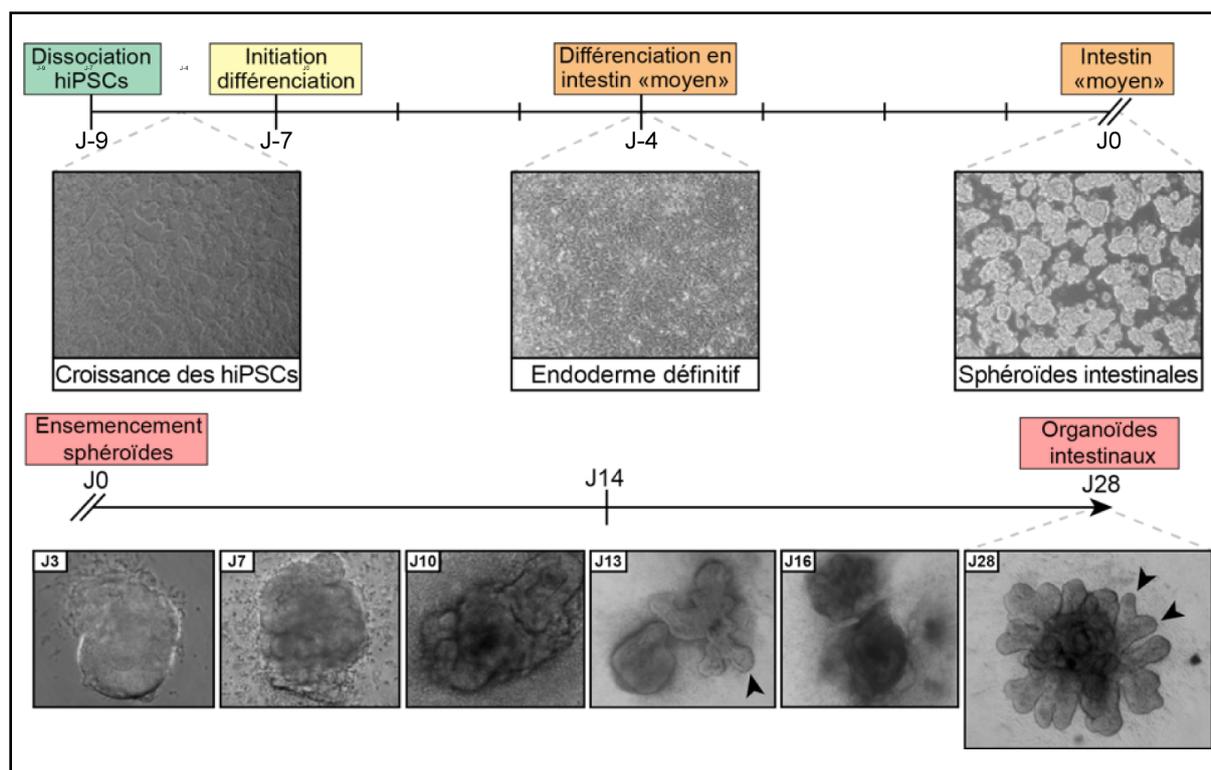
Anthony Beucher, un ancien post-doctorant de l'équipe, a pu mettre en place cette technologie au laboratoire suite à une formation dans le laboratoire de J.Wells. Cependant, au commencement de ce projet de thèse, la qualité et l'efficacité de la différenciation n'avait pas été établies.

De plus, le protocole initial datant de 2011 a depuis été révisé pour en améliorer l'efficacité et la fidélité du modèle (Workman et al., 2017; Watson et al., 2014). De ce fait, le protocole de différenciation a d'abord été testé et optimisé dans nos conditions expérimentales, avant d'être utilisé en tant que modèle d'étude de la régulation transcriptionnelle de la différenciation entéroendocrine.

## 1. Contrôle qualité des étapes clés du protocole de différenciation des HIOs *in vitro*

Les étapes clés du développement récapitulées dans le protocole *in vitro* incluent la formation de l'endoderme définitif, la régionalisation postérieure de l'endoderme, la formation du tube digestif et la croissance et morphogénèse intestinales. Ces étapes sont récapitulées sur la frise temporelle en **figure 37**.

L'efficacité de la différenciation peut dépendre de la lignée cellulaire utilisée (Spence et al., 2011) mais également de divers paramètres au cours des différentes étapes du protocole, comme la densité cellulaire, la stabilité des facteurs de croissance et la précision de la durée d'exposition aux différents milieux.



**Figure 37. Chronologie des étapes de la différenciation intestinale *in vitro***

Frise représentant la chronologie des étapes clés de la différenciation intestinale *in vitro*. Les graduations représentent une journée et les encadrés sont des images obtenues à l'aide d'un microscope optique, sauf pour les sphéroïdes intestinales qui ont été imagées avec un macroscopie. Ces images sont représentatives des différents stades de différenciation. Les flèches noires indiquent les bourgeons de cryptes se formant durant la morphogénèse des sphéroïdes en HIOs.

C'est pourquoi il est nécessaire de caractériser la différenciation *in vitro* en tissu intestinal de notre lignée hiPSCs (LV37; Jung et al., 2014) dans nos conditions expérimentales. Tout d'abord, l'influence de l'enzyme utilisée pour le passage des hiPSCs sur la densité cellulaire a été déterminée.

### 1.1. Influence de l'enzyme de dissociation sur la densité cellulaire

L'ensemencement des cellules représente une étape clé quant à l'efficacité ultérieure de formation de l'endoderme définitif. En effet, le nombre de cellules et leur répartition homogène dans la boîte de culture sont des paramètres essentiels à prendre en compte avant l'initiation de la différenciation. Il a pu être observé que si la densité des colonies formées est trop faible, la différenciation est inefficace et à l'inverse, qu'une densité trop importante mène à des régions indifférenciées à l'issue de la formation de l'endoderme définitif (non montré). Initialement, l'enzyme dispase était utilisée pour détacher les colonies sous forme de « particules » avant de les réensemencer, mais la répartition des cellules n'était pas homogène et la confluence de 85-90 % nécessaire pour débiter la différenciation n'était pas toujours atteinte. La dispase est une metalloprotéase non-spécifique qui agit en clivant notamment la fibronectine et le collagène IV par hydrolisation de liaison peptidique de résidus d'acides aminés non-polaires. Ces protéines sont des constituants majeurs de la matrice extracellulaire, nécessaires à l'adhésion cellule-matrice (Volovitz et al., 2016). Cependant, l'absence de quantification ne nous permettait pas d'ensemencer un nombre précis de cellules et leur croissance n'était donc pas reproductible entre les différentes boîtes de culture et expériences de différenciation. Un nombre insuffisant de cellules ensemencées empêchait d'atteindre la confluence adéquate mais un nombre trop important menait à une différenciation spontanée et prématurée des hiPSCs.

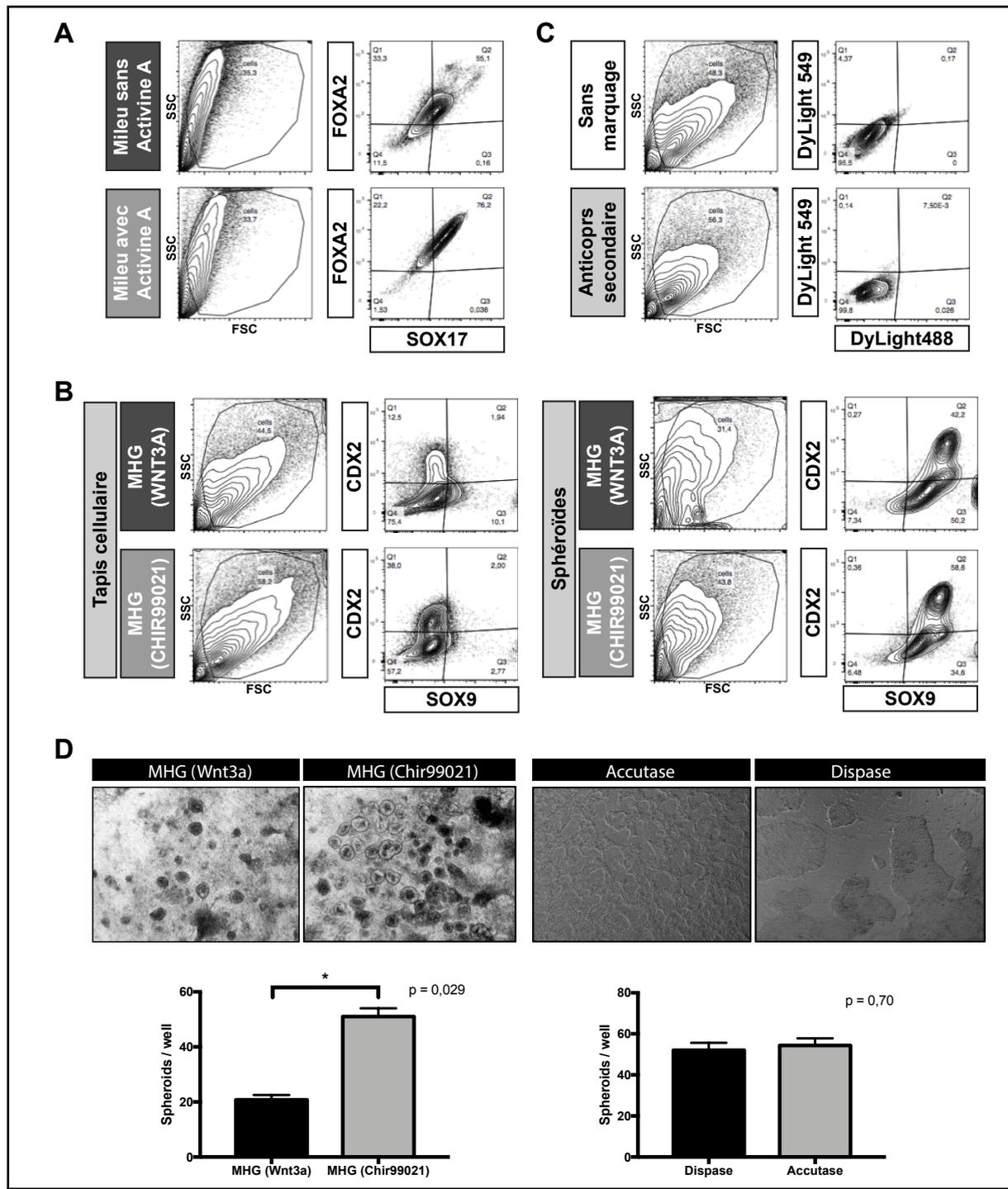
Sur les conseils du Dr. Maxime Mahé (UMR Inserm 1235, Nantes) et en adéquation avec la publication de Watson et al. (2014), une enzyme permettant de passer les cellules sous forme de cellules uniques, l'accutase, a été utilisée pour ensemencer un nombre de cellules précis, à savoir 100 000 cellules par cm<sup>2</sup>. L'accutase est constituée d'une combinaison d'enzymes naturelles mimant l'action protéolytique de la trypsine et de la collagénase (Zhang et al., 2012). Elle agit en dissociant les points d'interaction cellule-cellule et cellule-matrice par clivage des protéines d'adhésion cellulaire telles que les cadhérines. Comme observé sur le 4<sup>ème</sup> panneau de la **figure 38D**, deux jours après l'ensemencement des cellules passées à la dispase, il se forme des colonies qui ne sont pas encore suffisamment confluentes pour initier le protocole et qui ne sont pas uniformes. A contrario, les cellules passées à l'accutase sont bien réparties au sein de la boîte et ont atteint une confluence de 85-90% (**Figure 38D**, 3<sup>ème</sup> panneau).

Le passage des cellules à l'accutase et l'ensemencement d'un nombre prédéterminé de cellules semblent donc optimal quant au contrôle des paramètres de l'initiation du protocole de différenciation. Les sphéroïdes, structures assimilables à l'intestin primitif formées *in vitro*, ont été quantifiées dans les deux conditions. Cette quantification sera présentée dans la partie **1.3** détaillant la formation de ces structures. Une fois les cellules ayant atteint une confluence de 85-90%, le milieu permettant la formation de l'endoderme définitif peut être ajouté.

## 1.2. Formation de l'endoderme définitif

Pour générer des cellules intestinales à partir des hiPSCs, une étape mimant le processus de gastrulation et la formation de l'endoderme définitif est requise. La voie de signalisation Nodal est nécessaire et suffisante pour initier le développement de l'endoderme et du mésoderme chez tous les vertébrés (D'amour et al., 2005). Les ligands de cette voie sont des membres de la famille du TGF- $\beta$  qui vont interagir avec des co-facteurs transcriptionnels pour activer la transcription de gènes régulant le développement du mésendoderme. Une forte concentration de Nodal favorise le développement en endoderme, tandis que de faibles doses induisent le développement du mésoderme (Zorn et Wells, 2009). De ce fait, la différenciation des hiPSCs en endoderme définitif implique un traitement de 3 jours avec un milieu contenant une forte concentration d'activine A (100 ng/mL), qui est un activateur de la voie de signalisation Nodal. L'utilisation préférentielle de l'activine A par rapport à Nodal s'explique par une activité accrue et une meilleure efficacité de l'activine A dans la différenciation des hiPSCs en endoderme (D'amour et al., 2005). Ces 3 jours de traitement entraînent la formation d'un tapis de cellules différenciées en endoderme définitif (**Figure 37**) et exprimant des marqueurs caractéristiques comme FOXA2 et SOX17 (Stainier, 2002).

L'efficacité de la formation de l'endoderme a été évaluée après traitement des hiPSCs avec le milieu contenant l'activine A et un milieu sans activine A, composé uniquement du milieu basal avec la L-glutamine et les concentrations croissantes de sérum FBS (**Figure 33**). Un immunomarquage de FOXA2 et SOX17 a été réalisé sur l'ensemble des cellules constituant le tapis cellulaire et les cellules marquées ont ensuite été analysées au FACS (**Figure 38A**). Il faut néanmoins prendre en compte le fait que cette expérience a été réalisée lorsque les hiPSCs étaient dissociées à l'aide de la dispase et donc la densité et l'homogénéité des cellules n'étaient pas encore maîtrisées, ce qui peut entraver l'efficacité de la différenciation. L'analyse au FACS des immunomarquages sur les cellules traitées avec le milieu sans activine A a mis en évidence 55% de cellules double-positives FOXA2<sup>+</sup> SOX17<sup>+</sup> contre 76% avec les cellules traitées à l'activine A (**Figure 38A**).



**Figure 38. Contrôle de l'efficacité des étapes de la différenciation intestinale *in vitro***

(A-C) Analyse au FACS des immunomarquages : (A) Formation de l'endoderme définitif (FOXA2<sup>+</sup> SOX17<sup>+</sup>) en présence d'activine A ou non. (B) Postériorisation de l'endoderme et formation des sphéroïdes intestinales (CDX2<sup>+</sup> SOX9<sup>+</sup>) : Marquage des cellules constituant le tapis cellulaire après traitement avec WNT3A ou CHIR99021 (panneau de gauche), Marquage des sphéroïdes intestinales formées avec WNT3A ou CHIR99021 (panneau de droite). (C) Contrôle sans marquage et avec anticorps secondaire uniquement. (D) Comparaison de l'efficacité de formation des sphéroïdes intestinales par l'utilisation de WNT3A ou CHIR99021 : images représentatives de l'aspect des sphéroïdes et quantification par comptage (panneau de gauche); comparaison de la croissance des hiPSCs selon l'enzyme de dissociation (accutase ou dispase) et analyse quantitative par comptage des sphéroïdes formées (panneau de droite). \* $p < 0,05$ , différence significative entre les traitements (test de Mann-Whitney). Images obtenues avec un microscope optique. MHG = Mid-Hindgut.

Il est étonnant d'observer une telle fraction de cellules FOXA2<sup>+</sup> SOX17<sup>+</sup> en l'absence d'activine A et ceci suggère une différenciation spontanée des hiPSCs ou une contamination du milieu avec de l'activine A. Afin d'exclure cette éventualité, il sera nécessaire de reproduire cette expérience. La fluorescence observée correspond bien à un marquage spécifique des cellules par l'anticorps primaire comme le révèle l'absence de fluorescence dans les contrôles avec anticorps secondaires uniquement et sans marquages (**Figure 38C**). D'après ces observations, une efficacité de 76% de formation en endoderme définitif est obtenue, ce qui est légèrement inférieur à l'efficacité moyenne de 80-90% publiée dans le protocole initial (Spence et al., 2011), sachant que la lignée cellulaire utilisée diffère et peut avoir une influence sur l'efficacité de différenciation. De plus, la différenciation ayant été initiée sur des hiPSCs dont la croissance n'était pas contrôlée, l'efficacité du développement de l'endoderme définitif peut potentiellement être réduite dans l'expérience analysée. Diverses expériences de différenciation nous ont permis d'observer la formation d'un tapis cellulaire homogène, une viabilité cellulaire accrue et une meilleure reproductibilité après passage des hiPSCs à l'accutase. Les analyses du suivi de la différenciation ayant été réalisées préalablement au changement d'enzyme de dissociation des hiPSCs, ces observations n'ont pas été quantifiées. Dans tous les cas, l'endoderme définitif semble être formé de manière efficace et la postériorisation de l'endoderme a ensuite pu être initiée.

### 1.3. Formation de l'intestin « moyen » : les sphéroïdes intestinales

Pendant et après la gastrulation, des événements de régionalisation des axes antéro-postérieur et dorso-ventral se mettent en place. La détermination de l'axe antéro-postérieur de l'endoderme permet d'établir un tube digestif primitif composé d'une région antérieure et d'une région postérieure. L'intestin primitif moyen et postérieur donnera ensuite naissance à l'intestin grêle et au gros intestin. Il y a 4 voies de signalisation principales qui permettent la régionalisation postérieure de l'embryon chez les vertébrés : Wnt, FGF, BMP et l'acide rétinoïque; en régulant notamment l'expression des gènes de la famille des facteurs de transcription Cdx. Des études dans divers organismes modèles ont démontré que les voies de signalisation FGF et Wnt permettent la régionalisation postérieure de l'endoderme tout en inhibant le développement antérieur (Tsai et al., 2016; Zorn et Wells, 2009).

Ainsi, une fois l'endoderme formé, les cellules sont traitées pendant 4 jours avec le cocktail de facteur de croissance FGF4 + WNT3A pour permettre la régionalisation postérieure de l'endoderme, en accord avec le protocole publié en 2011. Cependant, une publication plus récente a mentionné l'efficacité accrue de la molécule CHIR99021 par rapport à WNT3A (Workman et al., 2016).

CHIR99021 est un dérivé d'aminopyrimidine, inhibiteur sélectif de la kinase GSK3. L'inhibition de GSK3 empêche la phosphorylation de la  $\beta$ -caténine, nécessaire à sa dégradation par le protéasome. Ainsi, l'inhibition de GSK3 correspond à une activation de la voie canonique Wnt. De ce fait, dans le but d'optimiser notre différenciation *in vitro*, les deux combinaisons possibles FGF4 + WNT3A ou FGF4 + CHIR99021 ont été utilisées afin d'induire la régionalisation postérieure de l'endoderme définitif formé au préalable et leur efficacité a été comparée. Cette étape de la différenciation est caractérisée par le développement de sphéroïdes se détachant du tapis cellulaire, qui formeront les HIOs, après maturation (**Figure 37**). Par simple quantification des sphéroïdes formées suite aux deux traitements possibles, une augmentation significative des sphéroïdes (+100%) a été mise en évidence avec l'utilisation de CHIR99021 (**Figure 38D**). L'aspect des sphéroïdes est également différent puisqu'elles sont de taille supérieure mais également d'une forme caractéristique de « donut », s'apparentant davantage à la morphogenèse de l'intestin primitif. Néanmoins, l'enzyme utilisée pour le passage des hiPSCs ne semble pas affecter le nombre de sphéroïdes récoltées puisqu'il est similaire dans les deux cas (**Figure 38D**).

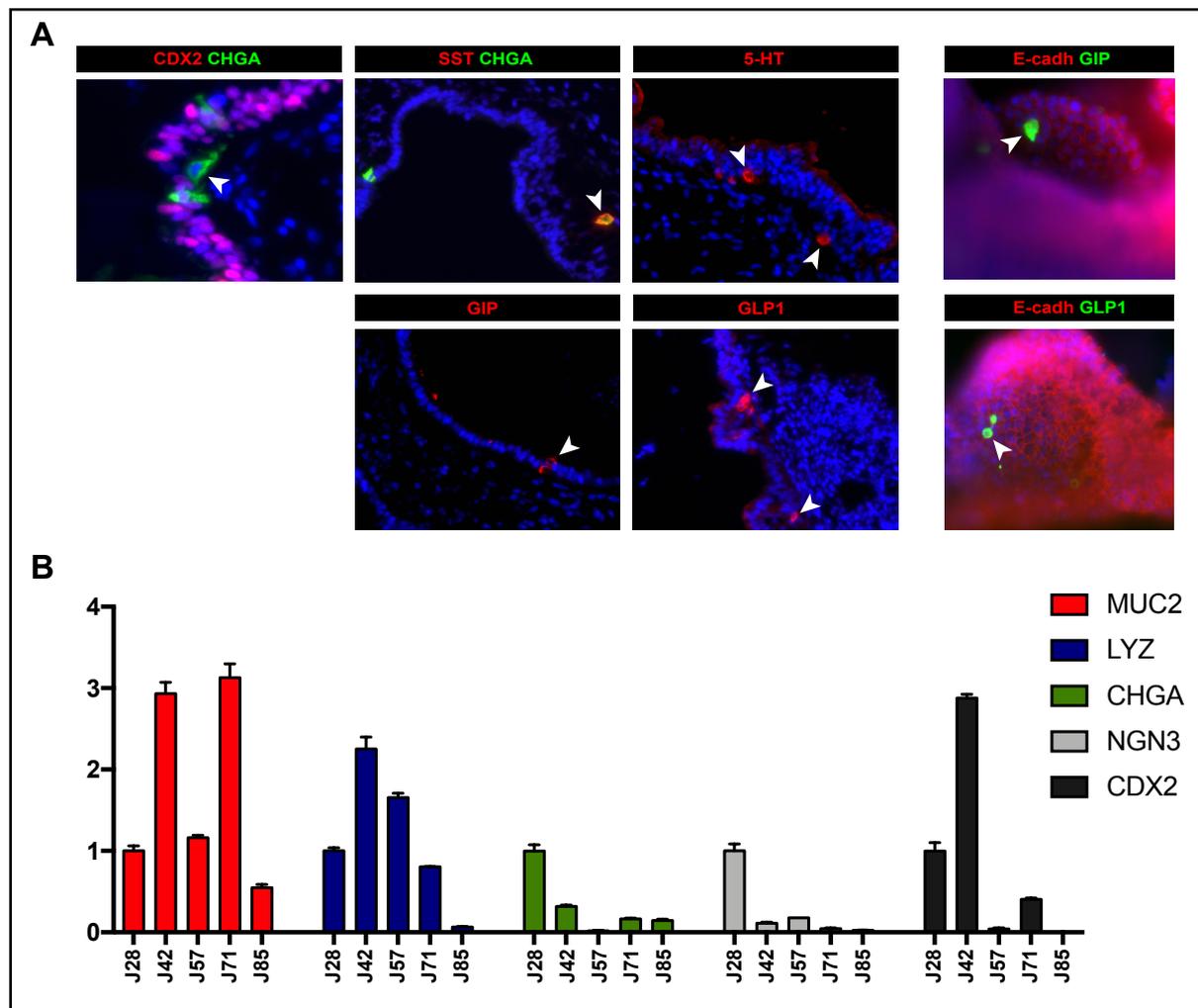
Des immunofluorescences sur les marqueurs de l'épithélium intestinal CDX2 et SOX9 ont été réalisées sur les cellules formant le tapis cellulaire, ainsi que sur les sphéroïdes formées après les 4 jours de traitement. L'analyse au FACS de ces immunomarquages nous a permis de confirmer ce qui a été observé par quantification des sphéroïdes, c'est à dire une meilleure efficacité de CHIR99021 dans la formation des sphéroïdes en comparaison à WNT3A. En effet, le marquage des sphéroïdes a révélé 58% de cellules doublement marquées CDX2<sup>+</sup> SOX9<sup>+</sup> dans celles formées avec CHIR99021, contre 42% dans celles formées avec WNT3A. Dans les deux cas, l'analyse au FACS a mis en évidence une population de cellules uniquement SOX9<sup>+</sup> (CDX2<sup>-</sup>), ce qui avait également été observé par Spence et al. (2011) par immunofluorescence. L'expression de CDX2 est retrouvée dans 90% des cellules épithéliales et l'expression de SOX9 se restreint peu à peu à une population de cellules prolifératives localisée à la base des structures assimilables aux villosités intestinales (Spence et al., 2011). Le marquage des cellules du tapis cellulaire a mis en évidence une augmentation de 18% de cellules marquées après traitement avec CHIR99021 en comparaison au traitement avec WNT3A (**Figure 38B**). Cette fois, une population de cellules CDX2<sup>+</sup> SOX9<sup>-</sup> a été mise en évidence, ce qui avait également été le cas par immunofluorescence dans la publication de Spence et al. (2011). Des cellules CDX2<sup>+</sup> sont donc retrouvées à la fois dans le tapis cellulaire et dans les sphéroïdes, alors que SOX9 n'est exprimé que dans les sphéroïdes. Cette observation s'explique par le fait que CDX2 est un marqueur d'intestin postérieur et marque également le mésenchyme entourant l'épithélium, tandis que l'expression de SOX9 se restreint aux zones prolifératives (Spence et al., 2011b).

Ces analyses attestent de la formation des sphéroïdes dans nos conditions expérimentales et indiquent une efficacité supérieure de la molécule CHIR99021 dans la postériorisation de l'endoderme. Les critères qualitatifs et quantitatifs étudiés indiquent la formation de sphéroïdes présentant des caractéristiques de l'intestin moyen. L'étape suivante a été de caractériser les HIOs obtenus à la fin du protocole de différenciation *in vitro*.

#### 1.4. Croissance et expansion des sphéroïdes en HIOs

Les sphéroïdes générées ont étéensemencées dans un environnement de culture favorisant la croissance tridimensionnelle, la différenciation et la morphogenèse intestinales (**Figure 33 et 37**). Cet environnement est composé d'un dôme de Matrigel et d'un milieu de culture contenant les trois facteurs de croissance Noggin, R-spondine1 et EGF. Le Matrigel est une matrice composée d'une préparation solubilisée extraite d'un sarcome murin, une tumeur riche en protéine de matrice extracellulaire (Hughes et al., 2010). L'agoniste de Wnt, R-spondine1; l'inhibiteur de BMP, Noggin et EGF, sont des facteurs de croissances essentiels pour la prolifération cellulaire dans les cryptes intestinales et le maintien d'une niche active de cellules souches (Sato et al., 2009). Ces conditions de culture permettent la croissance et la maturation des HIOs, caractérisée par le bourgeonnement de domaines ressemblant aux cryptes intestinales de l'épithélium (**Figure 37**). 28 jours après l'ensemencement des sphéroïdes (stade J28), les HIOs expriment des marqueurs des différents types de cellules intestinales : CDX2 et VIL1 (épithélium), CHGA (cellules entéroendocrines), MUC2 (cellules à mucus), LYZ (cellules de Paneth) et le marqueur des progéniteurs entéroendocrines NGN3 (Spence et al., 2011). Des immunomarquages sur des coupes d'organoïdes à J28 réalisés notamment par A. Beucher et A. Meunier ont montré la présence de cellules entéroendocrines (CHGA) à somatostatine (SST), GLP1, GIP et sérotonine (5-HT), parmi les cellules épithéliales (CDX2) dans des HIOs générés à partir de notre lignée hiPSCs LV37 (**Figure 39A**). Le marquage de la E-cadhérine sur des HIOs *in toto*, qui est une protéine d'adhésion cellulaire calcium-dépendante exprimée par les cellules épithéliales (Dogan et al., 1995), met en évidence la nature épithéliale du tissu formé à l'issue de la différenciation (**Figure 39A**). Des analyses par RT-qPCR ont décelé des variations de l'expression des gènes au sein des HIOs selon le stade d'analyse, dont NGN3 et CHGA qui sont plus exprimés au stade J28 qu'aux stades suivants (**Figure 39B**). Il est étonnant de constater tant de variations dans l'expression génique selon le stade des HIOs et il ne peut être exclu qu'elles soient également le reflet d'une hétérogénéité inhérente au modèle d'étude.

C'est pour ces raisons que nos études ont été menées sur le stade J28 dans un premier temps, puisque les HIOs semblent exprimer les marqueurs entéroendocrines à ce stade, dont *NGN3*, l'un de nos gènes d'intérêt. Une fois l'efficacité de la différenciation contrôlée et la production des HIOs optimisée, notre modèle d'étude de la régulation transcriptionnelle de la différenciation entéroendocrine a pu être mis en place, à commencer par la génération des clones hiPSCs mutants pour nos gènes d'intérêt.



**Figure 39. Les HIOs expriment des marqueurs d'épithélium intestinal au stade J28**

**(A)** Immunofluorescences sur des marqueurs d'épithélium intestinal CDX2, E-cadhérine, CHGA (cellules entéroendocrines) et hormones SST (somatostatine), 5-HT (sérotonine), GIP, GLP1 sur des HIOs au stade J28: panneaux de gauche, cryosections; panneaux de droite, *in toto*. **(B)** RT-qPCR sur les marqueurs CDX2, CHGA, NGN3 (marqueur pro-endocrine), MUC2 (cellules à mucus), LYZ (cellules de Paneth) sur des HIOs à différents stades : J28, J42, J57, J71, J85. Images obtenues avec un microscope à épifluorescence. Les flèches blanches indiquent des cellules marquées. Les RT-qPCR ont été réalisées en triplicats.

## Partie 2. Inactivation des gènes codant les facteurs de transcription NGN3, RFX6, ARX et PAX4

L'objectif étant d'étudier le rôle des facteurs de transcription NGN3, RFX6, ARX et PAX4 dans la différenciation entéroendocrine, il a d'abord fallu générer des lignées hiPSCs dans lesquelles les gènes codant ces facteurs de transcription ont été inactivés, avant de les différencier en tissu intestinal *in vitro*. Pour ce faire, l'outil de modification de génome CRISPR/Cas9 a été adopté. N'ayant encore jamais été utilisé au laboratoire, une mise au point de cette technologie était nécessaire pour cibler efficacement ces gènes dans les hiPSCs.

### 1. Mise au point de l'outil CRISPR/Cas9

Deux stratégies différentes ont été utilisées (**Figure 29**). La première stratégie consiste en l'utilisation d'un mutant de la Cas9 dont le domaine catalytique RuvC est muté, appelé Cas9 « nickase » (Cas9n) D10A. Seul l'autre domaine catalytique HNH reste actif, ce qui provoque une coupure sur un des brins de l'ADN au lieu d'une coupure double-brin. En effet, le domaine HNH clive le brin cible complémentaire du sgRNA et le domaine RuvC clive le brin complémentaire à la séquence cible, qui contient le motif de reconnaissance PAM. Une coupure double-brin peut cependant être mimée par l'utilisation de deux Cas9n agissant de pair et ciblant deux séquences suffisamment proches l'une de l'autre (stratégie de double-nickase). Comme les deux Cas9n doivent cliver leur brin d'ADN cible de manière efficace, cette stratégie limite les risques d'aboutir à des effets non-ciblés et est un outil généralement plus efficace par rapport à l'approche CRISPR/Cas9 «classique» qui induit une coupure double-brin (Ran et al., 2013, Mali et al., 2013). En effet, si la Cas9 sauvage reconnaît un autre locus dans le génome, la réparation de la cassure double-brin par le mécanisme NHEJ peut entraîner l'introduction d'une mutation qui peut avoir des conséquences délétères pour la cellule. En revanche, l'utilisation de la double-nickase permet d'éviter ce risque puisque si l'une des Cas9n se fixe ailleurs dans le génome, elle n'induirait qu'une coupure simple-brin qui sera réparée de manière plus fidèle par recombinaison homologue en utilisant le brin complémentaire comme matrice ou par BER (*Base Excision Repair*) (Vriend et Krawczyk, 2016 ; Dianov et Hubscher, 2013). Dans le cas de la double-nickase, l'écart maximal entre les deux sites de clivage est de 70 nucléotides, ce qui rend la délétion de larges régions des gènes cibles impossible (Ran et al., 2013). La stratégie privilégiée a donc été la délétion d'une courte séquence d'ADN incluant le codon d'initiation pour empêcher la traduction des protéines d'intérêt. La deuxième stratégie repose sur l'utilisation de la Cas9 sauvage pour induire la délétion de

domaines fonctionnels des gènes. Malgré le risque d'effets non-ciblés, cette stratégie a été utilisée dans le but de provoquer la délétion d'exons entiers et/ou de domaines fonctionnels des gènes, car cela facilite la génération et l'identification de mutations nulles. Cette stratégie sera décrite à la suite de la mise au point de la stratégie par double-*nickase*.

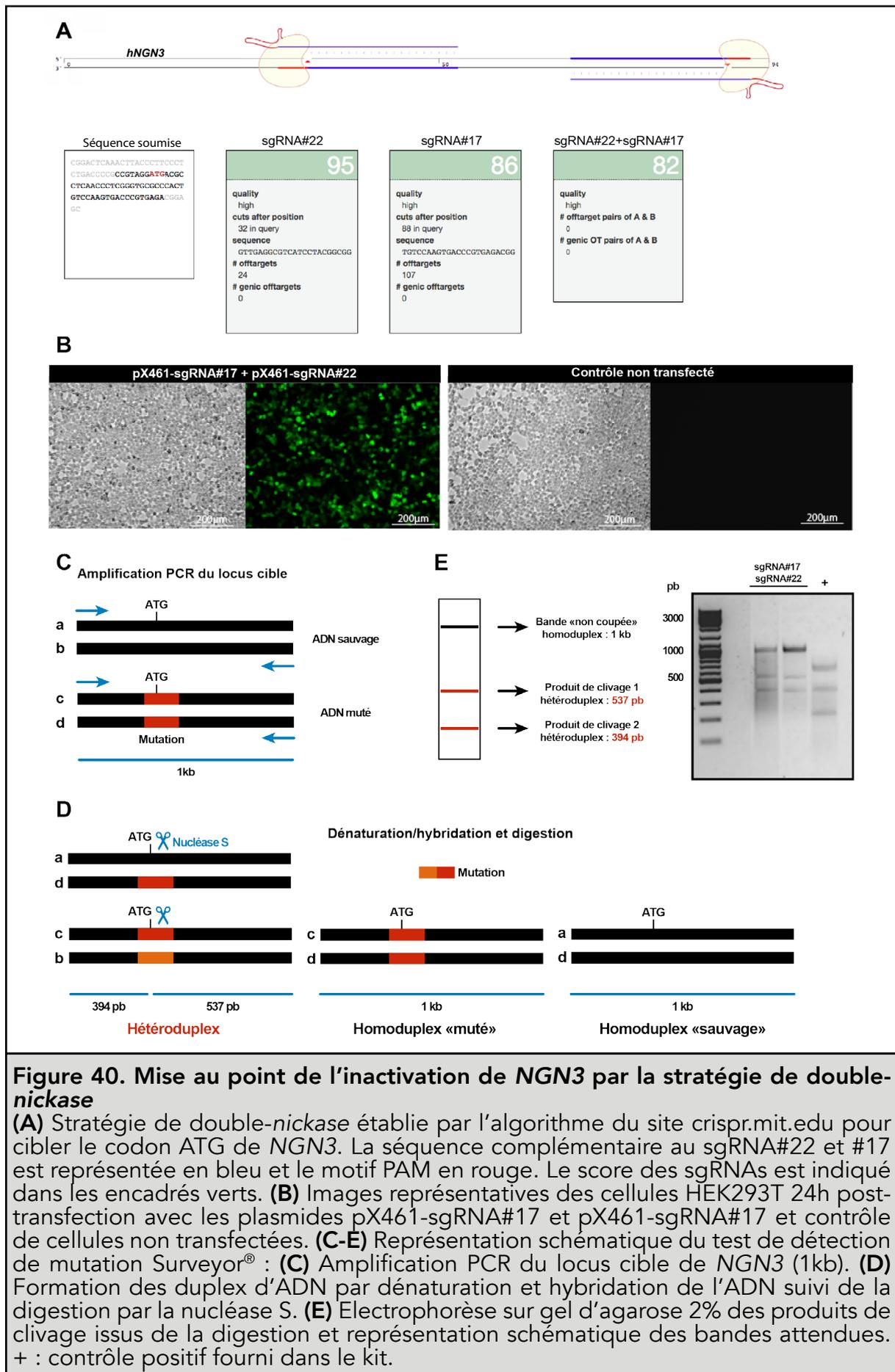
## 1.1. Approche utilisant une version mutée de la nucléase Cas9 : la double-*nickase*

### 1.1.1. Choix et validation des ARNs guides (sgRNAs)

Par souci de clarté, seule la procédure utilisée dans le cadre de l'inactivation de *NGN3* sera décrite en détail mais la même méthode a été utilisée pour cibler les autres gènes d'intérêt.

Afin d'induire la délétion du codon d'initiation de *NGN3*, deux sgRNAs ont été choisis de façon à ce qu'ils ciblent une séquence en amont et en aval du codon ATG. Les séquences de ces deux sgRNAs ont été déterminées grâce à l'interface mise à disposition par le laboratoire de Zhang (<http://crispr.mit.edu/>). En soumettant la séquence cible d'ADN et l'organisme sélectionné (*Homo Sapiens* dans notre cas), l'algorithme va déterminer la position et la séquence des guides possibles (20 nucléotides suivis d'une séquence PAM de type NGG), en les classant selon un score qui prend en compte la spécificité de reconnaissance de la séquence cible mais également les effets non ciblés possibles dans le génome de l'organisme. L'option « double-*nickase* » permet de choisir les deux guides qui, en agissant de pair, obtiennent le meilleur score basé sur les critères mentionnés précédemment. Les sgRNA#22 et sgRNA#17 ayant un score commun de 82 ont ainsi été sélectionnés (**Figure 40A**). La position exacte des potentiels sites de fixation non-spécifique est indiquée et listée pour chacun des guides potentiels. Ces deux sgRNAs ne présentent aucune cible aspécifique commune et celles propres à chacun des guides ne se trouvent pas dans des régions géniques (**Annexe 3**). Ces critères font de ces sgRNAs des candidats optimaux.

Les sgRNA#22 et sgRNA#17 ont chacun été clonés entre les sites de restriction *BbsI* du plasmide pX461 (Addgene) contenant la séquence codant la Cas9n D10A et le marqueur fluorescent eGFP (**Figure 30B**). L'efficacité de clivage au locus cible a été déterminée en réalisant un test de détection de mutation, le test Surveyor<sup>®</sup>, après transfection de cellules HEK293T avec les deux plasmides pX461-sgRNA#17 et pX461-sgRNA#22. Ces tests de fonctionnalité ont été effectués dans les cellules HEK293T, car leur culture et transfection imposent moins de contraintes techniques que celles des hiPSCs.



L'expression de la GFP 24h après la transfection témoigne d'une bonne efficacité de transfection, bien que non quantifiée, et n'est pas due à de l'autofluorescence des cellules, comme en atteste le contrôle négatif de cellules non transfectées (**Figure 40B**). Le test Surveyor<sup>®</sup> permet de détecter les mutations et polymorphismes dans une séquence d'ADN donnée. La nucléase Surveyor<sup>®</sup> (S) fait partie des endonucléases isolées à partir du céleri (nucléase CEL) et permet la détection des mésappariements de base et des petites insertions ou délétions (Vouillot et al., 2015). Pour cela, une séquence d'environ 1kb de *NGN3* contenant le codon d'initiation et les séquences complémentaires aux sgRNAs a été amplifiée par PCR, après avoir extrait l'ADN génomique des cellules HEK293T transfectées (**Figure 40C**). S'ensuit une étape de dénaturation et d'hybridation des brins d'ADN qui va provoquer la formation de deux sortes de duplex d'ADN : des homoduplex formés à partir de deux séquences sauvages ou mutées, ou des hétéroduplex formés d'une séquence sauvage et d'une séquence mutée ou de deux séquences possédant des mutations différentes. La nucléase S va alors reconnaître les mésappariements entre les deux séquences constituant les hétéroduplex et induire un clivage de l'extrémité 3' sur les deux brins (**Figure 40D**). Une fois les duplex digérés par la nucléase S, les produits de clivage sont analysés par électrophorèse sur gel d'une concentration d'agarose dépendante de la taille des fragments à observer. La présence de séquences sauvages est traduite par une bande de 1kb et les produits de clivage attendus, si les deux Cas9n coupent bien 3 nucléotides en aval du motif PAM, sont détectés par une bande de 537 pb et une autre de 394 pb, représentatives d'une délétion de 50 nucléotides (**Figure 40D-E**). D'après l'analyse sur gel d'agarose 2%, il y a un mélange de séquences sauvages (1kb) et de séquences mutées de la taille attendue. Le contrôle positif fourni dans le kit met en évidence deux produits de clivage de 217 pb et de 416 pb, ainsi qu'une séquence sauvage de 632 pb, témoignant de l'efficacité de la digestion de la nucléase S. Ainsi, les bandes de clivage observées dans nos échantillons sont bien dues aux mutations et ne sont pas des bandes aspécifiques (**Figure 40E**). Dans le cas du contrôle, les bandes correspondant aux produits de clivage sont de même intensité que celle correspondant à la séquence sauvage, tandis que dans notre cas, les bandes des produits de clivage sont de plus faible intensité que celle de la séquence sauvage. Ceci s'explique par le fait qu'il y ait plus de séquences sauvages que de séquences mutées, dû à l'absence de sélection des cellules transfectées et une efficacité d'induction de mutation limitée. La paire de guides sgRNA#17 et sgRNA#22 est donc validée pour induire la délétion d'une séquence de 50nt contenant le codon d'initiation de *NGN3*. Toutefois, cette approche, en fonction de la position des sgRNAs, présente le risque de générer des allèles hypomorphes et rend l'identification des clones mutants par génotypage difficile, puisque les petites mutations induites ne sont pas nécessairement visibles par amplification PCR.

Ainsi, l'utilisation de la Cas9 sauvage et de deux sgRNAs (coupures double-brin) a l'avantage de faciliter à la fois la génération d'allèles nuls et l'identification des clones mutants par génotypage, puisqu'elle permet la délétion de plus grandes régions d'ADN. Malgré les risques des effets non-ciblés, cette approche a ensuite été privilégiée pour cibler les autres gènes d'intérêt.

## 1.2. Approche utilisant la Cas9 sauvage

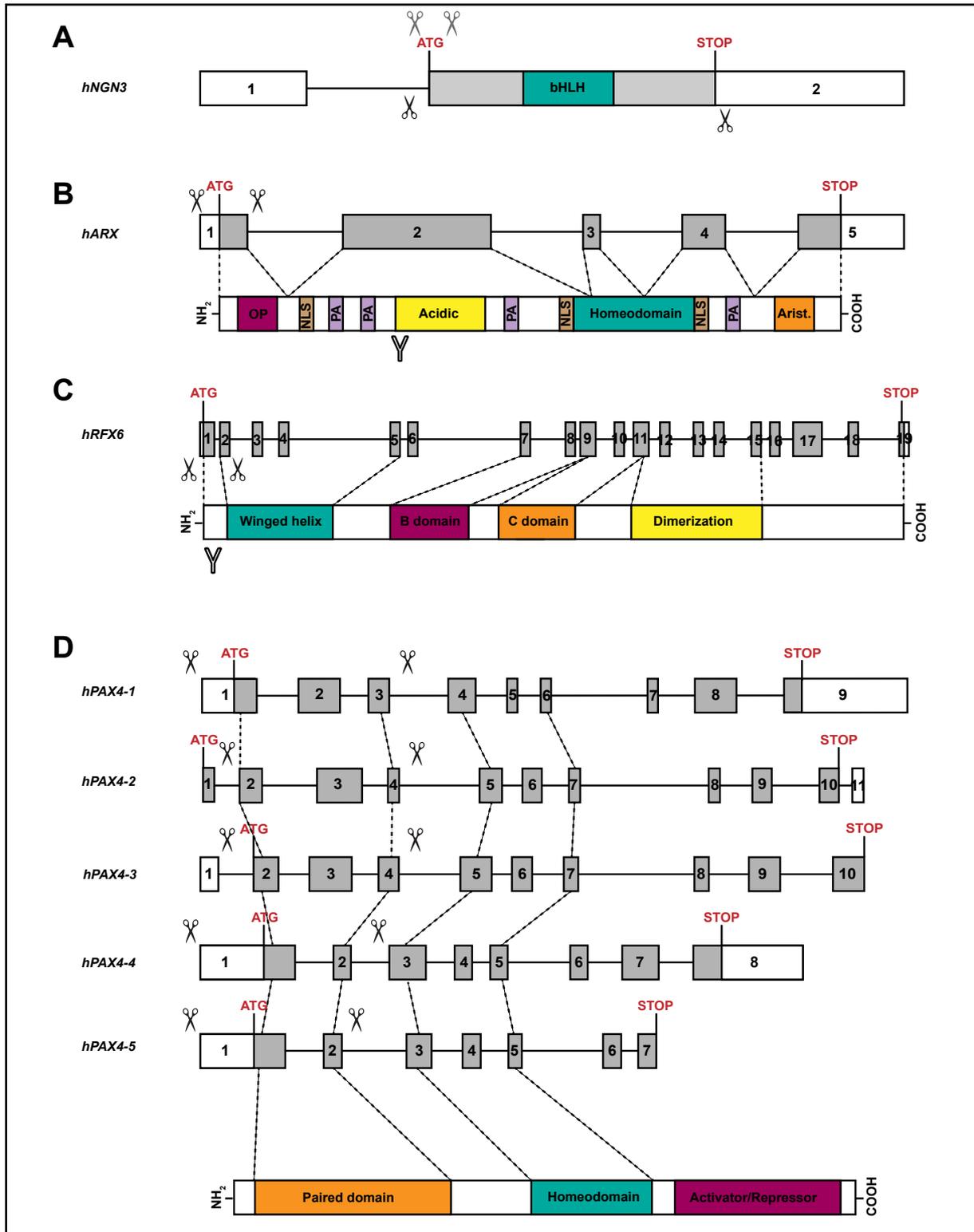
Ayant rencontré des difficultés de génotypage des clones avec l'approche susmentionnée (voir **2.1.1** de cette partie), l'approche avec la Cas9 sauvage a été mise en place pour pallier à ces problèmes. Les petites mutations de type insertions ou délétions générées étaient de trop petites tailles pour être visibles par simple amplification PCR et nécessitaient alors un séquençage systématique, tandis que la deuxième approche rend possible la délétion de larges domaines (plusieurs exons) qui facilite le génotypage par la suite. Toutefois, il faudra garder en tête l'augmentation du risque des effets non ciblés avec cette stratégie. Le choix des sgRNAs ainsi que leur clonage dans les plasmides exprimant la Cas9 et des marqueurs de sélection sont issus d'une collaboration avec Bernardo Reina San Martin à l'IGBMC. Les informations relatives à cette approche sont disponibles dans le matériel et méthodes (**Figure 30A et 31**) et la position des séquences cibles des sgRNAs sur les gènes est représentée sur la **figure 41**. Les stratégies utilisées pour chaque gène seront détaillées dans des parties dédiées. Par souci de temps, l'efficacité de clivage n'a cette fois pas été testée et les hiPSCs ont directement été transfectées avec les plasmides (**Figure 30A**).

## 2. Inactivation des facteurs de transcription dans les hiPSCs

### 2.1. Inactivation de *NGN3*

#### 2.1.1. Utilisation de la stratégie de double-*nickase*

Comme mentionné en **1.1.**, la première approche de double-*nickase* a d'abord été utilisée afin d'induire la délétion du codon d'initiation de la traduction de *NGN3* et d'empêcher la production de la protéine, ou encore d'induire une mutation de type insertion ou délétion qui décalerait le cadre de lecture, menant à la production d'une protéine *NGN3* tronquée non fonctionnelle. Comme décrit dans cette même partie, deux sgRNAs ont été choisis, l'un reconnaissant 20 nucléotides de la séquence du gène comprenant le codon ATG et l'autre ciblant 20 nucléotides se trouvant en aval du codon ATG (**Figure 41A**). Théoriquement, les Cas9n induisent une coupure simple-brin 3 nucléotides en aval ou en amont (selon le brin d'ADN) de la séquence PAM qu'elles reconnaissent après appariements de bases complémentaires des sgRNAs aux séquences cibles.



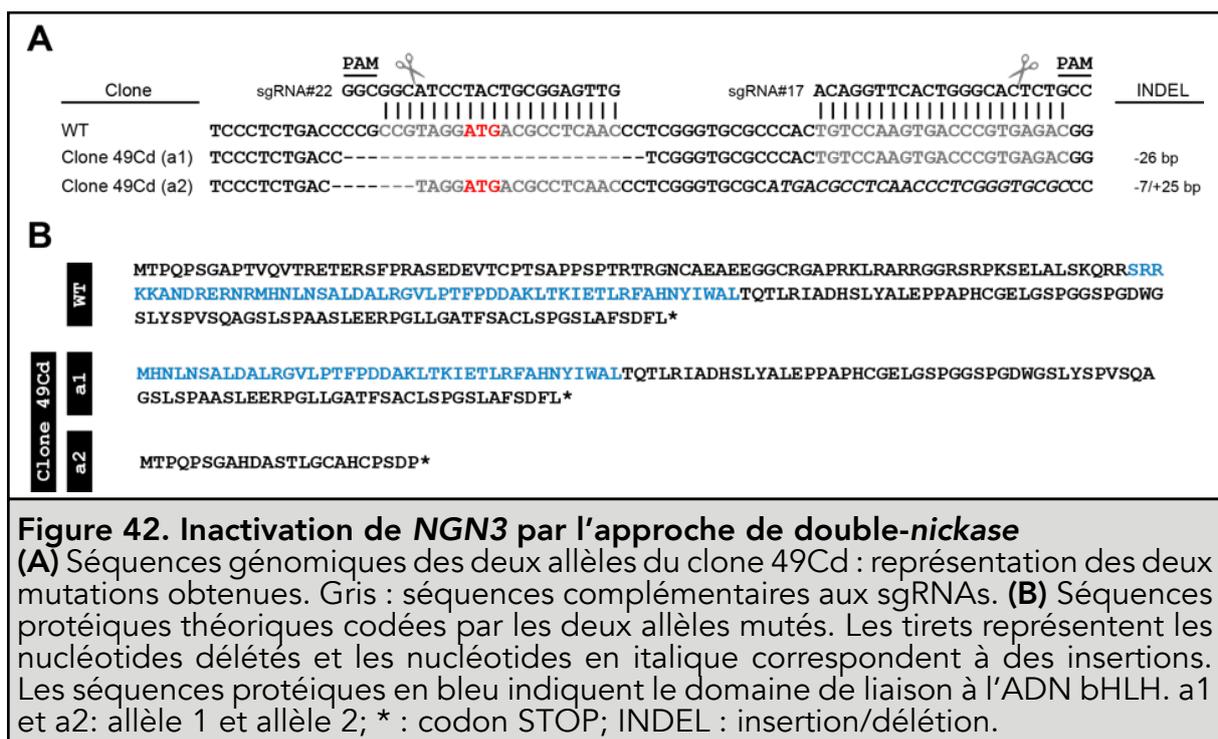
**Figure 41. Représentation schématique de la structure des gènes cibles et de leur protéine associée**

La portion des gènes en gris représente la séquence codante des gènes. Les ciseaux noirs indiquent la position des coupures théoriques dans l'approche CRISPR/Cas9 utilisant les propriétés de la Cas9 sauvage (coupure double-brin) et les ciseaux gris dans l'approche de double-nickase. Les sites de fixation des anticorps sur les protéines sont également représentés lorsqu'ils sont connus. **(A)** Structure de *NGN3* **(B)** Structure de *ARX* **(C)** Structure de *RFX6* **(D)** Structure des différents variants de *PAX4* (*hPAX4-1* à *5*).

Une délétion de 50 nucléotides comprenant le codon d'initiation est ainsi attendue. Les deux coupures simple-brin provoquées par l'action des Cas9n via les sgRNAs sont suffisamment proches pour être réparées de la même manière qu'une unique coupure double-brin. Cette réparation par le mécanisme NHEJ peut induire l'intégration de mutations de type insertions ou délétions (*indels*). Dans tous les cas, l'objectif est l'obtention de mutations menant à un décalage du cadre de lecture ou à un codon STOP prématuré, pour provoquer la délétion de *NGN3*. L'efficacité des sgRNAs ayant été validée auparavant sur des cellules HEK293T, les hiPSCs ont pu être transfectées avec les plasmides pX461-sgRNA#22 et pX461-sgRNA#17. Les plasmides portent le marqueur de sélection fluorescent eGFP et c'est par ce biais que les cellules ayant reçu au moins un des plasmides ont pu être sélectionnées par tri au FACS. Afin d'éviter l'obtention de clones hétérogènes, c'est-à-dire provenant de cellules différentes et donc potentiellement de génotypes différents, les cellules GFP-positives ont d'abord été réparties dans des plaques 96 puits avec une cellule par puits. Malheureusement, les cellules ne survivaient pas de cette manière et aucun clone n'a pu être formé. Effectivement, les hiPSCs sont maintenues en culture sous forme de colonies et les cellules pluripotentes de manière générale, bien que dépendant de la lignée cellulaire et des conditions de culture, ne survivent pas ou commencent à se différencier spontanément lorsqu'elles sont cultivées en cellules uniques (Chen et Pruett-Miller, 2018). Pour ces raisons, les cellules ont été réparties dans des boîtes de Pétri P60 et P35 à des concentrations différentes après tri au FACS, dans l'espoir d'obtenir des clones dérivant d'une seule cellule. Après une semaine à dix jours de culture, les clones isolés ont été dissociés mécaniquement pour en réensemencer la moitié dans une nouvelle boîte de culture pour le maintien et en extraire l'ADN génomique avec l'autre moitié, afin de procéder au génotypage.

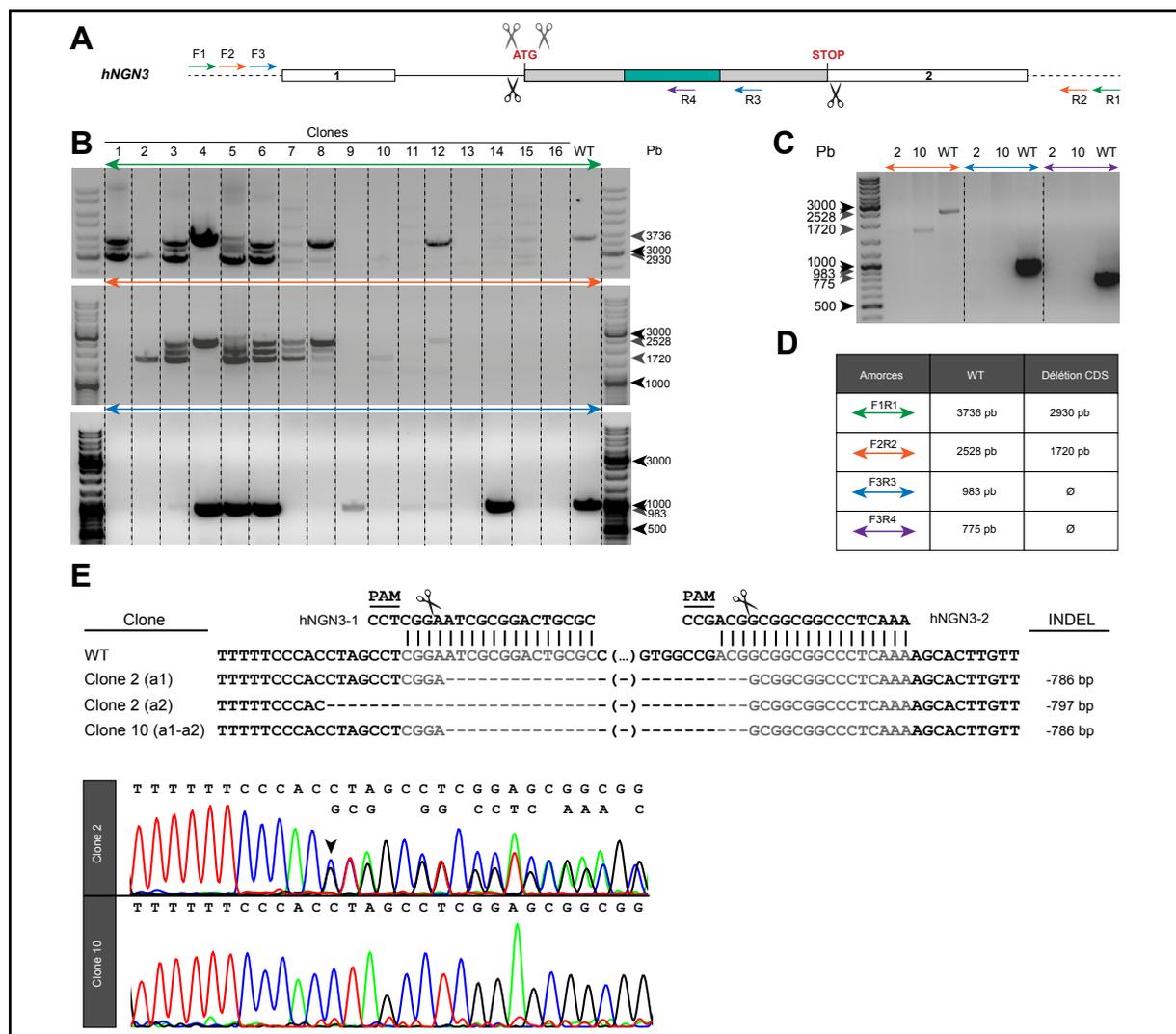
Une séquence d'environ 1 kb a été amplifiée par PCR à l'aide des amorces F3/R3 (**Figures 32 et 43A**). Les mutations induites par cette approche de double-*nickase* peuvent être de nature et de tailles variables, rendant l'analyse sur gel d'agarose des produits PCR peu informative. C'est pourquoi les produits PCR de 40 clones obtenus ont été séquencés avec les amorces F3/R3 menant à l'identification de 30 clones sauvages (WT) et 10 clones hétérogènes pour lesquels au minimum 3 séquences étaient visibles par séquençage, ce qui laisse présager d'un mélange de cellules de différents génotypes dans un même clone. Néanmoins, des mutations ont pu être décelées dans ces 10 clones, notamment dans le clone 49 qui présentait une séquence avec une délétion de 26 nucléotides, une séquence avec une insertion de 16 nucléotides et une séquence sauvage. De ce fait, ce clone a été sous-cloné en le dissociant sous forme de cellules uniques avec l'enzyme accutase, puis en ensemençant les cellules de manière à obtenir des clones isolés. Le séquençage nous a permis d'isoler deux clones C et D dont le séquençage a mis en évidence la présence des deux séquences

mutées, mais également la séquence sauvage en plus faible quantité, laissant penser à une contamination de cellules sauvages parmi le clone mutant. Afin d'éliminer cette contamination, ces clones ont à nouveau été sous-clonés. Cette fois, le clone 49Cd possédant la délétion de 26 nucléotides sur un allèle et l'insertion de 16 nucléotides sur l'autre, a pu être isolé. En analysant le séquençage avec plus de précision, l'insertion de 16 nucléotides initialement observée résulte en fait d'une délétion de 7 nucléotides et d'une insertion de 25 nucléotides (**Figure 42A**). La délétion présente sur l'allèle 1 provoque la suppression du codon d'initiation et la mutation sur l'allèle 2 induit un codon STOP prématuré altérant la traduction de la protéine. Cependant, un autre codon ATG en phase est présent dans l'exon 2 et peut éventuellement mener à la production d'une protéine incomplète à partir de l'allèle 1, dont le domaine de liaison à l'ADN bHLH n'est que partiellement perdu (**Figure 42B**). Ce clone 49Cd a tout de même été retenu pour les analyses décrites par la suite mais il faut tenir compte du fait que la fonction de la protéine peut n'être que partiellement affectée. L'isolation de clones purs a été particulièrement laborieuse et il ne peut être exclu que les nombreuses manipulations puissent avoir un impact sur le transcriptome des hiPSCs et notamment le maintien de leur pluripotence. L'identification des clones mutants a été compliquée par les mutations de petites tailles introduites, qui ne sont pas forcément identifiables par amplification PCR. Ainsi, la deuxième approche CRISPR/Cas9, visant à induire la délétion de larges portions de gènes, a été utilisée pour cibler les autres gènes d'intérêt et de nouveaux clones NGN3<sup>-/-</sup> ont également pu être générés de cette manière.



### 2.1.2. Seconde approche CRISPR/Cas9 : utilisation de la Cas9 sauvage

*NGN3* est composé de deux exons, la séquence codante entière est située sur l'exon 2 et inclut donc la séquence codant le domaine de liaison à l'ADN bHLH, qui est essentiel à la fonction de ce facteur de transcription (**Figures 41A et 43A**). Cette seconde stratégie nous a permis de générer des clones dont la séquence codante entière de *NGN3* a été délétée. Pour ce faire, les hiPSCs ont été transfectées avec le plasmide pX332 et sélectionnées par tri au FACS grâce à la GFP codée par le plasmide (**Figure 30A**).



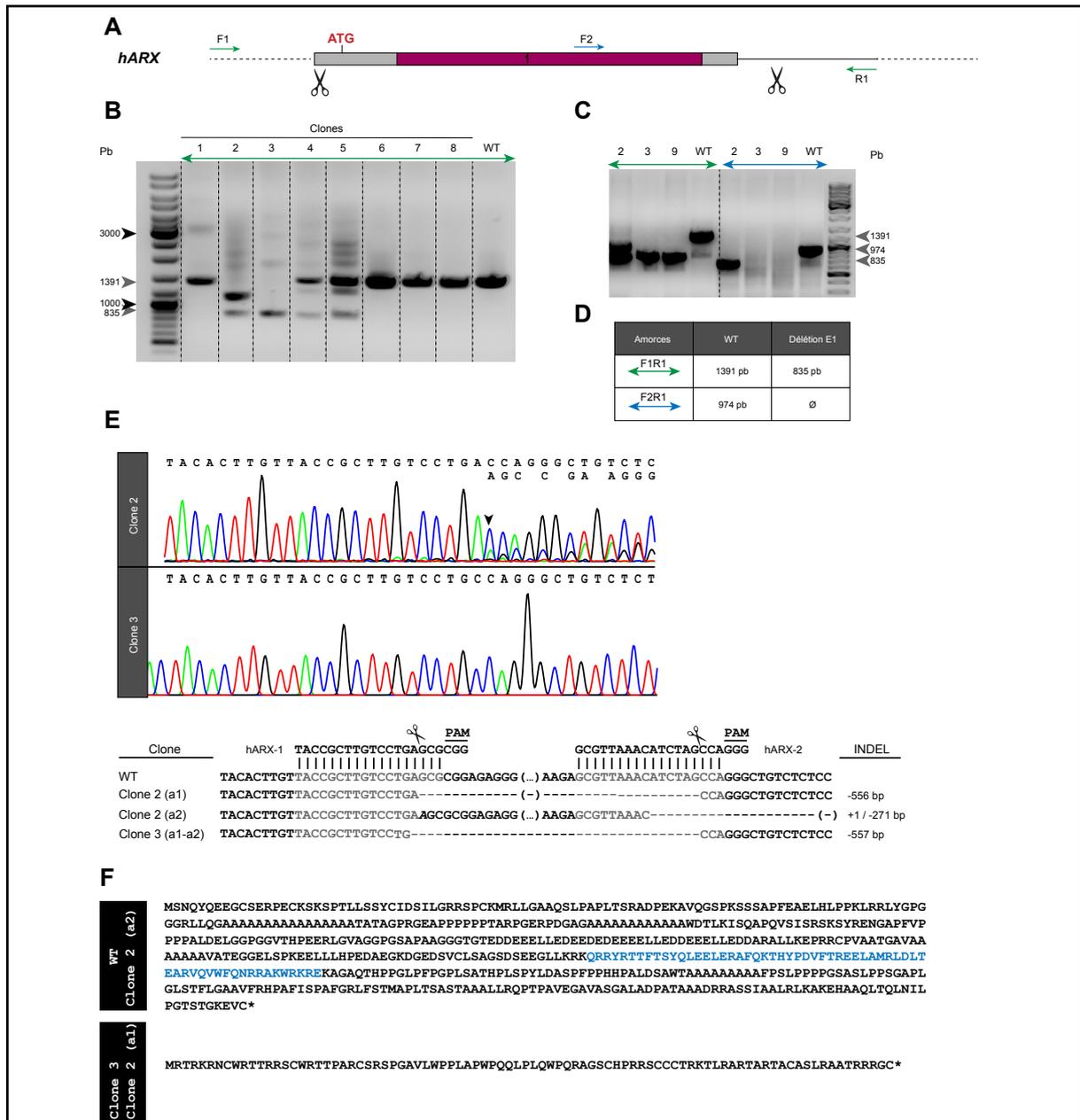
**Figure 43. Génération de deux mutants *NGN3*<sup>-/-</sup> par délétion de la séquence codante**

(A) Position des amorces utilisées pour le génotypage et des coupures théoriques induites par la Cas9. Ciseaux noirs: Coupures double-brin, ciseaux gris: double-nickase, rectangle bleu: séquence codante. (B) Génotypage par amplification PCR de 16 clones obtenus. (C) Confirmation du génotype des clones 2 et 10 par amplification PCR. (D) Paires d'amorces utilisées pour les génotypages et tailles des bandes attendues. (E) Séquences génomiques des clones 2 et 10 : représentation des trois mutations obtenues et séquençage des clones 2 et 10. a1 et a2: allèle 1 et allèle 2; INDEL: insertion/délétion.

16 clones ont été génotypés par amplification PCR de leur ADN en utilisant plusieurs combinaisons d'amorces externes et internes à la délétion attendue, car les séquences mutées sont préférentiellement amplifiées par rapport aux plus grandes séquences sauvages, ce qui aurait biaiser l'interprétation (**Figure 43A-D**). C'est ainsi que deux clones ont été mis en évidence grâce aux informations apportées par les trois amplifications PCR différentes : le clone 2 et le clone 10, qui ne présentaient que les bandes correspondant à la délétion attendue. Une seconde amplification PCR nous a permis de confirmer leur génotype et les produits PCR ont été séquencés pour plus de précision quant aux séquences d'ADN de ces clones. Le séquençage nous a ainsi permis de déterminer les séquences précises de leur mutation, à savoir une délétion de 786 pb sur les deux allèles pour le clone 10 et une délétion de 797 pb sur un allèle (a2) et une délétion de 786 pb sur l'autre (a1) pour le clone 2 (**Figure 43E**). La séquence codante est donc bien entièrement délétée et la traduction de la protéine est rendue impossible de par son absence. De ce fait, ces deux clones constituent un modèle robuste pour étudier l'inactivation de *NGN3* après différenciation de ces derniers en tissu intestinal *in vitro*.

## 2.2. Inactivation de ARX

ARX est un gène plus complexe composé de 5 exons. Afin de l'inactiver, l'exon 1 porteur du codon d'initiation et codant le domaine fonctionnel octapeptide (OP) a été ciblé (**Figure 41B**), notamment car des mutations humaines ont été retrouvées dans l'ADN codant ce domaine, dont c.196+2T>C qui provoque un saut d'exon de l'exon 1 (Kato, et al. 2004). L'ARNm de ARX dépourvu d'exon 1 provoquera un déphasage du cadre de lecture et sera probablement équivalente à une mutation nulle. Deux sgRNAs ciblant respectivement une séquence en amont et une autre en aval du codon ATG ont été choisis (**Figures 41B et 44A**). Les hiPSCs ont été transfectées avec le plasmide pX333 contenant la séquence codant la Cas9 et les sgRNAs (**Figures 30A et 31**), puis les cellules GFP-positives sélectionnées au FACS, pour ensuite être réensemencées. 8 clones isolés ont pu être analysés par amplification PCR du locus cible avec les amorces F1 et R1 permettant d'amplifier une séquence sauvage (WT) de 1391 pb et une séquence de 835 pb si la mutation attendue est présente (**Figure 44B et D**). Ceci nous a permis d'identifier le clone 3 qui semble posséder la mutation attendue sur les deux allèles et le clone 2 qui lui, semble posséder la mutation attendue sur un allèle mais également une mutation de plus petite taille sur l'autre allèle (**Figure 44B**). Les clones 1, 6, 7 et 8 sont similaires au contrôle sauvage et sont sans doute des clones sauvages. Les clones 4 et 5 ont la bande correspondant à la séquence mutée attendue mais également la bande sauvage d'intensité plus forte, suggérant des clones sauvages contaminés par des mutants, plutôt que des clones hétérozygotes.



**Figure 44. Génération d'un mutant  $ARX^{-/-}$  et d'un mutant  $ARX^{+/-}$  par délétion de l'exon 1**

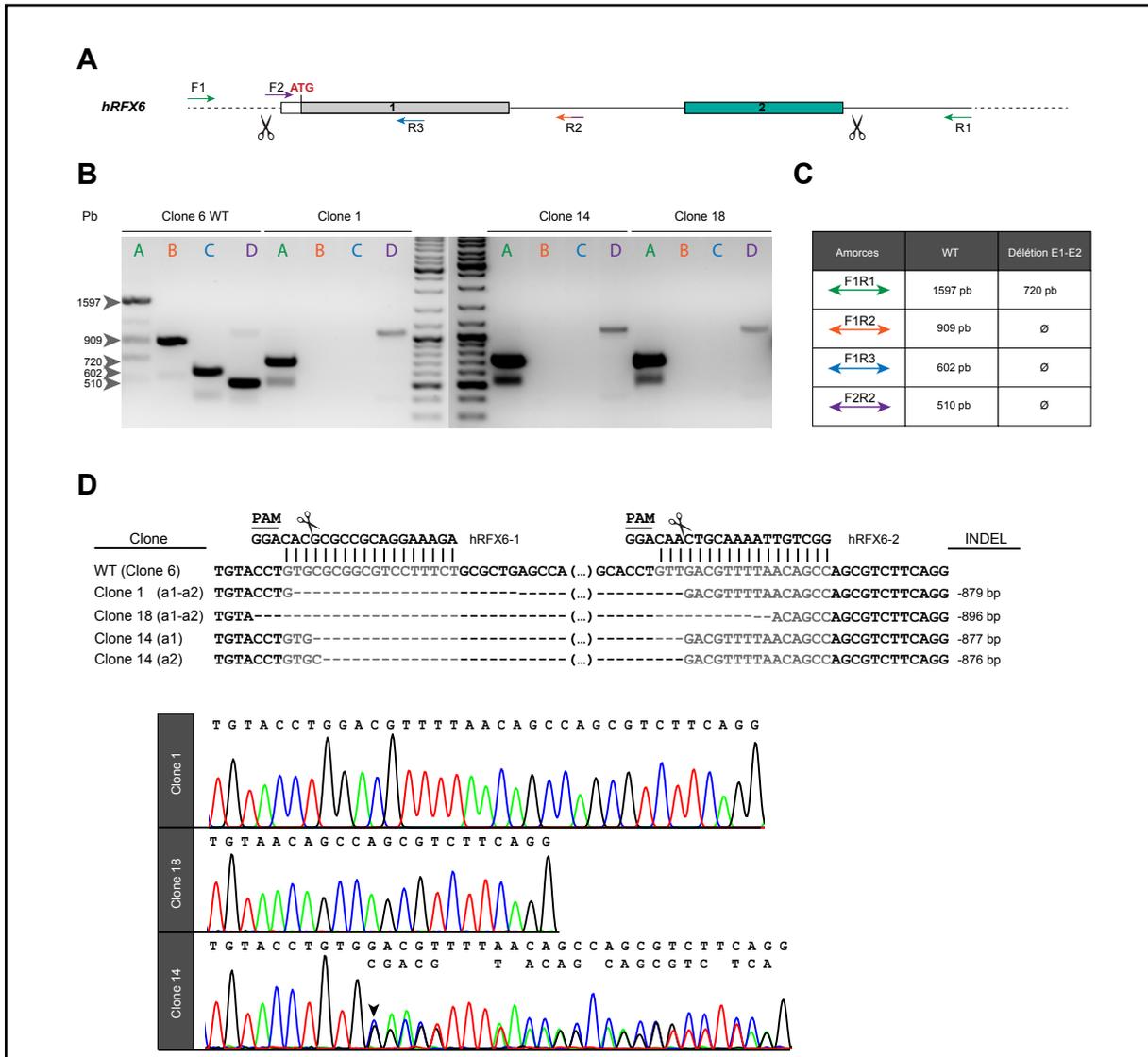
(A) Position des amorces utilisées pour le génotypage et des coupures théoriques induites par la Cas9. Rectangle rose: domaine octapeptide (OP). (B) Génotypage par PCR externe de 8 clones. (C) Confirmation du génotype des clones 2 et 3, ainsi que du clone 9, par PCR externe et interne à la délétion théorique. (D) Paires d'amorces utilisées pour les génotypages et tailles des bandes attendues. (E) Séquençage des produits PCR des clones 2 et 3 issus de la figure C et représentation des mutations obtenues. (F) Séquences protéiques théoriques des clones 2 et 3. a1 et a2: allèle 1 et allèle 2; \*: codon STOP; INDEL: insertion/délétion. Gris: séquences complémentaires aux sgRNAs.

Afin d'éviter le problème d'amplification préférentielle des séquences mutées, une autre combinaison d'amorces, dont une des amorces se fixe sur la séquence théoriquement délétée, a été utilisée. Ainsi, la séquence sauvage se traduit par une bande de 974 pb et aucune bande n'est censée être visible si la délétion a bien eu lieu. Le génotype du clone 3 est donc confirmé puisqu'aucune séquence n'est amplifiée, de même que le génotype du clone 2 puisqu'une séquence de plus petite taille que celle du contrôle sauvage est amplifiée, correspondant à la plus petite délétion sur un des deux allèles. Le clone 9, qui est un clone obtenu dans le cadre d'un essai de double inactivation *ARX-PAX4* décrit plus tard (**voir 2.4.**), semble également être un mutant homozygote puisque la séquence sauvage est absente par PCR (**Figure 44C**). Le séquençage de ces produits PCR a permis de confirmer le génotype des clones, à savoir une délétion de 557 pb sur les deux allèles du clone 3 et une délétion de 556 pb sur l'allèle 1 du clone 2 et une délétion de 271 pb, ainsi qu'une insertion de 1 pb sur l'allèle 2 (**Figure 44E**). Le clone 2 est vraisemblablement hétérozygote *ARX*<sup>+/-</sup>, car la mutation de l'allèle 2 est dans une région intronique, n'ayant théoriquement aucun impact sur la traduction de la protéine. Quant aux mutations attendues, elles induisent un codon STOP prématuré qui mènera à une protéine très petite et non fonctionnelle (**Figure 44F**). Le génotype du clone 3 est donc validé et il peut ainsi être utilisé pour étudier l'effet de l'inactivation totale de *ARX* dans la différenciation entéroendocrine, tandis que le clone 2 pourra sans doute être utilisé pour étudier l'inactivation partielle du facteur de transcription.

### 2.3. Inactivation de *RFX6*

*RFX6* est composé de 19 exons et possède un domaine de liaison à l'ADN de type *winged helix* codé par les exons 2 à 5. L'objectif a été la délétion de l'exon 1 et 2 et donc du codon d'initiation et une partie de domaine *winged helix*, afin d'empêcher l'initiation de la traduction (**Figures 41C et 45A**). Les hiPSCs ont été transfectées avec le plasmide pX335 contenant les deux sgRNAs et la Cas9 (**Figures 30A et 31**), puis les cellules GFP-positives sélectionnées, pour être réensemencées. 20 clones ont été analysés au total à l'aide d'une amplification PCR du locus cible avec les amorces F1 et R1 et seul trois d'entre eux semblaient être des mutants (non montrés). Puisqu'avec des amorces externes à la délétion, la séquence mutée est préférentiellement amplifiée, quatre paires d'amorces ont été utilisées pour confirmer le génotype de ces clones: une externe et trois internes à la délétion (**Figure 45B-C**). L'analyse des produits PCR sur gel d'agarose a mis en évidence un génotype similaire pour les trois clones 1, 14 et 18.

En effet, la PCR avec les amorces F1R1 a amplifié la séquence mutée attendue de 720 pb et les PCR avec les amorces se fixant à l'intérieur de la délétion théorique n'amplifient aucune séquence et ce pour les trois clones, en comparaison avec le contrôle sauvage, où les bandes correspondant à la séquence sauvage sont bien amplifiées (**Figure 45B**). Ces différentes PCR ont donc permis d'identifier 3 clones mutants pour lesquels les produits PCR ont été séquencés, afin de déterminer la séquence exacte des mutations.

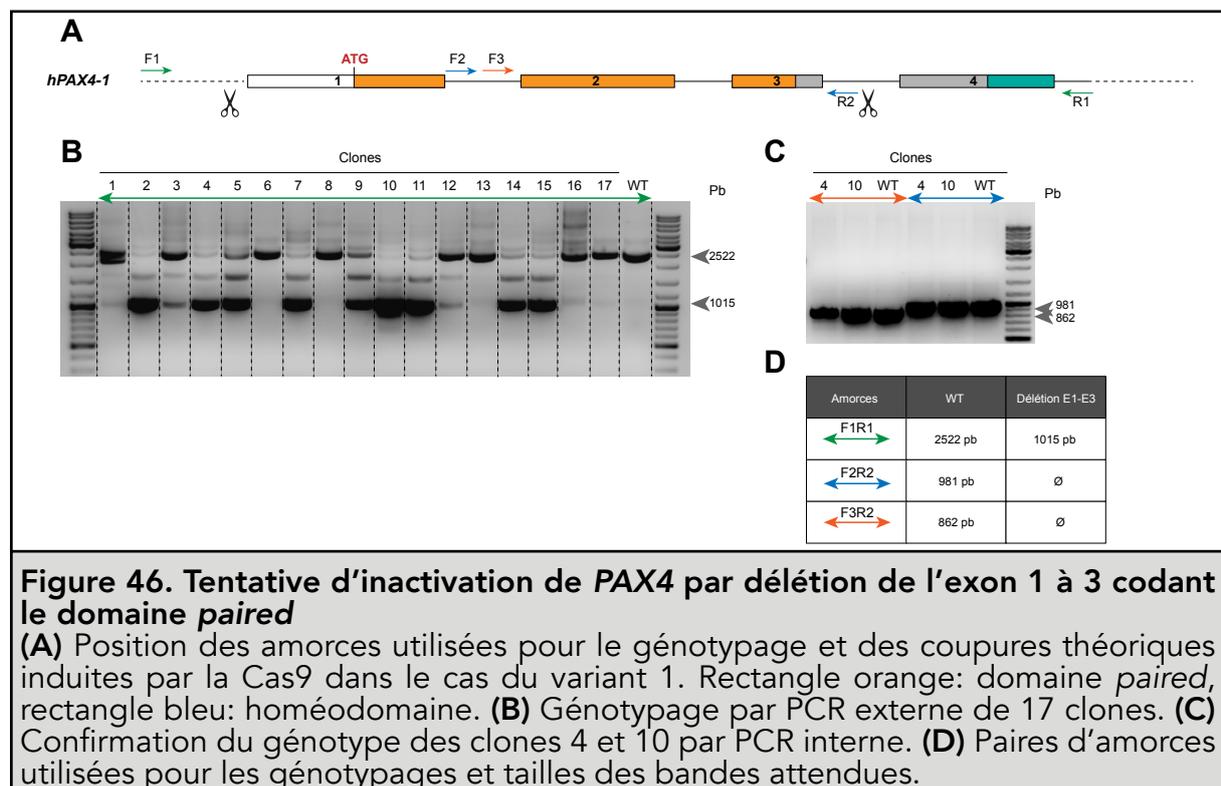


**Figure 45. Génération de trois mutants RFX6<sup>-/-</sup> par délétion des exons 1 et 2**  
**(A)** Position des amorces utilisées pour le génotypage par PCR et des coupures théoriques induites par la Cas9. Rectangle bleu: domaine *winged helix*. **(B)** Confirmation du génotype des clones 1, 14 et 18 par PCR externes et internes à la délétion théorique. **(C)** Paires d'amorces utilisées pour les génotypages et tailles des bandes attendues. **(D)** Représentation des séquences génomiques des clones 1, 14 et 18 : représentation des 4 mutations obtenues et séquençages des trois clones. a1 et a2: allèle 1 et allèle 2; INDEL: insertion/délétion. Gris: séquences complémentaires aux sgRNAs.

Le clone 1 possède la même délétion de 879 pb sur les deux allèles, le clone 18 possède une délétion de 891 pb également sur les deux allèles. Une délétion de 876 pb est retrouvée sur un allèle du clone 18 (a2) et une délétion de 877 pb sur l'autre (a1) (**Figure 45D**). Bien que le génotypage des hiPSCs est compliqué par l'amplification préférentielle des plus petites séquences, rendant l'identification des séquences sauvages difficile, la mise en commun des différentes amplifications PCR externes et internes aux délétions attendues avec les séquençages de ces produits PCR permet toutefois d'identifier de manière fiable les clones mutants. Les clones 1, 14 et 18 ont tous les trois été utilisés dans les analyses suivantes.

## 2.4. Le cas de PAX4

Le gène codant le facteur de transcription PAX4 est composé de deux domaines de liaison à l'ADN: un domaine *paired*, un domaine de type « homeoboîte » et un domaine activateur/répresseur de la transcription. Le domaine fonctionnel *paired* a été ciblé pour induire l'inactivation du gène, afin de reproduire la méthode employée pour générer les souris *Pax4<sup>-/-</sup>* (Sosa-Pineda et al., 1997). Cependant, il existe différents isoformes du gène (*PAX4-1* à *PAX4-5*) avec un nombre d'exons variable et le domaine *paired* n'est pas codé par les mêmes exons selon les variants (**Figure 41D**). Dans tous les cas, les sgRNAs ciblent des séquences de part et d'autre de la séquence codant ce domaine et induisent une délétion théorique d'environ 1,5 kb (**Figure 46A**). En revanche, dans le cas du variant *PAX4-2*, le codon d'initiation n'est pas inclu dans la délétion générée par l'action de la Cas9 et des sgRNAs (**Figure 41D**).

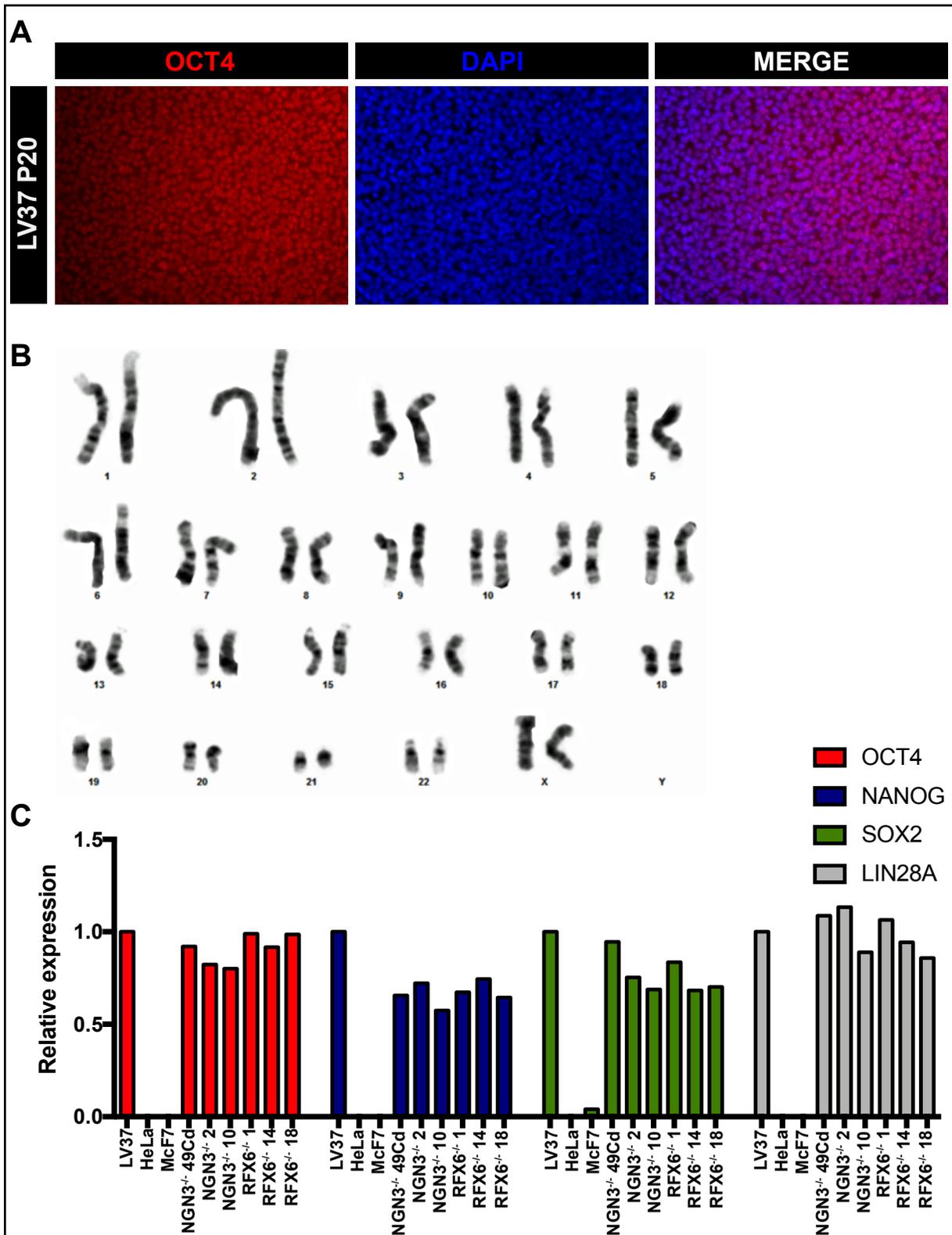


*PAX4* représente un cas particulier, car malgré l'analyse de nombreux clones suite à plusieurs transfections des hiPSCs, aucun clone mutant n'a pu être mis en évidence. En effet, le génotypage par amplification PCR du locus cible a révélé uniquement des mutants hétérozygotes ou hétérogènes, c'est-à-dire que la séquence mutée était à chaque fois accompagnée d'une séquence sauvage et ce, malgré les sous-clonages répétés. Par souci de clarté, seule l'analyse représentative de 17 clones est présentée dans ce manuscrit (**Figure 46B**). L'amplification PCR avec la paire d'amorce F1R1 a mis en évidence 11 clones possédant la mutation attendue : les clones 2, 3, 4, 5, 7, 9, 10, 11, 12, 14 et 15. En effet, la bande d'une taille de 1015 pb témoigne de la présence de la mutation dans ces clones. Néanmoins, la séquence mutée est accompagnée de la séquence sauvage représentée par une bande de 2522 pb dans la majorité des cas, sauf pour les clones 2, 4, 7, 10, 11, 14 et 15 où la bande sauvage n'est que présente en faible intensité, voire absente (**Figure 46B-D**). Pourtant, l'amplification PCR avec des amorces internes à la délétion révèle la présence de la séquence sauvage, notamment dans les clones 4 et 10 (**Figure 46C**). Le sous-clonage de ces deux clones n'a pas permis d'isoler de mutants (non montré).

De plus, une expérience de double inactivation à la fois de *ARX* et *PAX4* a également été réalisée, mais encore une fois, aucun mutant *PAX4*<sup>-/-</sup> n'a pu être généré, alors que ça a été le cas pour *ARX* (clone 9, **Figure 44C**). L'approche utilisée visant à induire la délétion d'exons entiers semble pourtant robuste puisque pour une vingtaine de clones analysés pour les autres gènes, il y avait toujours au moins un mutant. Dans le cas de *PAX4*, plus d'une centaine de clones ont été analysés et les analyses n'ont révélé aucun mutant, laissant penser que l'absence de *PAX4* pourrait être létale pour les hiPSCs, bien que cela n'ait jamais été démontré. La présence de différents isoformes a rendu l'établissement de la stratégie d'inactivation de *PAX4* difficile et peut compliquer sa délétion puisque selon l'isoforme, les sgRNAs ne ciblent pas la même région du gène, menant à des mutations aux conséquences différentes. La perspective de générer des mutants pour *PAX4* était intéressante, notamment dans le but de générer des double-mutants pour *ARX* et *PAX4*, puisque ces deux facteurs de transcription ont des fonctions antagonistes dans la différenciation entéroendocrine chez la souris (Beucher et al., 2012).

## 2.5. Analyse de la pluripotence des hiPSCs LV37 et des clones NGN3<sup>-/-</sup> et RFX6<sup>-/-</sup> générés

Avant de différencier les lignées générées en tissu intestinal *in vitro*, la pluripotence des hiPSCs LV37 et de certains clones générés a été analysée (**Figure 47**). Les différentes manipulations génétiques et procédures expérimentales peuvent avoir un impact sur la pluripotence et le caryotype des hiPSCs, notamment par la potentielle survenue de recombinaisons chromosomiques non souhaitées. Les hiPSCs LV37 (20 passages) présentent un caryotype femelle normal et l'immunofluorescence sur le marqueur de pluripotence OCT4 suggère le caractère pluripotent des cellules (**Figure 47A-B**). Ceci est confirmé par l'analyse par RT-qPCR sur les marqueurs de pluripotence *OCT4*, *NANOG*, *SOX2* et *LIN28A* (**Figure 47C**). Les clones NGN3<sup>-/-</sup> 2, 10 et 49Cd, ainsi que les clones RFX6<sup>-/-</sup> 1, 14 et 18 semblent également avoir conservé leurs caractéristiques de cellules pluripotentes malgré les manipulations génétiques et le nombre de passage élevé (P60), en comparaison aux LV37 non génétiquement modifiées et de nombre de passage peu élevé (P20), et à l'inverse des contrôles de cellules différenciées HeLa et McF-7 (**Figure 47C**).



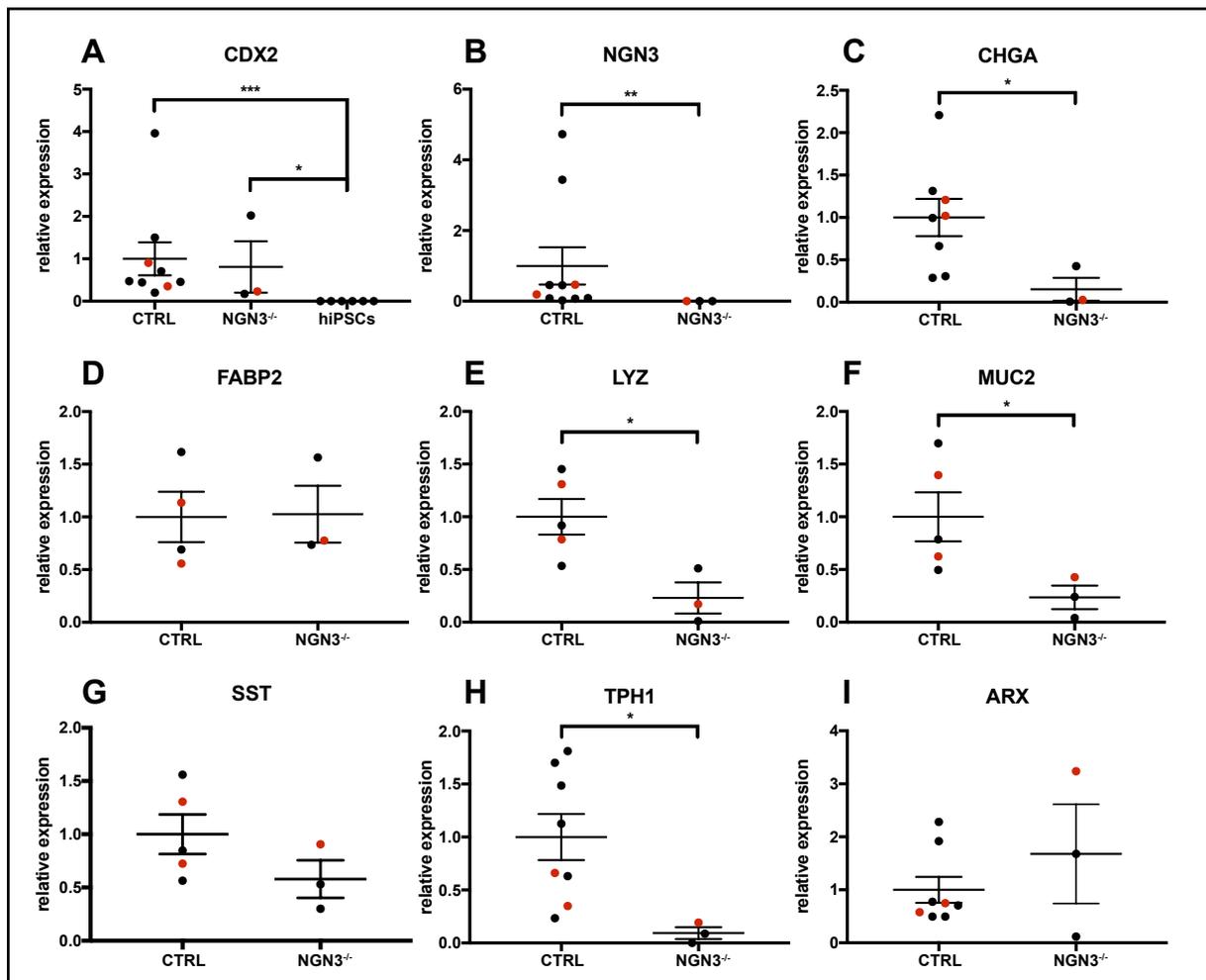
**Figure 47. Analyse du caryotype des hiPSCs LV37 et de la pluripotence des clones NGN3<sup>-/-</sup>, RFX6<sup>-/-</sup> et des LV37**

(A) Immunofluorescence sur le marqueur de pluripotence OCT4 sur des LV37 (P20 = 20 passages). (B) Caryotype des LV37 P20. (C) RT-qPCR sur les marqueurs de pluripotence OCT4, NANOG, SOX2 et LIN28A sur les clones NGN3<sup>-/-</sup> (Clones 2, 10: 60 passages; 49Cd: 45 passages) et RFX6<sup>-/-</sup> (clones 1, 14 et 18: 90 passages). LV37 : contrôle positif de cellules pluripotentes (30 passages), HeLa et McF-7 : contrôles négatifs de cellules différenciées. RT-qPCR réalisées en duplicat.

### Partie 3. Etude des conséquences de l'inactivation de NGN3 sur la différenciation entéroendocrine

L'objectif à long terme de ce projet de recherche est de procéder à une caractérisation des HIOs NGN3<sup>-/-</sup> et RFX6<sup>-/-</sup> afin d'étudier les conséquences de l'inactivation de ces facteurs de transcription sur le développement intestinal chez l'homme, en focalisant plus particulièrement sur la différenciation des cellules entéroendocrines. En effet, Ngn3 est nécessaire à la formation de tous les sous-types entéroendocrines et Rfx6 est initialement exprimé dans l'endoderme intestinal primitif mais également en aval de Ngn3 pour la détermination de certains sous-types entéroendocrines chez la souris (Piccand et al., manuscrit en préparation ; Mellitzer et al., 2010 ; Soyer et al., 2010 ; Jenny et al., 2002). De plus, la seule étude visant à caractériser la fonction de NGN3 dans la différenciation entéroendocrine chez l'homme suggère que sa fonction est conservée. Cependant, le modèle utilisé correspond à une perte de fonction partielle dans les HIOs et la seule observation réalisée est une absence de cellules CHGA<sup>+</sup> (Spence et al., 2011). Ainsi, il est nécessaire d'établir un modèle de perte de fonction totale et de caractériser en détail les modifications phénotypiques induites, notamment au niveau transcriptomique.

Les cellules entéroendocrines ne représentent que 1% des cellules de l'épithélium intestinal, ce qui correspond à un nombre très faible de cellules au sein des HIOs. Il est donc difficile de déceler des variations dans l'expression des gènes impliqués dans la différenciation de cette population rare de cellules. L'obtention de résultats interprétables quant à une variation d'expression génique au sein de cette faible population cellulaire nécessitera un effectif d'échantillons important couplée à une analyse fine des données obtenues. Le cœur du projet consiste en une analyse transcriptomique par séquençage des ARNm (RNA-seq) pour déterminer les gènes différentiellement exprimés dans les HIOs NGN3<sup>-/-</sup> et RFX6<sup>-/-</sup> en comparaison aux HIOs contrôles (CTRL). Préalablement à ces expériences de RNA-seq, des validations du caractère intestinal et épithélial des tissus générés, ainsi que la confirmation du génotype des clones ont été réalisées par RT-qPCR et analyse d'imagerie cellulaire sur des marqueurs prédéfinis. Cette partie détaillera les résultats obtenus dans le cadre de l'analyse des HIOs NGN3<sup>-/-</sup>.



**Figure 48. Analyse des HIOs NGN3<sup>-/-</sup> au stade J28 par RT-qPCR sur divers marqueurs intestinaux d'intérêt**

hiPSCs : cellules indifférenciées (contrôle négatif), CTRL: HIOs LV37 à J28 (contrôle positif), NGN3<sup>-/-</sup> : HIOs NGN3<sup>-/-</sup> clones 2 et 10 à J28. Chaque point représente un regroupement d'une quinzaine d'organoïdes. Les points rouges représentent les échantillons dont l'ARNm a été séquencé. (A) CDX2 : marqueur général d'épithélium intestinal, (B) NGN3 : progéniteurs entéroendocrines, (C) CHGA: cellules entéroendocrines, (D) FABP2 : entérocytes, (E) LYZ : cellules de Paneth, (F) MUC2 : cellules à mucus, (G) SST : somatostatine, (H) TPH1 : sérotonine, (I) ARX : progéniteurs des cellules L. Normalisation par le gène de référence RPLP0. \*p<0,05 ; \*\*p<0,01 ; \*\*\*p<0,001 Test de Mann-Whitney.

## 1. Analyses par RT-qPCR et immunofluorescences

### 1.1. RT-qPCR

Les HIOs différenciés à partir des clones NGN3<sup>-/-</sup> 2 et 10 (**partie 2.1.**) ont été récoltés au stade J28 de deux expériences de différenciation indépendantes et l'ARN en a été extrait afin de l'analyser par RT-qPCR (**Figure 48**). Pour évaluer le caractère épithélial et intestinal des tissus générés, l'expression du marqueur intestinal CDX2 a été étudiée. Il n'y a pas de différence significative (test de Mann-Whitney U=9, p=0,48)

entre l'ARNm de *CDX2* des HIOs CTRL et des *NGN3*<sup>-/-</sup> et son expression, bien que très variable, semble être caractéristique des HIOs puisqu'elle est absente dans les cellules indifférenciées (hiPSCs) (CTRL vs hiPSCs : U=0, p=0,0004 ; *NGN3*<sup>-/-</sup> vs hiPSCs : U=0, p=0,0238) (**Figure 48A**). L'inactivation de *NGN3* a été confirmée puisqu'aucun transcrite n'est détecté dans les HIOs *NGN3*<sup>-/-</sup> (*NGN3*<sup>-/-</sup> vs CTRL : U=0, p=0,007) (**Figure 48B**). La chromogranine A (CHGA) est une glycoprotéine soluble acide localisée dans les granules de sécrétion des cellules endocrines et neuroendocrines (Bartolomucci et al., 2011) et son expression nous permet donc de déterminer la présence de cellules entéroendocrines dans nos HIOs. Une baisse significative de *CHGA* est observée dans les HIOs *NGN3*<sup>-/-</sup> (*NGN3*<sup>-/-</sup> vs CTRL : U=2, p=0,048), suggérant une diminution du nombre de cellules entéroendocrines (**Figure 48C**). Cette observation pouvait être attendue puisqu'une diminution de 95 à 100% de l'ARNm de *ChgA* est observée chez la souris *Ngn3*<sup>Δint</sup> (Mellitzer et al., 2010). De plus, une absence de cellules *ChgA*<sup>+</sup> chez la souris déficiente pour *Ngn3* au stade embryonnaire et adulte et dans le modèle de perte de fonction partielle de *NGN3* dans les HIOs est observée (Spence et al., 2011 ; Mellitzer et al., 2010 ; Jenny et al., 2002). La présence des autres sous-types cellulaires de l'épithélium intestinal a été déterminée grâce à l'expression de marqueurs spécifiques. L'expression de *FABP2* (*FABP2/IFABP, Intestinal-Fatty Acid Binding Proteins*), marqueur des entérocytes (Storch et Corsico, 2008), est similaire dans les HIOs *NGN3*<sup>-/-</sup> et CTRL (U=5, p=0,086) (**Figure 48D**). Par contre, de manière surprenante, l'expression de *LYZ* (lysozyme) associée aux lysozymes sécrétés par les cellules de Paneth et l'expression de *MUC2* (mucine) associée aux mucines sécrétées par les cellules à mucus semblent être diminuées dans les HIOs *NGN3*<sup>-/-</sup> (*NGN3*<sup>-/-</sup> vs CTRL : *LYZ* : U=0, p=0,036 ; *MUC2* : U=0, p=0,036) (**Figure 48E-F**). Cependant, la variabilité inter-organoïdes et le faible effectif d'échantillons *NGN3*<sup>-/-</sup> ne permettent pas de conclure quant à un rôle plus large de *NGN3* chez l'homme, à savoir une implication dans la différenciation de toutes les lignées sécrétrices de l'épithélium intestinal. Chez la souris déficiente pour *Ngn3*, les autres types cellulaires de l'épithélium ne semblent pas affectés.

Seul le nombre de cellules à mucus est faiblement augmenté à des stades post-nataux précoces (Jenny et al. 2002), mais cette observation n'est pas retrouvée chez la souris *Ngn3*<sup>Δint</sup> (Mellitzer et al., 2010). Le niveau d'expression trop faible voire non détecté des gènes codant certaines hormones dans les expériences de RT-qPCR (*GCG, GIP* et *CCK*) nous empêche malheureusement de déceler une éventuelle différence d'expression entre les HIOs *NGN3*<sup>-/-</sup> et CTRL. Une explication plausible à cette absence d'expression serait que certains sous-types de cellules entéroendocrines ne soient pas formés dans notre système *in vitro*, hypothèse qui a déjà été émise, puisqu'une absence de cellules exprimant l'hormone *CCK* a été mise en évidence dans les HIOs (Sinagoga et al., 2018). Il se peut que la différenciation de certains

lignages entéroendocrines se produise plus tard dans le développement ou nécessite des facteurs qui sont absents du système *in vitro*. Il est également envisageable que le nombre de cellules de certains sous-types entéroendocrines soient trop faibles pour détecter les transcrits des gènes codant les hormones par RT-qPCR, d'autant plus que la faible quantité de matériel analysée constitue une réelle limite dans notre modèle. Effectivement, les cellules entéroendocrines représentent seulement 1% de l'épithélium intestinal *in vivo* et correspondent donc à un très petit nombre de cellules dans les HIOs. Pour contrecarrer ces limitations du système *in vitro*, une méthode de transplantation des HIOs sous la capsule rénale de souris immunodéficientes a été développée et permet aux HIOs une expansion et maturation additionnelles *in vivo*, achevant ainsi la différenciation de tous les sous-types entéroendocrines (Watson et al., 2014).

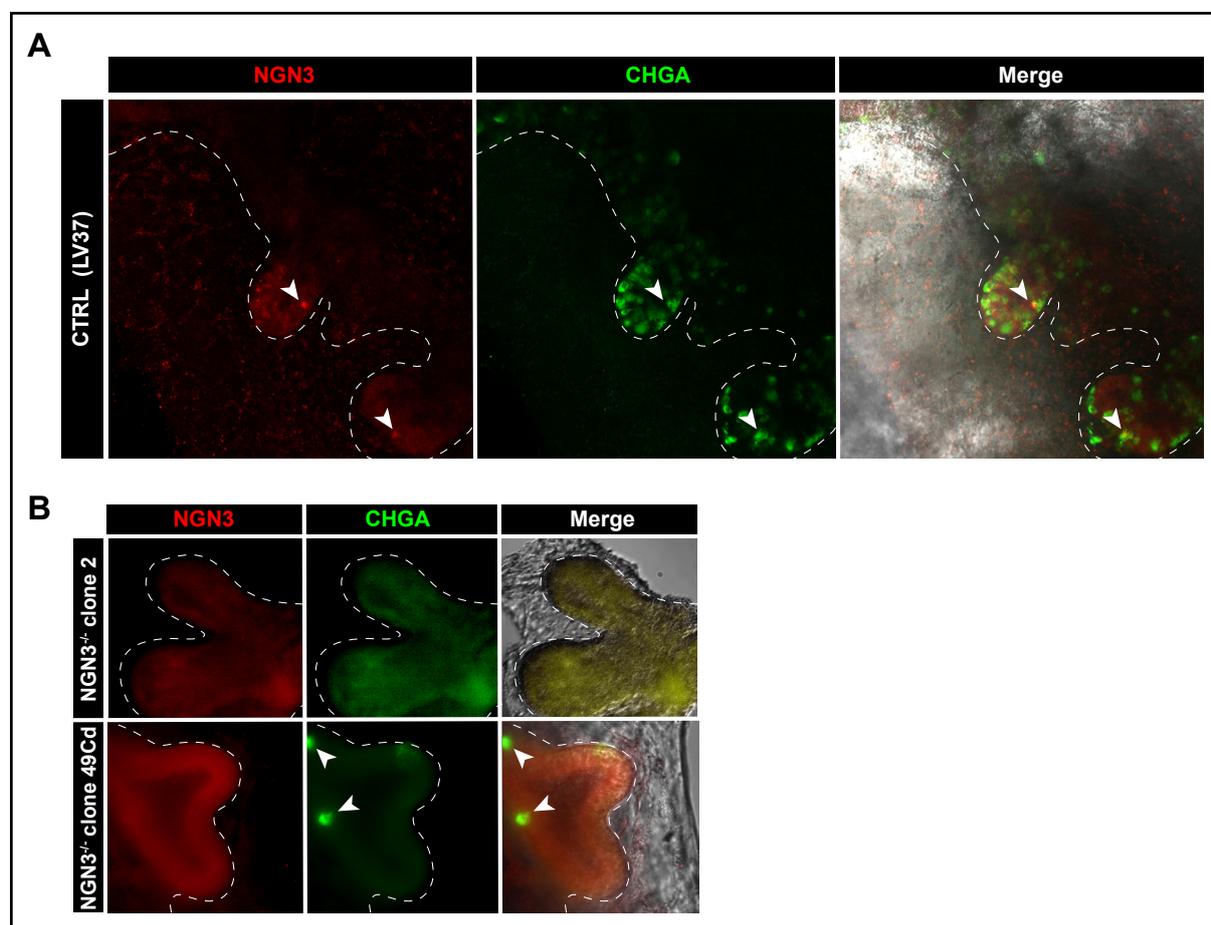
Toutefois, certains gènes codant des hormones et enzyme de biosynthèse de celles-ci ont tout de même pu être analysés dans notre système, tels que *SST* et *TPH1* (enzyme de biosynthèse de la sérotonine). Le niveau d'expression de *TPH1* est diminué de manière significative dans les *NGN3<sup>-/-</sup>* (*NGN3<sup>-/-</sup>* vs CTRL :  $U=0$ ,  $p=0,012$ ) (**Figure 48H**) comme c'est le cas chez la souris, alors que *SST* n'est pas différenciellement exprimé (*NGN3<sup>-/-</sup>* vs CTRL :  $U=3$ ,  $p=0,25$ ) (**Figure 48G**), contrairement à l'absence de *SST* observée chez les souris *Ngn3<sup>Δint</sup>* (Mellitzer et al. 2010). Cette différence de modulation de l'expression de *SST* entre le modèle murin et les HIOs *in vitro* peut être directement causée par la nature de l'organisme étudié ou les contextes différents (*in vitro* ou *in vivo*). Tout comme les gènes codant les hormones, les gènes codant les facteurs de transcription connus pour être impliqués dans la cascade de régulation transcriptionnelle de la différenciation entéroendocrine chez la souris, sont faiblement voire pas du tout exprimés dans les HIOs. *PAX4*, par exemple, n'est détecté dans aucun HIO alors que *ARX*, facteur de transcription à l'action antagoniste de *PAX4*, est exprimé dans les HIOs *NGN3<sup>-/-</sup>* et CTRL sans différence d'expression significative ( $U=8$ ,  $p>0,999$ ) (**Figure 48I**). Chez la souris, *Arx* est impliqué dans la différenciation des cellules L sécrétrices de GLP1 (Beucher et al., 2012) et son rôle en aval de *Ngn3* aurait ainsi pu expliquer un impact de l'inactivation de *NGN3* sur *ARX*, dans le cas où leurs fonctions seraient similaires chez l'homme.

En conclusion, les HIOs semblent être formés de tissu de nature épithéliale et intestinale qui expriment des marqueurs des différents types cellulaires de l'épithélium. Cependant, les gènes codant certaines hormones et facteurs de transcription ne sont pas détectés dans notre modèle, à cause d'une quantité insuffisante de matériel ou encore de l'immaturation des HIOs générés, qui sont de caractère foetal plutôt qu'adulte (Finkbeiner et al., 2015; Spence et al., 2011). L'inactivation de *NGN3*, confirmée par l'absence de transcrits *NGN3* chez les mutants, semble affecter la différenciation de certains sous-types de cellules entéroendocrines (baisse de *CHGA* et *TPH1*) et

pourrait éventuellement avoir un rôle dans la différenciation de l'ensemble des lignées sécrétrices chez l'homme (baisse de *LYZ* et *MUC2*), inconnu à ce jour, mais qui reste à confirmer.

## 1.2. Immunofluorescences *in toto*

Après analyse comparative des transcrits des HIOs *NGN3<sup>-/-</sup>* et CTRL par RT-qPCR, des immunofluorescences des marqueurs entéroendocrines ont été réalisées sur des HIOs entiers (*in toto*), afin de confirmer la délétion de *NGN3* et ses conséquences au niveau protéique (**Figure 49**). Les marquages des hormones (GLP1, GIP et 5-HT/sérotonine) n'ont malheureusement pas fonctionné, peut-être pour des raisons techniques ou bien pour des raisons similaires aux RT-qPCR, c'est-à-dire à cause du stade d'étude précoce (J28) et du caractère foetal des HIOs qui n'expriment pas forcément toutes les hormones produites *in vivo*.



**Figure 49. Immunofluorescences du marqueur entéroendocrine CHGA et NGN3 sur les HIOs *NGN3<sup>-/-</sup>* et CTRL au stade J28**

(A) Images prises au microscope confocal du double-marquage NGN3 et CHGA sur un HIO CTRL entier représentatif (n=4). Les flèches indiquent deux cellules double-positives *NGN3<sup>+</sup> CHGA<sup>+</sup>*. (B) Images prises au microscope à épifluorescence du double-marquage NGN3 et CHGA sur un HIO *NGN3<sup>-/-</sup>* issu du clone 2 et un issu du clone 49Cd (n=3 par clone). Les flèches indiquent deux cellules *CHGA<sup>+</sup>*. Les pointillés délimitent la structure des HIOs : partie droite sur le panneau A.

Seule l'immunofluorescence sur NGN3 et le marqueur de cellules entéroendocrines CHGA a pu être exploitée. Celle-ci nous a permis de détecter des cellules qui semblent co-exprimer NGN3 et CHGA, et un nombre important de cellules CHGA<sup>+</sup> dans les HIOs CTRL (**Figure 49A**). L'absence de cellules NGN3<sup>+</sup> dans les HIOs issus du clone 49Cd, du clone 2 (**Figure 49B**) et du clone 10 (non montré) confirme a priori la disruption de la protéine causée par la délétion de l'exon codant de *NGN3*. Il faudrait toutefois le confirmer par quantification du marquage sur un nombre plus importants de HIOs. Dans le cas des clones 2 et 10, aucune cellule CHGA<sup>+</sup> n'a pu être observée dans plusieurs HIOs (n=3 par clone) dont un exemple représentatif est illustré en **figure 49B**. Cette absence de cellules entéroendocrines concorde avec l'importante baisse du niveau de transcrits *CHGA* observée par RT-qPCR. Cependant, il subsiste quelques cellules entéroendocrines dans les HIOs issus du clone 49Cd (**Figure 49B**), qui diffère des clones 2 et 10 de par son génotype, car il possède une mutation nulle sur un allèle et une mutation menant à la perte partielle du domaine bHLH sur l'autre, qui pourrait permettre une activité résiduelle de NGN3 en comparaison aux mutants nuls 2 et 10 pour lesquels aucune activité du facteur de transcription ne peut théoriquement avoir lieu. Bien que le système de détection ne soit pas le même (microscope confocal ou microscope à épifluorescence), ces quelques cellules entéroendocrines restent toutefois rares en comparaison à celles retrouvées dans les HIOs CTRL (**Figure 49A**). L'abondance des cellules CHGA<sup>+</sup> dans les HIOs CTRL interroge sur la spécificité du marquage mais dans tous les cas, la différence dans le nombre de cellules marquées entre les mutants et CTRL reste flagrante. Il semblerait donc que les clones 2 et 10 soient bien des mutants nuls pour *NGN3* et que son absence provoquerait une perte totale de cellules entéroendocrines. A l'inverse, une activité résiduelle de NGN3, bien que non détectée par le marquage par immunofluorescence, pourrait persister dans le clone 49Cd puisqu'il semblerait qu'on y retrouve de rares cellules entéroendocrines comme en témoigne le marquage de CHGA. D'après les travaux effectués chez la souris, *Ngn3* est indispensable à la formation de toutes les cellules entéroendocrines (Mellitzer et al., 2010; Jenny et al., 2002), il est donc curieux que des cellules CHGA<sup>+</sup> soient retrouvées dans des HIOs dans lesquels aucune cellule NGN3<sup>+</sup> n'est identifiée. Une hypothèse pouvant expliquer l'absence de marquage NGN3 dans ce clone serait que l'anticorps utilisé pourrait reconnaître une région de la protéine qui est perdue ou bien que le niveau d'expression de NGN3 soit trop faible pour être détecté par immunofluorescence. Un argument pouvant appuyer cette dernière hypothèse est le nombre de cellules NGN3<sup>+</sup> qui paraît très faible en comparaison aux cellules CHGA<sup>+</sup> dans les HIOs CTRL.

### 1.3. Conclusion des analyses par RT-qPCR et immunofluorescences

En conclusion, l'ensemble des HIOs générés sont formés de tissu intestinal exprimant des marqueurs de cellules épithéliales à des niveaux similaires (*CDX2*, *FABP2*). Les différences observées pour les autres marqueurs sont donc vraisemblablement causées par la mutation génétique *NGN3*<sup>-/-</sup> et non par des problèmes dans les expériences de différenciation. Par contre, il n'est pas certain que le tissu formé soit entièrement de nature intestinale et il faudrait analyser des marqueurs d'autres tissus pour déterminer si les HIOs sont constitués d'un mélange de tissus d'origines différentes. Les HIOs ne sont de toute façon pas uniquement composés de cellules épithéliales mais également de cellules mésenchymateuses, dont des fibroblastes et cellules de muscles lisses (Spence et al., 2011), et il serait ainsi intéressant d'évaluer la proportion de ces tissus par rapport à l'épithélium intestinal et de déterminer si des tissus additionnels non attendus se forment. Ces analyses par RT-qPCR et immunofluorescence nous permettent d'affirmer que les HIOs différenciés à partir des clones 2 et 10 sont bien inactivés pour *NGN3*, puisqu'aucun transcrite n'est détecté et qu'aucune cellule ne semble exprimer le facteur de transcription, ce qui est en concordance avec le génotypage des clones par PCR et séquençage des produits PCR (**voir partie 2, 2.1.2.**). De plus, l'expression de l'ARNm de *CHGA* est diminuée dans ces HIOs et une perte de cellules entéroendocrines a été mise en évidence par l'absence de cellules marquées *CHGA*<sup>+</sup>. Il est surprenant d'observer une absence totale de cellules *CHGA*<sup>+</sup> alors que son ARNm semble encore exprimé. Toutefois, il a déjà été mis en évidence que le niveau d'expression des ARNm n'est pas forcément corrélé avec l'expression protéique des hormones (Beucher et al., 2012 ; Mortensen et al., 2003). Il a été montré dans des études sur cellules entéroendocrines isolées, par exemple, qu'une cellule entéroendocrine exprimant une ou deux hormones au niveau protéique, est capable d'exprimer jusqu'à six transcrits d'hormones différentes (Egerod et al., 2012).

Malgré le fait que la majorité des hormones n'aient pu être détectées dans les HIOs mutants et CTRL, le niveau d'expression de *SST* et *TPH1* a pu être quantifié par RT-qPCR et a révélé une expression quasiment nulle de *TPH1* dans les HIOs *NGN3*<sup>-/-</sup> issus des clones 2 et 10 et une absence de différence d'expression de *SST*. L'expression inchangée de *SST* est surprenante au vu de l'absence de cellules *CHGA*<sup>+</sup>. De plus, cette observation n'est pas concomitante avec l'altération de la différenciation de toutes les cellules entéroendocrines observée chez la souris déficiente pour *Ngn3*, puisque l'épithélium de ces souris est dépourvu de cellules *ChgA*<sup>+</sup> et l'expression d'aucune hormone n'est détectée (Mellitzer et al., 2010; Jenny et al., 2002). Spence et al. ont également mis en évidence une absence de cellules *CHGA*<sup>+</sup> dans des HIOs dont l'expression de *NGN3* a été altérée grâce à des vecteurs lentiviraux exprimant

des shRNAs dirigés contre *NGN3* (Spence et al., 2011). Cette observation suggère que la différenciation des cellules D sécrétrices de somatostatine n'est pas affectée dans les HIOs *NGN3*<sup>-/-</sup>.

D'après les RT-qPCR sur les HIOs *NGN3*<sup>-/-</sup>, l'inactivation de *NGN3* semble affecter toutes les lignées sécrétrices, y compris les cellules de Paneth et les cellules à mucus. Pourtant, l'inactivation de *Ngn3* spécifiquement dans l'intestin n'a aucun impact sur la différenciation des autres types cellulaires de l'épithélium chez la souris (Mellitzer et al., 2010), à l'exception d'une légère augmentation des cellules à mucus chez la souris *Ngn3*<sup>-/-</sup> (Jenny et al., 2002), bien qu'il a été suggéré que certains progéniteurs exprimant *Ngn3* pouvaient se différencier en cellules de Paneth et cellules à mucus, correspondant ainsi à un progéniteur commun à tous les lignages sécréteurs (Schonhoff et al., 2004). Nos observations sont toutefois à confirmer puisque cela reste des résultats préliminaires sur un nombre d'échantillon limité.

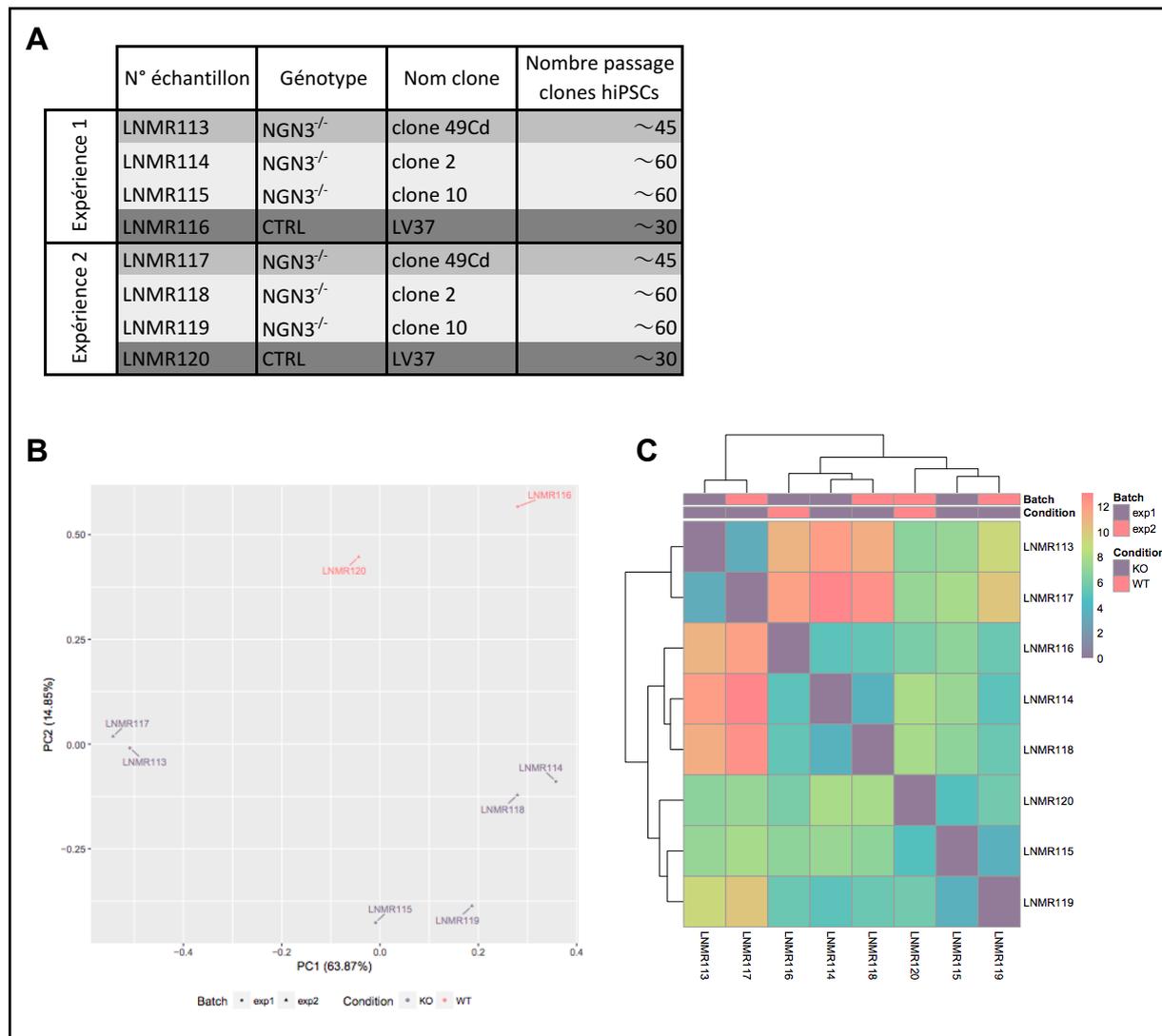
Concernant les HIOs issus du clone 49Cd, aucune cellule *NGN3*<sup>+</sup> n'est présente mais il subsiste quelques rares cellules entéroendocrines (*CHGA*<sup>+</sup>), possiblement dû à la mutation non nulle présente sur un des allèles. En effet, cette mutation pourrait mener à la production d'une protéine dont seulement une partie du domaine bHLH a été délétée, provoquant une activité résiduelle de *NGN3* qui serait possiblement suffisante à la différenciation de certains sous-types entéroendocrines. Ceci dit, la présence de cellules *CHGA*<sup>+</sup> n'atteste pas forcément de la production d'hormones mais uniquement de la présence de cellules endocrines possédant des granules sécrétoires. Ainsi, il est possible que ce clone soit capable de former certains types de cellules entéroendocrines mais qui n'expriment aucune hormone. Pour déterminer si tel est le cas, il faudrait directement analyser l'expression et la production des hormones. Le clone 49Cd ne possédant pas la mutation nulle attendue, il n'a pas été pris en compte dans la suite des analyses, afin d'éviter d'inclure un phénotype potentiellement incomplet.

## 2. Séquençage des ARNm (RNA-seq)

Afin d'étudier le transcriptome des HIOs *NGN3*<sup>-/-</sup> plus en détails et de valider nos observations préliminaires, une expérience de RNA-seq a été réalisée et les gènes différentiellement exprimés ont été analysés. Chaque échantillon est constitué d'une quinzaine d'organoïdes rassemblés et chaque génotype est représenté par deux échantillons issus de deux différenciations indépendantes (**Figure 50A**). Les données de RNA-seq ont été analysées par la bio-informaticienne de l'équipe Constance Vagne. Les détails de la méthode d'analyse sont disponibles dans la section dédiée du matériel et méthodes (**voir 3.3**).

## 2.1. Différence entre chaque échantillon

Tout d'abord, une analyse de la différence entre chaque échantillon, basée sur les niveaux d'expression de l'ensemble des gènes, a été réalisée (**Figure 50**). L'analyse en composantes principales (PCA/ACP) met en évidence trois principaux regroupements, l'un avec les HIOs issus du clone 49Cd, un deuxième avec les HIOs NGN3<sup>-/-</sup> issus des clones 2 et 10 et le dernier avec les HIOs CTRL (**Figure 50B**).



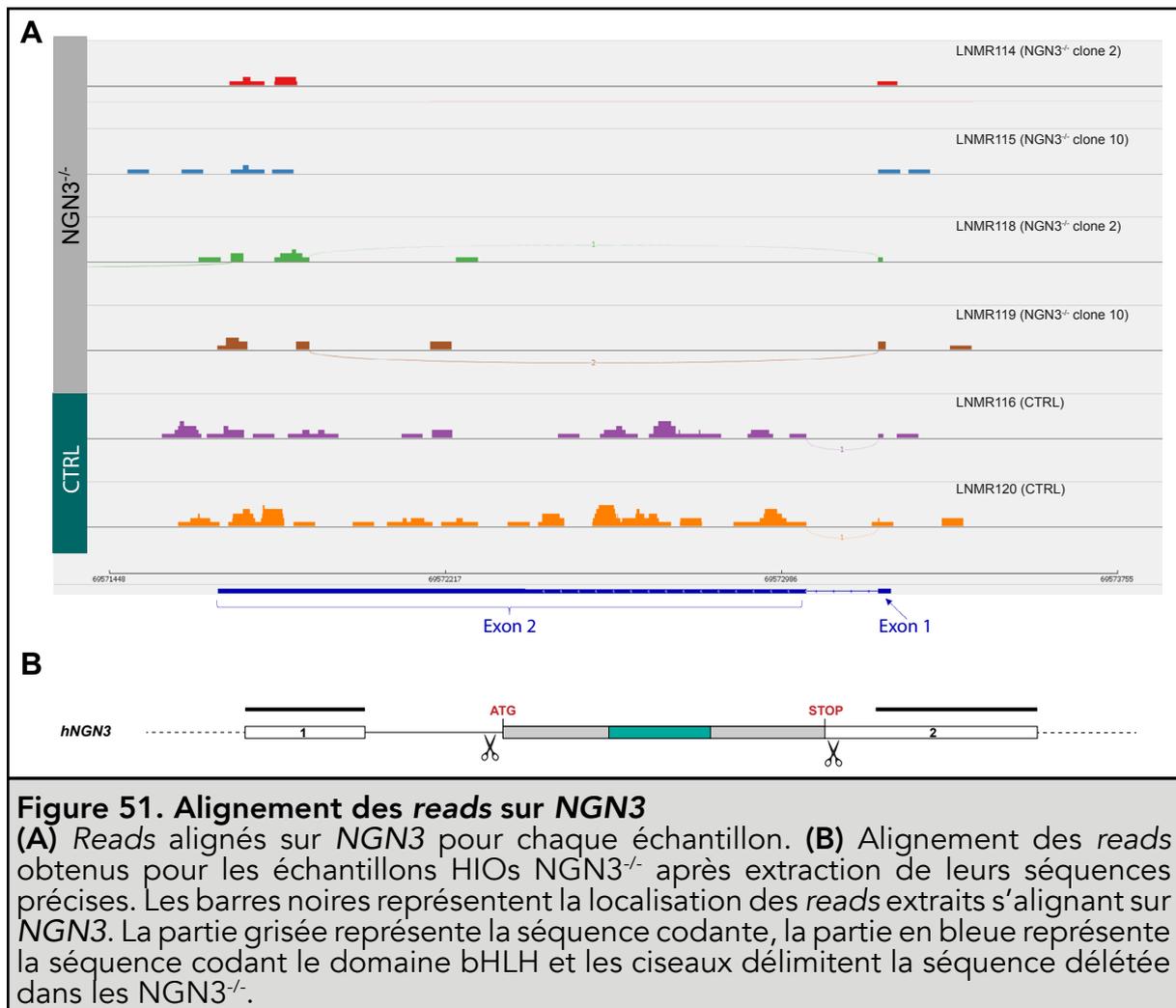
**Figure 50. Analyse de la différence entre chaque échantillon des expériences de RNA-seq**

(A) Tableau récapitulatif des échantillons dont l'ARNm a été séquencé : deux expériences de différenciation distinctes ont été menées. Pour chaque expérience les clones des différents génotypes ont été différenciés en HIOs. Les numéros d'échantillons sont associés au génotype et au clone correspondant. Le nombre de passage des clones lors du début des expériences de différenciation est donné à titre indicatif à  $\pm 2$  passages. Gris foncé : CTRL, gris moyen : clone 49Cd, gris clair : NGN3<sup>-/-</sup> clones 2 et 10. (B) Analyse en composantes principales (PCA/ACP). (C) Dendrogramme des échantillons en fonction du génotype et de l'expérience. Condition: WT=CTRL, KO=NGN3<sup>-/-</sup>; Batch: exp1=expérience 1, exp2=expérience 2.

L'hétérogénéité des échantillons est également représentée par un dendrogramme qui permet d'identifier leur regroupement selon l'expérience et le génotype. A priori, il n'y a pas de regroupement des échantillons selon l'expérience et cet effet, appelé communément *batch effect*, n'a donc pas été pris en compte dans la normalisation des données de séquençage. Comme observé sur la PCA, les HIOs issus du clone 49Cd forment un groupe bien distinct, confirmant ainsi la différence entre HIOs issus de ce clone et HIOs issus des clones 2 et 10, appuyant notre décision de ne pas inclure ces échantillons parmi les HIOs NGN3<sup>-/-</sup> dans les analyses. Les HIOs issus du clone 2 et des deux expériences différentes semblent former un groupe, de la même manière que les HIOs issus du clone 10. Par contre, les HIOs CTRL ne se regroupent pas entre eux (**Figure 50C**). Ces regroupements semblent cohérents puisque les mutants forment des groupes par génotype, malgré une hétérogénéité entre les deux échantillons HIOs CTRL.

## 2.2. Alignements des *reads* obtenus sur la séquence d'ADN codant NGN3

Dans le but de confirmer à nouveau le génotype des HIOs NGN3<sup>-/-</sup> issus des clones 2 et 10, les *reads* obtenus ont tout d'abord été alignés sur le génome de référence humain Hg38. Les *reads* s'alignant sur la séquence d'ADN codant NGN3 ont été identifiés et leurs séquences extraites afin de les localiser précisément sur le gène (**Figure 51**). Ceci nous permet de déceler d'éventuels *reads* s'alignant sur la séquence codante de NGN3 qui est supposée être déléetée dans les HIOs NGN3<sup>-/-</sup>. D'après le génotypage initial des clones 2 et 10, la délétion induite comprend toute la séquence codante située sur la première partie de l'exon 2 du gène. L'alignement créé à partir du logiciel IGV (*Integrative Genomics Viewer*, Robinson et al., 2011) montre une absence de *reads* alignés sur la majeure partie de l'exon 2 du gène pour tous les échantillons NGN3<sup>-/-</sup> (LNMR114, 115, 118, 119) par comparaison avec les échantillons CTRL (LNMR116, 120), pour lesquels plusieurs *reads* s'alignent sur cette région (**Figure 51A**). Les séquences des *reads* s'alignant sur l'exon 1 et une partie de l'exon 2 ont été extraites pour les échantillons NGN3<sup>-/-</sup>. Aucune séquence ne s'aligne au niveau de la délétion attendue (**Figure 51B**) et cette observation démontre donc de manière fiable que les HIOs NGN3<sup>-/-</sup> issus des clones 2 et 10 sont bien déficients pour NGN3 et l'analyse de l'expérience de RNA-seq de ces échantillons devrait donc être représentative de l'inactivation de NGN3 dans notre modèle humain *in vitro*.



### 2.3. Nature intestinale et épithéliale des HIOs générés

Bien que les analyses précédentes par RT-qPCR et immunofluorescences indiquent la présence de marqueurs intestinaux dans les HIOs générés, l’analyse des données de RNA-seq autorise une étude plus fine et détaillée de l’expression de marqueurs des différentes populations cellulaires de l’épithélium intestinal. L’expression des gènes caractéristiques des types cellulaires retrouvés dans l’épithélium intestinal a été comparée dans les HIOs NGN3<sup>-/-</sup> et les HIOs CTRL pour déterminer la nature intestinale des tissus et déceler les éventuelles conséquences de la mutation sur la différenciation épithéliale intestinale (**Figure 52**). Les marqueurs utilisés pour identifier ces cellules sont extraits de différents articles issus de la littérature récente, à savoir d’une expérience similaire de RNA-seq sur des HIOs (Workman et al., 2016) et d’une expérience de RNA-seq sur des cellules épithéliales isolées (*single cell* RNA-seq) à partir d’intestin grêle et d’entéroïdes de souris, qui a permis d’identifier des marqueurs spécifiques aux sous-types cellulaires de l’épithélium (Haber et al., 2017).

	Gènes	Log2 FC	P-valeur ajustée	Moyenne CTRL	Moyenne NGN3 <sup>-/-</sup>
Cellules souches	LGR5	-0,31	6,57E-01	3213	2594
	OLFM4	0,48	2,85E-01	330	459
	ASCL2	-2,06	4,80E-02	49	11
	BMI1	-0,26	7,63E-01	1293	1080
	TERT	-0,57	7,20E-01	92	62
Cellules de Paneth	LYZ	0,65	4,84E-01	6110	9584
	DLL1	0,08	9,73E-01	2471	2614
	DLL4	-0,31	6,79E-01	179	144
	XBP1	0,31	4,79E-01	5099	6304
Cellules à mucus	MUC2	1,02	1,53E-01	358	726
	AGR2	0,27	8,93E-01	19428	23405
	KIT	0	1,00E+00	728	729
	SPEDF	-0,35	9,24E-01	75	59
	TFF3	-0,53	9,05E-01	3734	2580
Entérocytes	SLC2A2	0,96	4,07E-01	68	133
	SLC26A6	-0,12	9,48E-01	1180	1088
	SLC9A1	0,18	8,37E-01	1608	1820
	VIL1	-0,61	6,08E-01	5661	3712
	CBR1	0,13	9,15E-01	627	684
	EPHX2	-0,15	9,44E-01	798	719
	MEP1A	-1,74	4,87E-03	744	223
Cellules en "touffes"	LRMP	0,97	6,44E-01	7	14
	DCLK1	-0,45	7,57E-01	3092	2267
	PTGS2	0,05	9,83E-01	305	315
	HOPX	0,13	9,53E-01	218	238
	EGFR	0,35	5,38E-01	3494	4467

**Figure 52. Tableau récapitulatif de l'expression de marqueurs spécifiques aux sous-types cellulaires de l'épithélium intestinal**

Les gènes sont regroupés par type cellulaire. Bleu foncé : lignage absorbant, bleu moyen : lignage sécréteur, bleu clair : autres types cellulaires.

Le Log2(FC) (Fold Change) représente la variation d'expression entre HIOs CTRL et HIOs NGN3<sup>-/-</sup>. Une variation d'expression est jugée significative si la p-valeur ajustée est inférieure à 0,05. La moyenne des reads normalisés est indiquée pour les CTRL et les mutants NGN3<sup>-/-</sup>. En rouge sont indiqués les gènes significativement sous-exprimés dans les HIOs NGN3<sup>-/-</sup>.

Dans les cryptes, sont localisées les cellules souches qui donnent naissance à tous les types cellulaires de l'épithélium, dont deux types ont été décrits : celles localisées entre les cellules de Paneth à la base des cryptes, appelées CBC (*Crypt Base Columnar*) et exprimant *Lgr5*, et celles en position +4 exprimant *Bmi1* aussi appelées LRC (*Label-Retaining Cells*).

Ces deux marqueurs sont fortement exprimés dans nos HIOs et aucune différence significative n'est observée entre les mutants et les CTRL. D'autres marqueurs, comme *Ascl2* et *Olfm4* (Munoz et al., 2012) permettent l'identification des cellules souches CBC et *Tert* permet celle des cellules souches LRC (Montgomery et al., 2011). Ces marqueurs sont bien exprimés dans nos HIOs et ne sont pas différenciellement

exprimés selon le génotype, hormis *ASCL2*, dont l'expression est diminuée chez les mutants. Il est difficile de déterminer si ceci résulte d'une conséquence de la mutation puisque le niveau d'expression de ce gène est assez faible. L'inactivation de *NGN3* ne semble globalement pas affecter la production des cellules souches.

Les cellules de Paneth sont les seules cellules différenciées des cryptes et sécrètent des peptides antimicrobiens (lysozymes) chargés de protéger les muqueuses intestinales contre certains micro-organismes pathogènes. Ainsi, l'expression du gène *LYZ* codant les lysozymes est spécifique de ce type cellulaire. Ni l'expression de *LYZ* ni celle des autres marqueurs *DLL1*, *DLL4* et *XBP1* (Workman et al., 2016 ; van Es et al., 2012) n'est altérée dans les HIOs *NGN3*<sup>-/-</sup>. La différenciation des cellules de Paneth ne serait donc pas dépendante de l'expression de *NGN3*, contrairement à ce qui a été observé par RT-qPCR.

Parmi les cellules sécrétrices, on retrouve également les cellules à mucus dans les villosités, qui sécrètent un mucus jouant un rôle dans la protection de l'épithélium intestinal. Elles sont identifiables grâce à l'expression de *MUC2* codant la mucine 2, membre de la famille des mucines, constituant principal du gel protecteur formé. L'expression de ce gène, comme celle d'autres marqueurs tels que *AGR2*, *KIT*, *SPEDF* et *TFF3* (Workman et al., 2016 ; Bergström et al., 2014 ; Rothenberg et al., 2012), n'est pas non plus affectée par l'inactivation de *NGN3*, suggérant l'absence d'implication de *NGN3* dans le lignage des cellules à mucus. Ces observations ne sont pas en accord avec les résultats obtenus par RT-qPCR, puisque toutes les lignées sécrétrices semblaient affectées par la mutation. Plusieurs hypothèses peuvent justifier ces discordances : la première, et la plus probable, est que ces différences soient liées au fait que les échantillons analysés par séquençage ne correspondent pas forcément à ceux étudiés par RT-qPCR (**Figure 48**, points rouges), ou bien qu'une des deux approches soit plus sensible en terme de détection du niveau d'expression des ARNm.

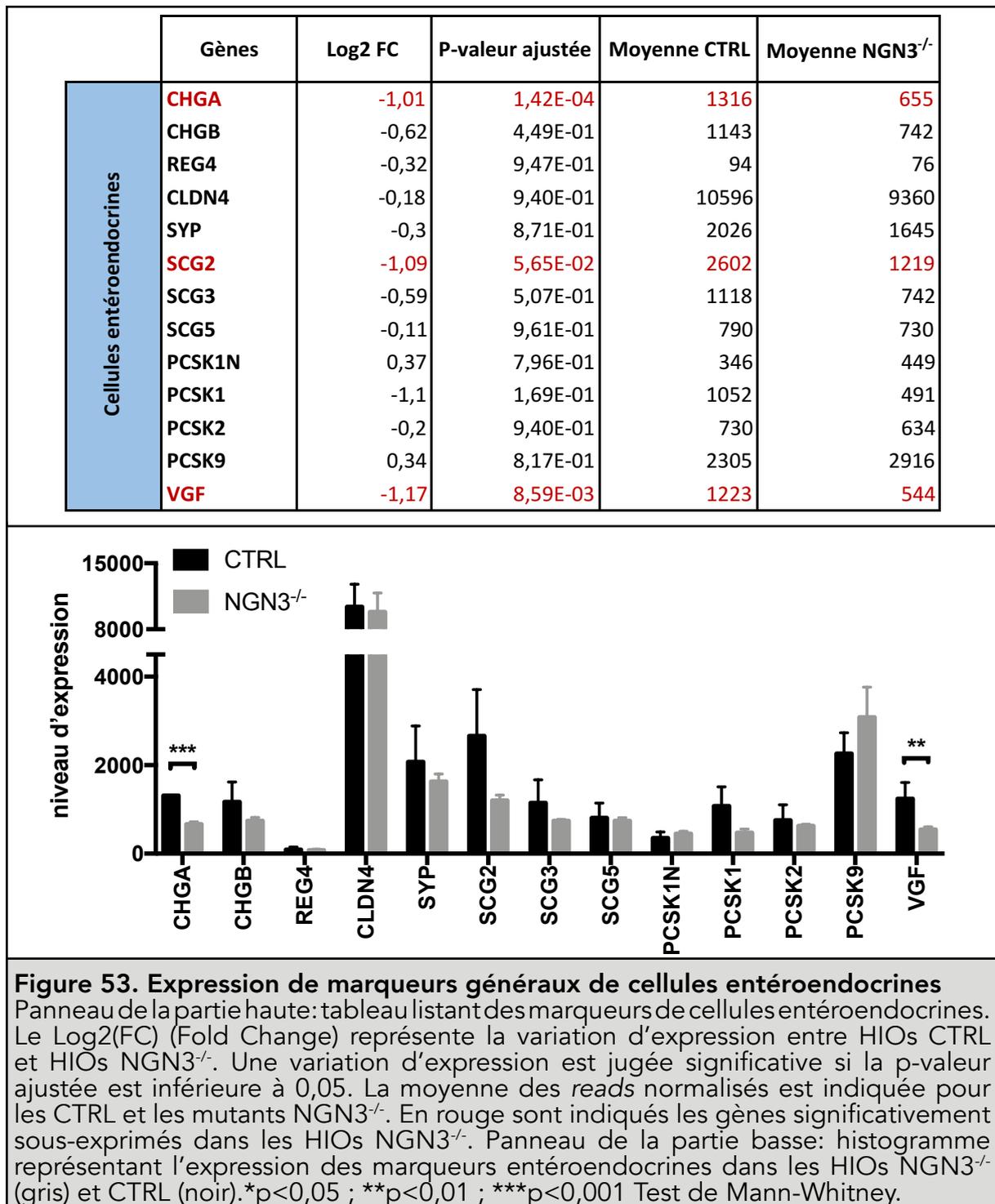
Les entérocytes, cellules absorbantes, sont les plus abondantes dans l'épithélium intestinal et sont impliquées dans la digestion et notamment l'absorption des nutriments. Le transport des nutriments est médié par des transporteurs présents à la membrane de ces cellules, qui constituent ainsi des marqueurs de choix de ce type cellulaire. Parmi ces transporteurs, les gènes codant l'échangeur d'anions *SLC26A6* (*Solute-linked carrier 26*), l'antiport  $\text{Na}^+/\text{H}^+$  *SLC9A1* et le transporteur de glucose *GLUT2*, codé par le gène *SLC2A2*, ont une expression similaire entre les différents génotypes. C'est également le cas du marqueur de la villine *VIL1* et des marqueurs d'entérocytes de l'intestin proximal *CBR1* et *EPHX2* (Haber et al., 2017). Étonnamment, le marqueur d'entérocytes de l'intestin distal *MEP1A* est significativement sous-exprimé dans les HIOs mutants. Il reste à déterminer si la diminution de ce marqueur provient d'une nature moins distale du tissu formé dans les HIOs *NGN3*<sup>-/-</sup> ou d'une diminution du nombre de ce type d'entérocytes.

Les marqueurs de cellules « en touffes » *DCLK1*, *PTGS2*, *HOPX* et *EGFR* sont bien exprimés dans les HIOs et leur expression n'est pas modulée par la mutation. Ainsi, les différents types de cellules sont bien formés dans les HIOs d'après le séquençage et *NGN3* ne semble pas nécessaire à la spécification des cellules à mucus, des cellules de Paneth, des cellules « en touffes » et des entérocytes, comme chez la souris (Mellitzer et al., 2010 ; Jenny et al., 2002).

*Ngn3* étant exprimé par des progéniteurs communs à tous les sous-types entéroendocrines, il est attendu que son inactivation mène à la perte de ces cellules dans les HIOs *NGN3*<sup>-/-</sup>. Cependant, il faut garder à l'esprit que ces cellules constituent seulement 1% des cellules épithéliales et qu'il n'est donc pas aisé de déceler des changements d'expression de gènes dans ces cellules et particulièrement dans un tel modèle.

#### **2.4. Etude de l'expression de marqueurs généraux de cellules entéroendocrines**

Bien qu'il soit difficile de considérer ces cellules comme une population homogène, certains marqueurs ont été décrits comme étant exprimés par l'ensemble des sous-types entéroendocrines. C'est notamment le cas du marqueur générique *CHGA*, de chromogranine C (*SCG2*) et de *VGF* (*Nerve growth factor inducible*), qui est connu pour être enrichi dans les cellules entérochromaffines de l'intestin grêle et du côlon (Roberts et al., 2018 ; Lund et al., 2018), dont l'expression est significativement altérée dans les HIO *NGN3*<sup>-/-</sup> (**Figure 53**). Un certain nombre d'autres marqueurs de ces cellules exprimés dans les HIOs ne semblent pas affectés chez les mutants : *REG4* (Grün et al., 2015), *CLDN4* (Nagatake et al., 2014), *CHGB*, *SCG3*, *SCG5*, *SYP*, *PCSK1*, etc. (Roberts et al., 2018 ; Engelstoft et al., 2015). Il est donc possible que l'absence de *NGN3* ait un impact sur la différenciation de certains sous-types entéroendocrines dans notre modèle *in vitro*, mais pas de tous. Ces marqueurs ne permettent peut-être pas l'identification de l'intégralité des types de cellules entéroendocrines et peuvent être plus représentatifs d'un type de cellule précis, comme *VGF* par exemple, qui est plus spécifiquement exprimé dans les cellules entérochromaffines (Lund et al., 2018). Cette hypothèse sera explorée par l'étude de l'expression des gènes codant les différentes hormones produites par les cellules entéroendocrines ou leur enzyme de biosynthèse.

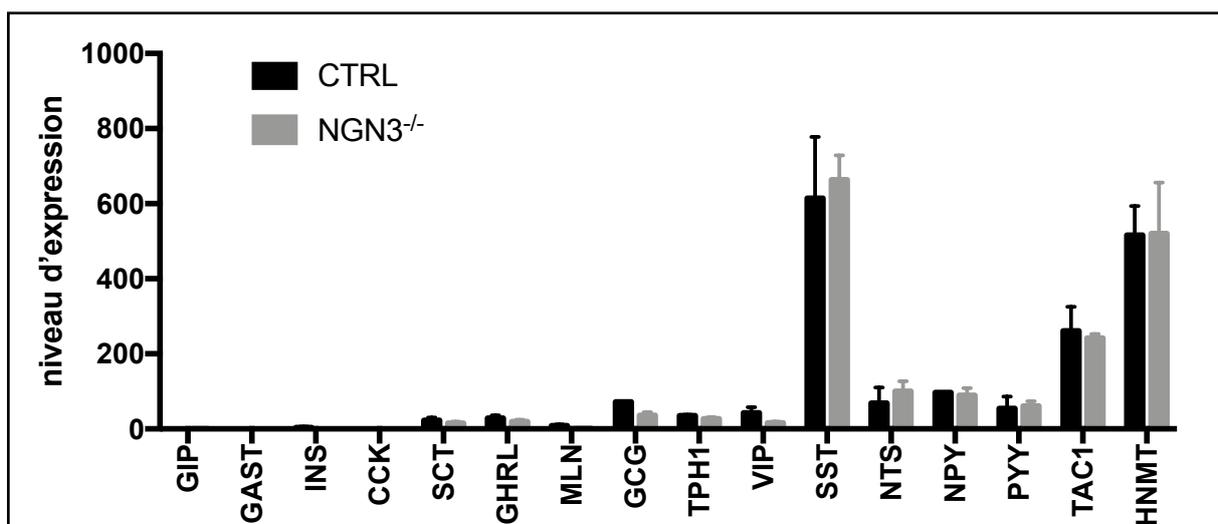


## 2.5. Impact de l'inactivation de *NGN3* sur la différenciation entéroendocrine

### 2.5.1. Expression des gènes codant les produits de sécrétion des cellules entéroendocrines ou leur enzyme de biosynthèse

L'expression des marqueurs entéroendocrines *CHGA*, *SCG2* et *VGF* étant diminuée dans les HIOs *NGN3*<sup>-/-</sup>, l'expression des gènes codant les produits de sécrétion des cellules entéroendocrines (hormone, enzyme de biosynthèse, neuropeptide...) a été analysée afin de déterminer quels sous-types entéroendocrines sont affectés par l'inactivation de *NGN3*. De façon très inattendue, l'expression d'aucun des gènes codant les produits de sécrétion n'est affectée par la mutation (**Figure 54**). De façon similaire aux analyses par RT-qPCR, il est important de noter que le niveau d'expression de certains gènes est nul ou très faible, empêchant l'éventuelle mise en évidence d'une altération dans la formation de certains sous-types entéroendocrines. C'est notamment le cas de *GIP*, *GAST*, *INS* et *CCK* qui ne sont pas du tout exprimés dans nos HIOs.

Ces observations étaient en partie attendues puisque l'insuline n'est exprimée que dans les cellules  $\beta$  des îlots pancréatiques et l'hormone gastrine n'est produite que dans les cellules G de l'épithélium gastrique. Les gènes codant les hormones SCT, GHRL, MLN, GCG, PPY, TPH1 et VIP ne sont que faiblement exprimés, attestant tout de même de la présence des sous-types endocrines associés, mais ne permettant pas d'exposer un potentiel effet de la mutation. En revanche, l'expression de *SST*, *NTS*, *NPY*, *PYY*, *TAC1* (précurseur de la substance P) et *HNMT* (enzyme synthétisant l'histamine) ont un niveau d'expression élevé mais non différenciellement exprimé



**Figure 54. Niveau d'expression des gènes codant les hormones produites dans le tractus digestif ou leur enzyme de biosynthèse**  
Le niveau d'expression des gènes est comparé entre les HIOs *NGN3*<sup>-/-</sup> (gris) et les HIOs CTRL (noir).

entre les HIOs CTRL et *NGN3*<sup>-/-</sup>. Pour ces gènes, il est donc possible de conclure de manière plus assurée quant à une absence d'effet de la mutation de *NGN3* sur la différenciation des cellules productrices de ces hormones dans notre modèle. De façon surprenante, les gènes les plus exprimés *HNMT*, *TAC1* et *SST* sont plutôt exprimés dans la partie proximale du tractus gastrointestinal (**Figure 16**) (Chen et al., 2009). L'ensemble de ces observations va à l'encontre des données de la littérature concernant la fonction de *Ngn3* dans la différenciation entéroendocrine chez la souris, puisqu'une absence totale de cellules entéroendocrines avait été reportée chez la souris déficiente pour *Ngn3* (Mellitzer et al., 2010 ; Jenny et al., 2002). En revanche, ces observations concordent avec nos résultats obtenus par RT-qPCR qui indiquaient une absence d'effet sur l'expression de *SST*. Comme l'expression du marqueur général de cellules entéroendocrines *CHGA* est affectée (diminution de l'ARNm et absence de la protéine), on peut imaginer que *NGN3* contrôle uniquement les programmes génériques de la différenciation entéroendocrine et non la spécification des sous-types qui produisent les hormones chez l'homme, ou au moins dans ce modèle *in vitro*. Mais ceci est en contradiction avec l'analyse des biopsies intestinales des patients porteurs de mutation dans *NGN3* qui présentent une absence de cellules entéroendocrines et des hormones associées (Wang et al., 2006). Il est également possible que *NGN3* soit seulement nécessaire à la différenciation de certains sous-types entéroendocrines que notre approche transcriptomique n'a pas permis de mettre en évidence. Cette hypothèse peut être propre au système *in vitro*, qui permet la génération de tissus intestinaux de nature fœtal et qui n'ont pas toute la complexité de l'organe *in vivo*. Cependant le rôle de *Ngn3/NGN3* semble pourtant conservé entre la souris et l'homme puisque les HIOs dont l'expression de *NGN3* a été altérée avec des shRNAs ne présentent pas de cellules entéroendocrines et des expériences d'induction de l'expression exogène de *NGN3* dans des HIOs ont montré une augmentation du nombre de cellules entéroendocrines mais également de l'expression des différentes hormones (Sinagoga et al., 2018 ; Múnera et al., 2017 ; McCracken et al., 2014). Les raisons pouvant expliquer nos résultats surprenants et les hypothèses abordées constitueront des éléments de discussion par la suite.

### **2.5.2. Expression des gènes codant des facteurs de transcription régulant la différenciation entéroendocrine**

*Ngn3* est inscrit dans une cascade de facteurs de transcription régulant la différenciation entéroendocrine chez la souris. De plus, les HIOs dont l'expression de *NGN3* est altérée par des shRNAs ne présentant aucune cellule entéroendocrine (Spence et al., 2011), il est attendu que cette régulation transcriptionnelle soit perturbée par l'inactivation de *NGN3* dans notre modèle également.

Gènes	Log2 FC	P-valeur ajustée	Moyenne CTRL	Moyenne NGN3 <sup>-/-</sup>
ATOH1/MATH1	0,55	9,16E-01	35	51
HES1	0,26	8,32E-01	1317	1580
NEUROD1	-0,2	9,54E-01	418	364
ARX	0,84	5,55E-01	94	168
NKX2-2	0,7	7,04E-01	60	98
PAX6	0,16	9,39E-01	717	804
PAX4	-2,24	NA	2,4	0,6
PDX1	1,02	5,27E-02	456	924
FOXA1	-0,18	9,05E-01	3073	2704
FOXA2	0,71	2,72E-01	617	1008
RFX6	-0,98	2,58E-01	99	50
INSM1	0,51	7,83E-01	590	841
ISL1	0,04	9,89E-01	832	859

**Figure 55. Tableau récapitulatif de l'expression de facteurs de transcription régulant la différenciation des cellules de l'épithélium intestinal**

Le Log2(FC) (Fold Change) représente la variation d'expression entre HIOs CTRL et HIOs NGN3<sup>-/-</sup>. Une variation d'expression est jugée significative si la p-valeur ajustée est inférieure à 0,05. En orange est indiqué un gène pour lequel l'analyse statistique révèle une tendance à l'augmentation de l'expression. NA : expression du gène insuffisante pour l'analyse statistique.

Des facteurs de transcription connus pour être impliqués dans la différenciation entéroendocrine chez la souris ont donc été analysés dans nos HIOs NGN3<sup>-/-</sup> (**Figures 20 et 55**).

Dans un premier temps, la différenciation des lignages sécréteurs (cellules de Paneth, à mucus et entéroendocrines) à partir des cellules souches intestinales requiert l'inhibition de la voie Notch et l'expression sous-jacente de Math1/Atoh1 (Shroyer et al., 2007 ; Yang et al., 2001 ; Jensen et al., 2000), tandis que Hes1 est exprimé dans les progéniteurs du lignage absorbant, les entérocytes (Sikandar et al., 2010 ; Fre et al., 2005). Tous deux possèdent une expression similaire dans les HIOs NGN3<sup>-/-</sup> et CTRL, coïncidant avec l'expression inchangée des marqueurs des entérocytes, des cellules à mucus et des cellules de Paneth (**Figure 52**). La différenciation des cellules souches en progéniteurs des lignées sécrétrices et en progéniteurs des entérocytes ne semble donc pas affectée par l'absence de NGN3 d'après l'analyse des marqueurs de ces types cellulaires et des facteurs de transcription impliqués dans la régulation de leur différenciation.

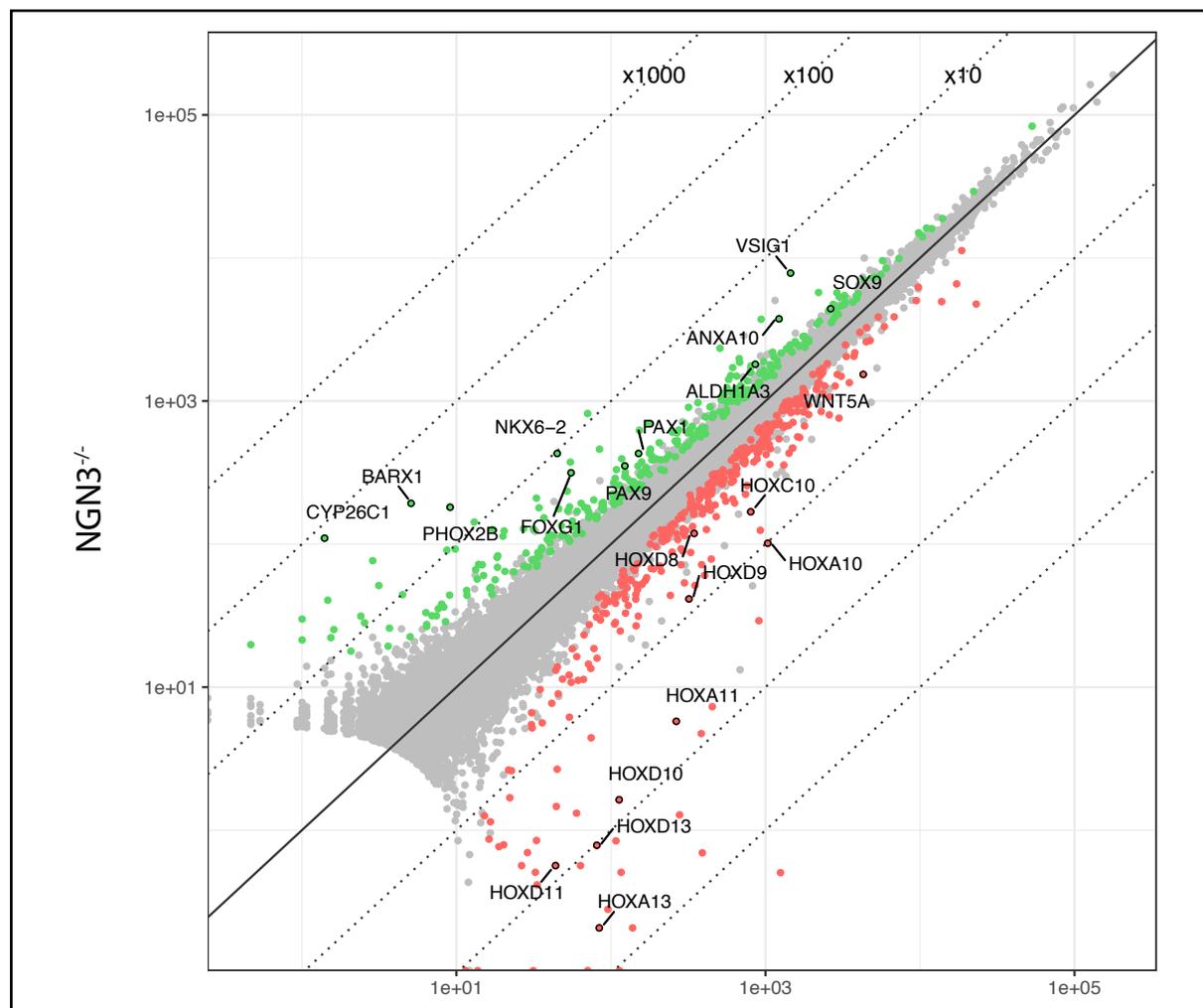
Plusieurs facteurs de transcription agissent en aval de Ngn3 pour réguler la différenciation des sous-types entéroendocrines mais les mécanismes moléculaires précis ne sont pas bien connus.

Une des principales cibles de Ngn3 est Neurod1, dont l'expression est importante dans la différenciation des cellules sécrétrices de Cck et Sct (Rindi et al., 1999). Dans notre modèle d'étude, les gènes codant les hormones CCK et SCT sont peu ou pas exprimés et l'expression de *NEUROD1*, bien que présente, n'est étonnamment pas différentiellement exprimés entre les HIOs *NGN3<sup>-/-</sup>* et CTRL. Les facteurs de transcription *Arx*, *Pdx1*, *Nkx2.2*, *Pax4*, *Insm1* et *Pax6* sont eux aussi impliqués dans la formation des sous-types entéroendocrines au niveau proximal et distal du tractus digestif (Beucher et al., 2012 ; Chen et al., 2009 ; Desai et al., 2008 ; Gierl et al., 2006 ; Larsson et al., 1998). *ARX*, *NKX2.2* et *PAX6* sont exprimés dans les HIOs *NGN3<sup>-/-</sup>* à un niveau similaire aux HIOs CTRL, alors que *PAX4* n'est exprimé dans aucun des HIOs. Concernant *PDX1*, une forte tendance à l'augmentation est observée chez les mutants. Les souris déficientes pour *Pdx1* spécifiquement dans l'intestin présentent une diminution du nombre de transcrits *Sst* dans l'intestin grêle proximal, sans pour autant affecter le nombre de cellules D associées (Chen et al., 2009). Par contre, la diminution des transcrits *Gip* dans l'intestin de ces souris est associée à un nombre de cellules K moindre (Ikeguchi et al., 2018). Dans notre cas, *GIP* n'est pas exprimé mais l'ARNm de *SST* est détecté dans les HIOs et n'est pas différentiellement exprimé chez les mutants. Il ne peut être exclu que l'augmentation de *PDX1* provoquée par l'inactivation de *NGN3* induise une modification de la spécification des lignages entéroendocrines, notamment des cellules K, mais cela reste à déterminer. De plus, son augmentation peut probablement expliquer l'absence d'effet sur *SST*. D'autres facteurs comme *Foxa1* et *Foxa2* ont été mis en évidence chez la souris comme régulateurs de la différenciation des cellules L (*Glp1*, *Glp2*, *Pyy*) et D (*Sst*) dans l'intestin grêle (Ye et al., 2009), mais ne sont pas différentiellement exprimés dans les HIOs *NGN3<sup>-/-</sup>*. Il semblerait donc que la différenciation des cellules D ne soit pas dépendante de *NGN3*, de même que la différenciation des cellules L, bien que ces observations doivent être confirmées. Le rôle de *Insm1* et *Isl1* dans la différenciation des sous-types entéroendocrines en aval de Ngn3 a également été démontré (Terry et al., 2014 ; Gierl et al., 2006), alors que leur expression est inchangée dans nos HIOs mutants par rapport aux HIOs CTRL. Cette absence d'effet de l'inactivation de *NGN3* sur ces facteurs de transcription est vraiment surprenante compte tenu de la littérature existante. D'autant plus qu'une étude très récente de surexpression de *NGN3* dans les HIOs a montré l'augmentation d'expression des facteurs de transcription régulant la différenciation entéroendocrine comme *NKX2.2*, *PAX4*, *RFX6*, *PAX6*, *ARX* et *NEUROD1* (Sinagoga et al., 2018), qui sont des cibles directes de *NGN3* initialement identifiées dans le pancréas (Smith et al., 2010, 2004, 2003 ; Watada et al., 2003, Huang et al., 2000). Là encore ces observations sont en faveur d'un rôle conservé de *NGN3* dans la différenciation entéroendocrine, qui n'est pas mis en évidence dans notre modèle.

De manière intéressante, l'unique régulation transcriptionnelle observée est l'augmentation de l'expression de *PDX1*, qui est plutôt exprimé dans l'antré de l'estomac et la partie proximale de l'intestin grêle (Willet et Mills, 2016). Cette observation peut être mise en relation avec la diminution d'expression du marqueur distal d'entérocytes *MEP1A* observée précédemment (**Figure 52**). Ceci suggère une identité régionale différente entre les HIOs mutants et CTRL qui doit être approfondie et qui pourrait expliquer l'absence du phénotype attendu.

## 2.6. Gènes les plus différenciellement exprimés de manière significative dans les HIOs *NGN3*<sup>-/-</sup>

Après avoir analysé les gènes d'intérêts prédéfinis et afin de comprendre l'absence du phénotype attendu, les gènes les plus différenciellement exprimés ont été passés au crible afin d'identifier ceux potentiellement impliqués dans le développement et la physiologie intestinale (**Figure 56**).



**Figure 56. Gènes les plus différenciellement exprimés dans les HIOs *NGN3*<sup>-/-</sup>**  
 Les gènes les plus sous-exprimés dans les HIOs *NGN3*<sup>-/-</sup> sont indiqués par des points rouges et les plus surexprimés, par des points verts. Le nom de certains gènes d'intérêt est indiqué sur le nuage de point.

### 2.6.1. Gènes les plus sous-exprimés de manière significative dans les HIOs NGN3<sup>-/-</sup>

De manière intéressante, un certain nombre de gènes les plus significativement sous-exprimés appartiennent à la famille de gènes *HOX*, qui jouent un rôle dans le développement du tractus digestif, notamment dans la mise en place de l'axe proximo-distale (**Figure 57**).

C'est le cas des gènes *HOXA10*, *A11*, *A13*, *D8*, *D9*, *D10*, *D11*, *D13* et *C10*, dont une partie est décrite en **figure 7**, connus pour leur implication dans les interactions entre l'endoderme et le mésoderme au cours du développement. Leur expression précoce dans le mésoderme est nécessaire à la morphogenèse et à la régionalisation intestinale. Par la suite, ces interactions endoderme-mésoderme sont importantes pour la différenciation épithéliale (Le Guen et al., 2015). Il a été mis en évidence chez la souris et le poulet que l'ensemble des gènes *Hox* a un patron d'expression très spécifique le long de l'axe proximo-distal de l'intestin au cours du développement. On ne peut cependant affirmer avec certitude dans quelle région précise sont exprimés ces gènes chez l'homme puisqu'une variabilité inter-espèce importante a été rapportée (Yuasa et al., 2003 ; Beck et al., 2000).

Dans notre cas, les gènes *HOX* sous-exprimés semblent plutôt correspondre à ceux retrouvés dans la partie distale du tube digestif (gros intestin, côlon, rectum). En effet, les *HOXA10*, *A11*, *A13*, *D10*, *D11* et *D13* sont utilisés dans l'identification de l'intestin distal et notamment pour déterminer la nature distale des HIOs (Munera et al., 2017). De plus, *HoxD8* et *HoxD9* sont exprimés dans le caecum et le côlon chez la souris (Kawazoe et al., 2002).

Gènes	Log2 FC	P-valeur ajustée	Moyenne CTRL	Moyenne NGN3 <sup>-/-</sup>
<b>HOXA13</b>	<b>-8,13</b>	<b>4,99E-10</b>	84	0,2
<b>HOXD13</b>	<b>-6,77</b>	<b>5,57E-08</b>	81	0,8
<b>HOXD11</b>	<b>-6,46</b>	<b>8,45E-06</b>	44	0,6
<b>HOXD10</b>	<b>-6,24</b>	<b>3,41E-15</b>	113	1,6
<b>HOXA11</b>	<b>-5,5</b>	<b>7,64E-27</b>	265	5,8
<b>HOXA10</b>	<b>-3,36</b>	<b>2,36E-43</b>	1035	101
<b>HOXD9</b>	<b>-2,96</b>	<b>1,36E-24</b>	320	41
<b>HOXC10</b>	<b>-2,25</b>	<b>1,53E-11</b>	803	168
<b>HOXD8</b>	<b>-1,54</b>	<b>1,72E-11</b>	346	119
<b>WNT5A</b>	<b>-1,48</b>	<b>5,08E-12</b>	4292	1537

**Figure 57. Tableau présentant des gènes significativement sous-exprimés dans les HIOs NGN3<sup>-/-</sup>.**

Le Log2(FC) (Fold Change) représente la variation d'expression entre HIOs CTRL et HIOs NGN3<sup>-/-</sup>. Une variation d'expression est jugée significative si la p-valeur ajustée est inférieure à 0,05.

Par contre, le rôle de *HoxC10* n'est pas bien décrit dans la littérature. La diminution de l'expression de ces gènes de spécification de la régionalisation distale du tube digestif semble donc indiquer une nature plutôt proximale des HIOs *NGN3*<sup>-/-</sup> par rapport aux HIOs CTRL, hypothèse déjà émise plus tôt puisque *PDX1* (marqueur proximal) est surexprimé et *MEP1A* (marqueur distal d'entérocytes) est sous-exprimé chez les mutants. Pourtant, l'expression de *Ngn3* est restreinte aux progéniteurs entéroendocrines chez la souris et de ce fait, aucune altération du développement et de la régionalisation intestinale n'est attendue en l'absence de *NGN3*.

*WNT5A*, un ligand de la famille des Wnt, fait également partie des gènes les plus significativement surexprimés et est essentiel dans le développement et l'élongation de l'intestin grêle chez la souris. Effectivement, l'intestin grêle des souris déficientes pour *Wnt5a* est dramatiquement raccourci (Cervantes et al., 2009). Il a également été montré qu'il favorise la régénération des cryptes dans le côlon par l'intermédiaire de la signalisation TGF- $\beta$  après dommage (Miyoshi et al., 2012). Ceci suggère, là aussi, que les HIOs *NGN3*<sup>-/-</sup> formés semblent de nature plus proximale en comparaison aux HIOs CTRL. Afin de confirmer cette hypothèse, l'expression de marqueurs des régions proximales a été analysée et est présentée dans la partie suivante.

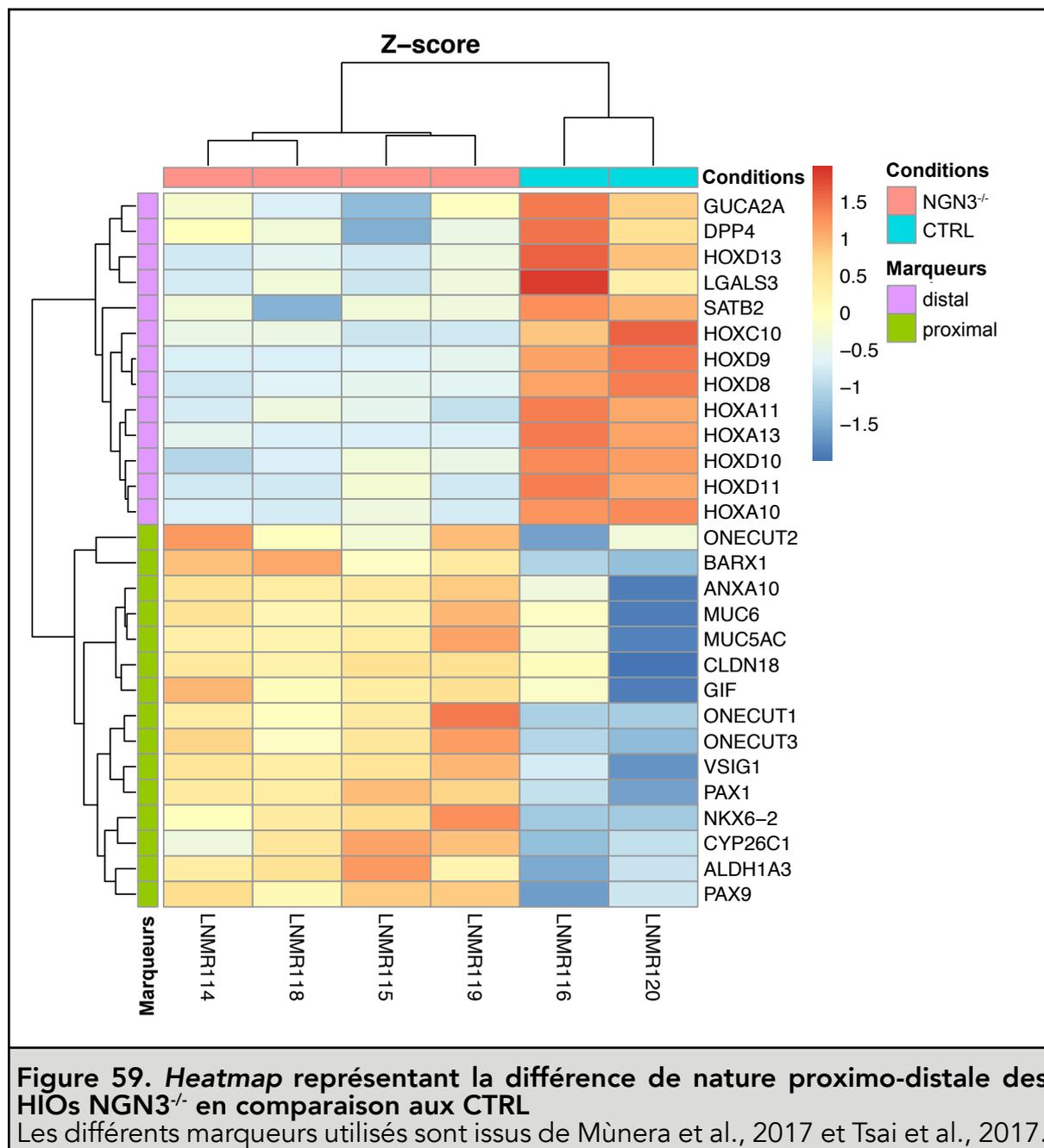
### 2.6.2. Etude de la nature proximale des HIOs

L'expression de plusieurs marqueurs de l'intestin proximal est listée dans la **figure 58**. La cathepsine E, codée par le gène *CTSE*, est un marqueur de différenciation gastrique (Konno-Shimizu et al., 2013) et s'avère être surexprimé dans les HIOs *NGN3*<sup>-/-</sup>. C'est également le cas de *GATA4*, requis pour la différenciation de l'épithélium gastrique (Jacobsen et al., 2002).

	Gènes	Log2 FC	P-valeur ajustée	Moyenne CTRL	Moyenne <i>NGN3</i> <sup>-/-</sup>
Marqueurs proximaux	<i>CTSE</i>	1,99	2,23E-03	938	3734
	<i>GATA4</i>	1,34	9,32E-04	596	1508
	<i>TFF1</i>	1,38	8,26E-03	2206	5747
	<i>MUC5AC</i>	1,98	2,74E-01	725	2867
	<i>MUC6</i>	1,71	3,98E-01	85	278
	<i>CAPN8</i>	1,05	8,93E-02	1015	2099
	<i>TFF2</i>	2,13	1,25E-01	1153	5056

**Figure 58. Tableau représentant le niveau d'expression de certains marqueurs proximaux du tube digestif**

Le Log2(FC) (Fold Change) représente la variation d'expression entre HIOs CTRL et HIOs *NGN3*<sup>-/-</sup>. Une variation d'expression est jugée significative si la p-valeur ajustée est inférieure à 0,05. En bleu sont indiqués les gènes significativement surexprimés dans les HIOs *NGN3*<sup>-/-</sup>.



Afin de compléter l'analyse, des marqueurs des différents types cellulaires de l'épithélium gastrique, qui sont notamment utilisés dans la validation de la différenciation *in vitro* d'organoïdes gastriques, ont été étudiés (McCracken et al., 2014). L'expression du marqueur de cellules à mucus *TFF1* est significativement augmentée chez les mutants et les autres marqueurs de ces cellules, *MUC5AC* et *MUC6*, présentent également une augmentation du niveau d'expression, bien que non significative. *CAPN8* est aussi un marqueur des cellules à mucus de l'estomac (Engelstoft et al., 2015) et présente une augmentation presque significative de son expression dans les HIOs NGN3<sup>-/-</sup>. Le marqueur de cellules glandulaires de l'antra, *TFF2*, est également surexprimé mais de manière non-significative. Manifestement, tous les marqueurs analysés sont surexprimés chez les mutants même si pour certains,

la différence d'expression n'est pas significative entre les HIOs NGN3<sup>-/-</sup> et CTRL. Ces observations sont concomitantes avec l'augmentation de *PDX1* mise en évidence précédemment, puisqu'il est notamment exprimé dans l'antra de l'estomac et la partie proximale de l'intestin (Willet et Mills, 2016).

Une analyse plus globale de marqueurs proximaux et distaux, représentée sous forme d'une *heatmap* (**figure 59**) confirme assez clairement la nature proximale des HIOs NGN3<sup>-/-</sup> en comparaison aux HIOs CTRL.

### 2.6.3. Perturbation de la régulation de la prolifération cellulaire dans les HIOs NGN3<sup>-/-</sup>

Un certain nombre de gènes sous-exprimés dans les HIOs NGN3<sup>-/-</sup> ont un rôle connu dans la prolifération des cellules de l'épithélium intestinal (**Figure 60**). Par exemple, *NKX2.3* est exprimé dans le mésenchyme de l'intestin grêle mais sa délétion chez la souris provoque une hyperprolifération des cellules dans les cryptes intestinales et une formation altérée des villosités (Pabst et al., 1999) et *SOCS3* est un suppresseur de tumeur qui limite la prolifération des cellules épithéliales intestinales (Shaw et al., 2017). Il a également été mis en évidence dans des entéroïdes murins que *Myb* est impliqué dans la croissance, la prolifération et la différenciation cellulaire de l'épithélium intestinal (Germann et al., 2014). *FOXF1* participe aussi à ces processus. En effet, les souris déficientes pour *FoxF1* présentent un axe crypte-villosité perturbé, caractérisé par une prolifération excessive (Madison et al., 2009). Le long ARN non codant *FENDRR* associé à *FOXF1* (*FOXF1 adjacent non-coding developmental regulatory RNA*) est également lié à la prolifération et l'invasion cellulaire dans des cancers gastriques (Luo et al., 2018).

Gènes	Log2 FC	P-valeur ajustée	Moyenne CTRL	Moyenne NGN3 <sup>-/-</sup>
<b>NKX2.3</b>	<b>-2,42</b>	5,89E-06	387	72
<b>SOCS3</b>	<b>-1,69</b>	9,01E-05	2590	802
<b>MYB</b>	<b>-1,49</b>	1,42E-03	188	67
<b>FOXF1</b>	<b>-1,28</b>	9,13E-03	2607	1073
<b>FENDRR</b>	<b>-1,88</b>	5,25E-03	1315	357

**Figure 60. Tableau représentant l'expression de gènes impliqués dans la régulation de la prolifération cellulaire de l'épithélium intestinal**

Le Log2(FC) (Fold Change) représente la variation d'expression entre HIOs CTRL et HIOs NGN3<sup>-/-</sup>. Une variation d'expression est jugée significative si la p-valeur ajustée est inférieure à 0,05. En rouge sont indiqués les gènes significativement sous-exprimés dans les HIOs NGN3<sup>-/-</sup>.

Ces observations sont d'autant plus intéressantes puisque les souris dans lesquelles *Ngn3* a été spécifiquement inactivé dans l'intestin, possèdent des cryptes élargies dues à une prolifération aberrante mais également des villosités plus courtes à cause d'une augmentation du renouvellement cellulaire (Mellitzer et al., 2010). Les HIOs *NGN3*<sup>-/-</sup> présentent une diminution de l'expression de gènes impliqués dans la régulation de la prolifération cellulaire. En effet, les modèles murins déficients pour ces gènes présentent un phénotype avec une altération de la prolifération cellulaire comparable à ce qui est observé chez les souris *Ngn3*<sup>Δint</sup>. *Ngn3* pourrait donc avoir un rôle dans la morphogenèse de l'épithélium intestinal, comme c'est le cas lors du développement pancréatique (Magenheim et al., 2011), et qui serait potentiellement conservé dans le développement intestinal chez l'homme, d'après nos observations dans les HIOs *NGN3*<sup>-/-</sup>. Il se pourrait que l'altération de la prolifération soit **indirectement** due aux conséquences de la perte de *NGN3*.

#### 2.6.4. Gènes les plus surexprimés de manière significative dans les HIOs *NGN3*<sup>-/-</sup>

Les gènes les plus différentiellement surexprimés ont également été passés au crible afin d'identifier ceux potentiellement impliqués dans le développement et la physiologie intestinale. La **figure 61** présente une liste de gènes parmi ceux les plus surexprimés dans les HIOs *NGN3*<sup>-/-</sup> de façon significative.

On retrouve parmi eux *BARX1*, qui est exprimé dans le mésenchyme gastrique et est essentiel dans le développement de l'estomac à partir de l'endoderme intestinal. En outre, une approche permettant la différenciation des cellules souches embryonnaires en tissu gastrique en induisant l'expression de *BARX1* a été développée (Noguchi et al., 2015). D'autres gènes exprimés dans les régions proximales du tube digestif sont surexprimés, comme *NKX6.2* qui est exprimé dans le mésenchyme du duodénum (Nelson et al., 2005) et *VSIG1* qui code une protéine d'adhésion requise pour le développement de la partie glandulaire de l'estomac (Oidovsambuu et al., 2011). *ANXA10* est un marqueur de la muqueuse gastrique et des glandes de Brunner localisées dans la sous-muqueuse du duodénum (Lu et al., 2011). De plus, *PAX1* et *PAX9* sont également impliqués dans le développement de l'intestin primitif antérieur (Muller et al., 1996). Par ailleurs, l'acide rétinoïque (AR) joue un rôle crucial dans le développement des organes dérivés de l'intestin primitif antérieur et il est utilisé dans le protocole de différenciation des organoïdes gastriques dans l'étape de régionalisation postérieure des sphéroïdes (McCracken et al., 2014). Sa distribution est étroitement contrôlée par les gènes codant les synthétases *ALDH1As* et les enzymes métaboliques *CYP26s* (Xi et al., 2015). Ces gènes définissent donc l'exposition tissulaire à l'acide rétinoïque et sont essentiels pour la morphogenèse des organes. Parmi ceux-ci, les gènes *ALDH1A3* et *CYP26C1* sont surexprimés dans les HIOs *NGN3*<sup>-/-</sup>.

Gènes	Log2 FC	P-valeur ajustée	Moyenne CTRL	Moyenne NGN3 <sup>-/-</sup>
CYP26C1	6,26	4,87E-03	1,4	110
BARX1	5,25	5,86E-05	5,1	193
PHOX2B	4,32	2,36E-07	9,1	181
NKX6.2	3,26	6,71E-08	45	430
FOXG1	2,51	3,02E-09	55	316
VSIG1	2,43	5,96E-09	1453	7851
ANXA10	1,62	2,88E-03	1223	3759
PAX1	1,51	2,81E-07	151	430
PAX9	1,51	8,81E-06	123	352
ALDH1A3	1,07	1,93E-04	859	1809
SOX9	0,74	6,28E-03	2642	4426

**Figure 61. Tableau présentant des gènes significativement surexprimés dans les HIOs NGN3<sup>-/-</sup>.**

Le Log2(FC) (Fold Change) représente la variation d'expression entre HIOs CTRL et HIOs NGN3<sup>-/-</sup>. Une variation d'expression est jugée significative si la p-valeur ajustée est inférieure à 0,05.

Ces observations confirment donc encore une fois notre hypothèse du caractère plus proximal des HIOs NGN3<sup>-/-</sup> par rapport aux HIOs CTRL.

SOX9, lui, est corrélé à la voie de signalisation Wnt et joue un rôle important dans les cryptes intestinales en régulant notamment la prolifération cellulaire (Bastide et al., 2007) ce qui renvoie une fois de plus à l'hypothèse de dérégulation de la prolifération cellulaire vue dans la partie précédente (**voir 2.6.3**).

De manière inattendue, le gène *PHOX2B* qui est exprimé dans les progéniteurs du système nerveux entérique (ENS) (Buehler et al., 2014) et *FOXG1*, également impliqué dans la différenciation neuronale (Pancrazi et al., 2015), font partie des gènes les plus surexprimés dans les HIOs NGN3<sup>-/-</sup>. Il est alors envisageable que du tissu nerveux soit formé dans nos HIOs et que la différenciation neuronale soit exacerbée chez les mutants NGN3<sup>-/-</sup> pour une raison inconnue.

## 2.7. Conclusions et points de discussion relevés

En conclusion de l'analyse des HIOs *NGN3*<sup>-/-</sup>, il peut être affirmé de manière certaine que les clones hiPSCs 2 et 10 sont bien déficients pour *NGN3*, comme en atteste le génotypage des hiPSCs par PCR et par séquençage des produits PCR et les différentes analyses sur les HIOs dérivés de ces clones : RT-qPCR, immunofluorescences, ainsi que RNA-seq. Le clone 49Cd, par contre, semble avoir une activité résiduelle de *NGN3* et n'est donc pas considéré comme déficient pour *NGN3* et n'a pas été pris en compte dans l'analyse de l'expérience de RNA-seq. Il faudra établir une caractérisation plus fine pour déterminer s'il pourrait être incorporé dans l'analyse en tant que clone hétérozygote possédant un allèle fonctionnel et un allèle tronqué.

L'analyse par RT-qPCR montre une diminution significative de *LYZ* et *MUC2* dans les HIOs *NGN3*<sup>-/-</sup>, respectivement marqueur des cellules de Paneth et cellules à mucus, alors qu'aucune différence n'est observée dans les résultats de RNA-seq. Ceci peut s'expliquer par l'effectif réduit des *NGN3*<sup>-/-</sup> (n=3) dans le cadre des RT-qPCR. Un seul de ces échantillons a été analysé par RNA-seq, ainsi que des échantillons supplémentaires issus d'une différenciation indépendante. L'ajout d'échantillons permet d'amoindrir la variabilité et améliore donc la fidélité de l'analyse. Cette observation n'est donc pas très fiable et il faudrait réaliser des immunofluorescences de ces marqueurs sur les HIOs pour déterminer si un nombre plus faible de cellules de Paneth et cellules à mucus est effectivement recensé.

L'inactivation de *NGN3* ne provoque pas le phénotype attendu, c'est-à-dire la perte de toutes les cellules entéroendocrines observée chez la souris (Mellitzer et al., 2010 ; Jenny et al., 2002), dans le modèle de HIOs dans lequel l'expression de *NGN3* a été altérée via des shRNAs (Spence et al., 2011) et chez les patients porteurs de mutations ponctuelles dans *NGN3* (Wang et al., 2006). L'absence d'expression des hormones produites par les différents types cellulaires, ainsi que la dérégulation transcriptionnelle de la différenciation entéroendocrine mises en évidence dans les modèles murins, n'ont pas été retrouvées dans les HIOs *NGN3*<sup>-/-</sup>. Il y a bien une forte diminution des marqueurs de cellules entéroendocrines généraux *CHGA* et *SCG2*, mais à cause de la faible expression voire de l'absence d'expression de la plupart des hormones, les sous-types entéroendocrines affectés n'ont pas pu être mis en évidence.

Toutefois, il semblerait que la différenciation des cellules D sécrétrices de SST ne soit pas affectée puisque le gène codant cette hormone est bien exprimé et de manière similaire dans les HIOs *NGN3*<sup>-/-</sup> et CTRL, ainsi que le facteur de transcription *FOXA1*, impliqué notamment dans la différenciation de ces cellules. L'augmentation de *PDX1* pourrait expliquer l'expression inchangée de SST.

A l'inverse, il se pourrait que la différenciation des cellules entérochromaffines (EC) soit altérée puisqu'il y a une diminution significative de l'expression de *CHGA* et une absence de cellules marquées *CHGA*<sup>+</sup>. Or, il a été montré que *CHGA* n'est pas forcément un marqueur générique de cellules entéroendocrines mais plutôt un marqueur de cellules entérochromaffines (Engelstoft et al., 2015). De plus, une diminution significative de *VGF*, gène connu pour être enrichi dans les cellules entérochromaffines de l'intestin grêle et du colon (Lund et al., 2018) est observée. *TPH1* code l'enzyme de biosynthèse de la sérotonine, qui est produite par les cellules entérochromaffines, et présente une baisse d'expression dans les RT-qPCR. Cette observation n'a pas été confirmée par l'approche de RNA-seq mais le niveau d'expression très faible de ce gène pourrait expliquer l'absence de différence entre les HIOs *NGN3*<sup>-/-</sup> et CTRL.

Les groupes de gènes significativement sous-exprimés sont majoritairement des marqueurs des régions distales du tractus gastro-intestinal et des gènes impliqués dans des processus de régulation de la prolifération cellulaire. Les gènes significativement surexprimés sont impliqués dans le développement des régions proximales ou participent à des régulations développementales de l'estomac et des parties proximales de l'intestin. Nos observations quant à la régulation de la prolifération cellulaire coïncident avec la réduction des villosités et l'élargissement du compartiment des cryptes dû à une hyperprolifération et un renouvellement aberrant des cellules dans l'épithélium intestinal des souris *Ngn3*<sup>Δint</sup> (Mellitzer et al., 2010).

Les analyses des données de RNA-seq sont assez probantes concernant le fait que les HIOs *NGN3*<sup>-/-</sup> étudiés soient constitués de tissu intestinal de nature plus proximale que les HIOs CTRL. En effet, les HIOs *NGN3*<sup>-/-</sup> possèdent un niveau d'expression plus élevé de marqueurs d'intestin proximal et un niveau d'expression moindre des marqueurs distaux. Ceci pourrait s'expliquer par l'absence de *NGN3* et donc un effet propre au génotype, qui favoriserait la formation d'intestin proximal. Un rôle de *Ngn3* dans le développement de l'estomac est envisageable, puisque des souriceaux *Ngn3*<sup>-/-</sup> présentent un estomac de taille réduite et une muqueuse plus épaisse dans la partie proximale de l'intestin (Jenny et al., 2002; Lee et al., 2002). Cependant, il est également envisageable que ces observations soient indépendantes de la mutation et provoquées par une variation de la durée d'exposition des hiPSCs au cocktail FGF4 + CHIR99021 durant la formation de l'intestin moyen au cours de la différenciation intestinale *in vitro*. En effet, il a été montré que plus cette exposition est longue, plus les HIOs adoptent une nature proximale (Tsai et al., 2016).

La nature proximale du tissu généré dans les HIOs NGN3<sup>-/-</sup> pourrait expliquer l'absence du phénotype sévère attendu. En effet, Ngn3 n'est pas indispensable à la différenciation de tous les sous-types endocrines dans l'épithélium gastrique, contrairement à l'épithélium de l'intestin grêle. La délétion de *Ngn3* empêche la formation des cellules G sécrétrices de gastrine et D sécrétrices de somatostatine dans l'estomac chez la souris, alors que les cellules à sérotonine et ghréline sont toujours formées (Jenny et al., 2002 ; Lee et al., 2002). L'hétérogénéité dans les tissus formés et le caractère proximal des HIOs NGN3<sup>-/-</sup> pourraient masquer d'éventuels effets de l'inactivation de *NGN3*. Des immunofluorescences sur des marqueurs proximaux et distaux pourraient confirmer la nature des tissus générés selon le génotype.

D'autre part, les HIOs NGN3<sup>-/-</sup> ont été générés avec des clones d'un nombre de passage plus élevé (P60) que les HIOs CTRL (P30). Aucune donnée ne permettent d'affirmer que la qualité de la différenciation dépend du nombre de passage des cellules de départ mais c'est une hypothèse non négligeable puisque l'accumulation des passages peut provoquer des altérations du caryotype des cellules et en affecter leur pluripotence. De plus, contrairement aux hiPSCs utilisées dans la différenciation des HIOs CTRL, les clones NGN3<sup>-/-</sup> sont des cellules nucléofectées qui peuvent être affectées par cette manipulation, notamment par la possibilité d'introduction de mutations aléatoires qui peuvent également provenir de coupure non-ciblée dans l'ADN via le système CRISPR/Cas9. De nouveaux clones NGN3<sup>-/-</sup>, ainsi que des contrôles isotypiques ayant également subi la nucléofection, ont été récemment générés avec des hiPSCs d'un nombre de passage peu élevé (P20) pour pallier à ce problème.

La potentielle présence de tissus neuronaux ou autres tissus non-intestinaux, qui est accrue dans les HIOs NGN3<sup>-/-</sup>, pourrait également influencer la pertinence des analyses en masquant les effets propres au tissu intestinal. Effectivement, bien que non décrit dans la littérature, il est possible que certains expérimentateurs obtiennent une différenciation importante de tissus nerveux qui expriment également des marqueurs endocrines. Des immunofluorescences sur des marqueurs neuronaux pourraient être envisagées pour déterminer la présence de tissu nerveux. Enfin, il a été observé chez la souris que le fond génétique utilisé pouvait influencer sur le phénotype provoqué par certaines mutations génétiques induites. De ce fait, l'absence du phénotype attendu suite à la délétion de *NGN3* pourrait être propre à notre lignée de cellules hiPS utilisée. Il faudrait donc réaliser les mêmes expériences avec une autre lignée de cellules hiPS pour vérifier cette hypothèse.

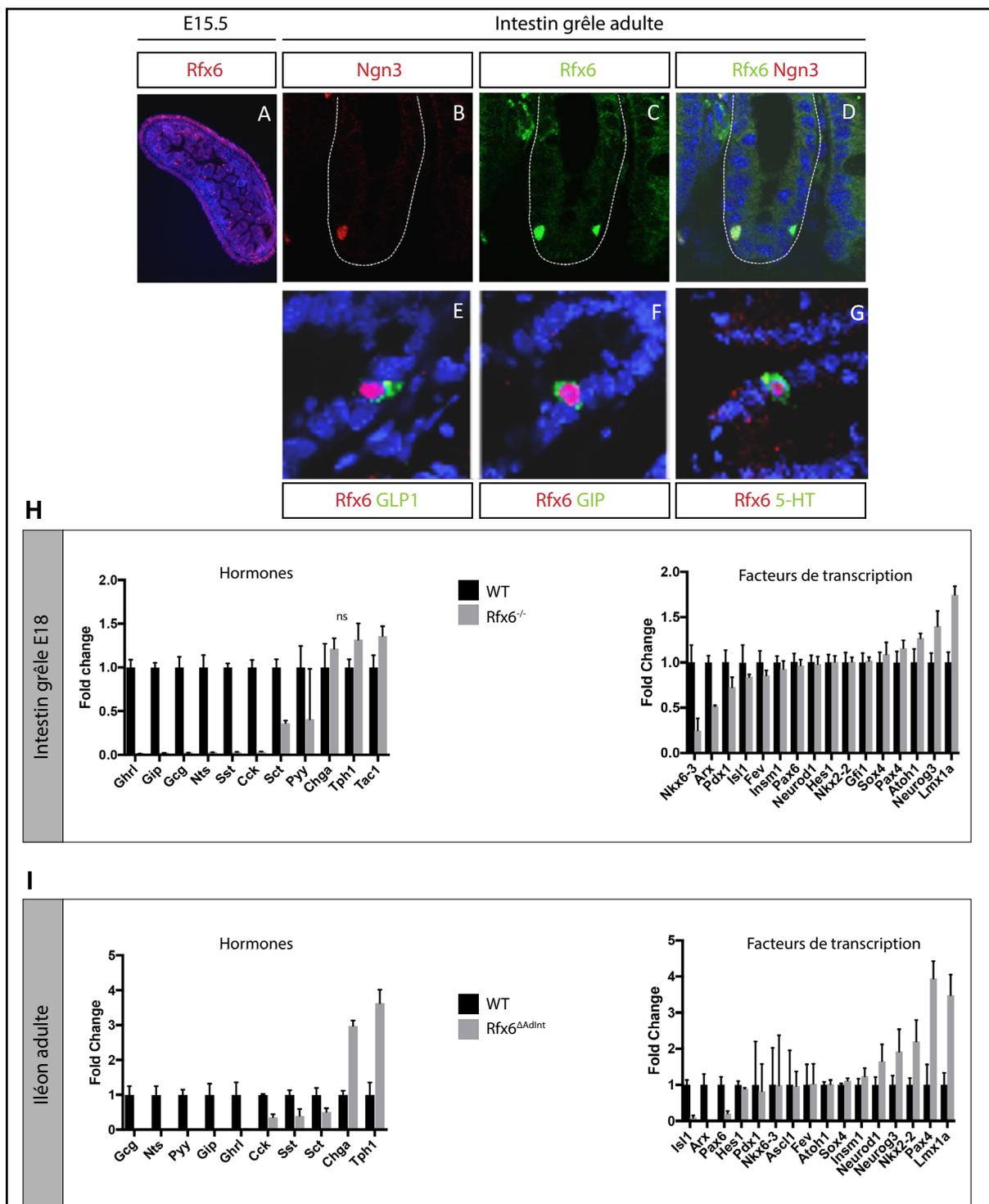
## Partie 4. Etude des conséquences de l'inactivation de *RFX6* sur la différenciation entéroendocrine chez l'homme

En parallèle de l'étude des HIOs *NGN3<sup>-/-</sup>*, les HIOs *RFX6<sup>-/-</sup>* ont également été analysés de manière similaire, à savoir par RT-qPCR et immunofluorescences sur des marqueurs spécifiques de l'épithélium intestinal, afin de valider le caractère intestinal des tissus générés et confirmer le génotype des clones. Puis, une expérience de RNA-seq a été réalisée pour identifier les gènes différentiellement exprimés dans les HIOs *RFX6<sup>-/-</sup>* par rapport aux HIOs CTRL. Des données similaires issues d'expériences sur des embryons de souris au stade embryonnaire E18 déficients pour *Rfx6* (*Rfx6<sup>-/-</sup>* E18) et sur des souris adultes déficientes pour *Rfx6* spécifiquement dans l'intestin (*Rfx6<sup>ΔAdInt</sup>*) ont été générées par l'équipe (Julie Piccand, Anthony Beucher, Aline Meunier et Adèle de Arcangelis). Ces données font l'objet d'un manuscrit qui est actuellement en préparation. L'ensemble des résultats d'analyses des HIOs *RFX6<sup>-/-</sup>* sera comparé à ces données issues de modèles murins. Considérant l'absence de données concernant le rôle de *Rfx6* dans l'intestin dans la littérature, des données issues du manuscrit en préparation sont présentées dans un premier temps afin d'introduire les données obtenues dans notre modèle *in vitro*.

### 1. Etudes des modèles murins déficients pour *Rfx6* (Piccand et al., manuscrit en préparation)

*Rfx6* est d'abord exprimé pendant le développement embryonnaire dans l'endoderme de l'intestin primitif postérieur et son expression devient peu à peu restreinte à l'estomac, l'endoderme intestinal et le pancréas. Dans le pancréas en développement, il est exprimé dans le lignage endocrine où il co-localise avec une population de cellules *Ngn3<sup>+</sup>*, puis son expression est maintenue dans les cellules endocrines adultes (Piccand et al., 2014 ; Smith et al., 2010 ; Soyer et al., 2010). Un patron d'expression similaire est observé dans l'intestin puisque l'expression de *Rfx6* apparaît sporadique au sein de l'intestin de souris au stade embryonnaire E15 (**Figure 62A**) et que l'on retrouve à la fois des cellules doublement marquées pour *Rfx6* et *Ngn3* dans les cryptes (**Figure 62B-D**), ainsi qu'une co-localisation de *Rfx6* avec les hormones entéroendocrines *Glp1*, *Gip* et 5-HT (sérotonine) dans l'intestin de souris adulte (**Figure 62E-G**).

Les expériences de RNA-seq sur les embryons *Rfx6<sup>-/-</sup>* à E18 ont mis en évidence une absence d'expression de la majorité des gènes codant les hormones ou enzyme de biosynthèse de celles-ci (**Figure 62H**).



**Figure 62. Etude de l'expression de Rfx6 et des conséquences de son inactivation sur la différenciation entéroendocrine chez la souris**

Données acquises par J. Piccand, A. Beucher, A. Meunier et A. De Arcangelis.

(A-G) Expression de Rfx6 par immunofluorescence sur des cryosections d'intestin grêle de souris : (A) Rfx6 (rouge) dans l'intestin de souris à E15.5, (B-D) Co-localisation de Rfx6 (vert) avec Ngn3 (rouge) dans des cryptes intestinales de souris adulte, (E-G) Co-localisation de Rfx6 (rouge) avec les hormones Glp1, Gip et 5-HT (vert) dans l'intestin grêle de souris adulte. (H-I) Variation de l'expression des gènes (Fold Change) codant les hormones/enzymes de biosynthèse entéroendocrines et des facteurs de transcription régulant la différenciation épithéliale entre les souris déficientes pour Rfx6 et les souris contrôles (WT): (H) Rfx6<sup>-/-</sup> E18, (I) Rfx6<sup>ΔAdInt</sup>.

Une diminution de l'expression de *Sct* est observée, ainsi qu'une baisse non significative de *Pyy*. De manière surprenante, seule l'expression de *ChgA* ne change pas par rapport aux souris contrôles et une augmentation de *Tph1* (enzyme de biosynthèse de la sérotonine) et de *Tac1* (codant notamment la substance P) est observée chez les mutants (**Figure 62H**). Concernant les souris *Rfx6*<sup>ΔAdInt</sup>, les analyses ont été faites 11 semaines après l'injection de tamoxifène qui permet l'induction de la délétion de *Rfx6* spécifiquement dans l'intestin chez la souris adulte. Un phénotype similaire aux embryons de souris *Rfx6*<sup>-/-</sup> a été mis en exergue chez les souris adultes *Rfx6*<sup>ΔAdInt</sup> puisque l'expression de la majorité des gènes codant les hormones est absente ou drastiquement diminuée et l'expression de *Tph1* est augmentée. Cette fois, l'expression de *ChgA* est également augmentée de manière significative chez les mutants (**Figure 62I**).

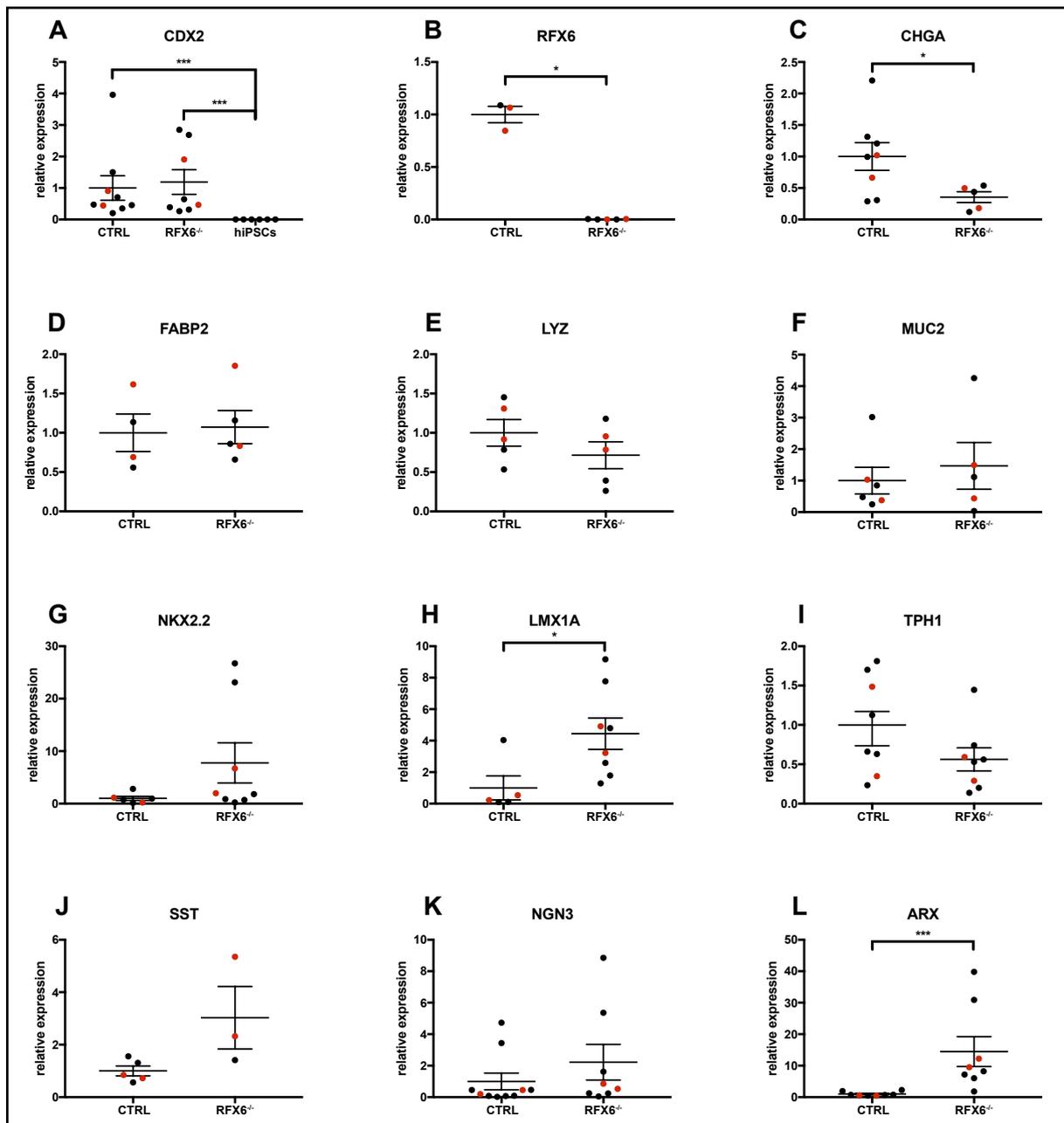
L'expression des facteurs de transcription régulant la différenciation des cellules épithéliales intestinales et notamment la différenciation entéroendocrine a été étudiée. Chez les embryons *Rfx6*<sup>-/-</sup>, la majorité de ces facteurs de transcription ont une expression non significativement différente (**Figure 62H**). *Nkx6.3* (impliqué dans la différenciation des cellules G sécrétrices de gastrine ; May et Kaestner, 2010), *Arx* (impliqué dans la différenciation des cellules L sécrétrices de Glp1 ; Beucher et al., 2012) et en moindre mesure *Pdx1* (impliqué dans la différenciation des cellules G et des cellules K sécrétrices de Gip ; Ikeguchi et al., 2018), ont un niveau d'expression moindre chez les mutants et au contraire, *Lmx1a* (impliqué dans la régulation de la biosynthèse de sérotonine ; Gross et al., 2016) et *Ngn3* sont augmentés (**Figure 62H**). Chez les souris adultes *Rfx6*<sup>ΔAdInt</sup>, une baisse de l'expression de *Isl1* (impliqué dans la différenciation des cellules à Glp1, Gip et Cck ; Terry et al., 2014), *Pax6* (impliqué dans la différenciation des cellules à Glp1 et Glp2 ; Hill et al., 1999) et *Arx* est observée (**Figure 62I**). Comme chez les embryons, *Lmx1a* est augmenté, ainsi que *Ngn3*. De plus, une augmentation de *NeuroD1* (impliqué dans la différenciation des cellules sécrétrices de *Sct* et *Cck* ; Naya et al., 1997), *Nkx2.2* (impliqué dans la régulation de la biosynthèse de sérotonine ; Gross et al., 2016) et *Pax4* (impliqué dans la différenciation des cellules D sécrétrices de somatostatine) est provoquée par l'absence de *Rfx6* chez les souris adultes (**Figure 62I**). En résumé, ces données indiquent un rôle essentiel de *Rfx6* dans la différenciation entéroendocrine embryonnaire et adulte chez la souris. Il semblerait qu'il agisse en aval de *Ngn3* et en amont de *Arx* et que la différenciation des cellules à sérotonine soit indépendante de *Rfx6*. Les processus de régulation de la différenciation entéroendocrine semblent relativement conservés entre l'homme et la souris et un rôle similaire de *Rfx6* est donc attendu dans notre modèle humain *in vitro*.

## 2. Analyse par RT-qPCR et immunofluorescences

### 2.1. RT-qPCR

Les HIOs différenciés à partir des clones RFX6<sup>-/-</sup> 1, 14 et 18 ont été récoltés au stade J28 de plusieurs expériences de différenciation indépendantes et l'ARN en a été extrait afin d'analyser les différents HIOs par RT-qPCR (**Figure 63**). Pour évaluer le caractère épithélial et intestinal des tissus générés, l'expression du marqueur *CDX2* a été étudiée. Il n'y a pas de différence significative entre l'ARNm de *CDX2* des HIOs RFX6<sup>-/-</sup> et CTRL (test de Mann-Whitney U=35, p=0,96) et l'expression de *CDX2* semble être caractéristique des HIOs puisqu'elle est absente dans les cellules indifférenciées (hiPSCs) (CTRL vs hiPSCs : U=0, p=0,0004 ; RFX6<sup>-/-</sup> vs hiPSCs : U=0, p=0,0007) (**Figure 63A**), comme décrit dans l'analyse des HIOs NGN3<sup>-/-</sup> (**Figure 48A**). L'inactivation de *RFX6* a été confirmée puisque son expression est absente dans les HIOs RFX6<sup>-/-</sup> (RFX6<sup>-/-</sup> vs CTRL : U=0, p=0,036) (**Figure 63B**). De plus, une baisse significative de *CHGA* est retrouvée dans les HIOs RFX6<sup>-/-</sup> (RFX6<sup>-/-</sup> vs CTRL : U=6, p=0,045), suggérant une diminution du nombre de cellules entéroendocrines (**Figure 63C**). Chez la souris adulte déficiente pour *Rfx6* spécifiquement dans l'intestin (Rfx6<sup>ΔAdInt</sup>), une augmentation de *ChgA* associée à une augmentation de *Tph1* est observée (**Figure 62I**) (Piccand et al., manuscrit en préparation). A l'inverse, lors d'études préliminaires dans un modèle *in vitro* d'entéroïdes dérivés des cryptes intestinales de ces souris, établi par A. De Arcangelis au laboratoire, une baisse de *ChgA* associée à une baisse de *Tph1* a été mise en évidence. Il est donc probable que les différences observées dans la variation de l'expression de *ChgA* entre les différents modèles soient liées au caractère *in vitro* ou *in vivo* des expériences. L'expression de *ChgA* semble systématiquement varier dans le même sens que celle de *Tph1*, suggérant que *ChgA* marque préférentiellement les cellules à sérotonine (EC).

La présence des autres sous-types cellulaires de l'épithélium intestinal dans nos HIOs a été déterminée grâce à l'expression de marqueurs spécifiques. L'expression de *FABP2* (entérocytes) (RFX6<sup>-/-</sup> vs CTRL : U=8, p=0,73) (**Figure 63D**) de *LYZ* (cellules de Paneth) (RFX6<sup>-/-</sup> vs CTRL : U=7,5, p=0,34) (**Figure 63E**) et de *MUC2* (cellules à mucus) (RFX6<sup>-/-</sup> vs CTRL : U=12, p=0,66) (**Figure 63F**) ne semblent pas modifiées dans les HIOs RFX6<sup>-/-</sup>. En accord avec les observations faites lors de l'analyse des HIOs NGN3<sup>-/-</sup>, il y a une très grande variabilité au sein des HIOs de même génotype et le niveau d'expression faible voire nul de certaines hormones en RT-qPCR (*GCG*, *GIP* et *CCK*) ne permet pas de déceler une différence d'expression entre les mutants et les CTRL (non montré).



**Figure 63. Analyse des HIOs RFX6<sup>-/-</sup> au stade J28 par RT-qPCR sur divers marqueurs intestinaux d'intérêt**

hiPSCs : cellules indifférenciées (contrôle négatif), CTRL: HIOs LV37 à J28 (contrôle positif), RFX6<sup>-/-</sup> : HIOs RFX6<sup>-/-</sup> clones 1, 14 et 18 à J28. Chaque point représente un regroupement d'une quinzaine d'organoides. Les points rouges représentent les échantillons dont l'ARNm a été séquencé. (A) CDX2 : marqueur général d'épithélium intestinal, (B) RFX6, (C) CHGA: cellules entéroendocrines, (D) FABP2 : entérocytes, (E) LYZ : cellules de Paneth, (F) MUC2 : cellules à mucus, (G) NKX2.2 : régulation biosynthèse sérotonine, (H) LMX1A : régulation biosynthèse sérotonine, (I) TPH1 : sérotonine, (J) SST: somatostatine, (K) NGN3: progéniteurs entéroendocrines, (L) ARX: progéniteurs des cellules L. Normalisation par le gène de référence RPLP0. \*p<0,05 ; \*\*p<0,01 ; \*\*\*p<0,001 Test de Mann-Whitney.

Néanmoins, certains gènes codant les hormones ou leur enzyme de biosynthèse ont tout de même pu être analysés, comme *SST* ou bien *TPH1*, dont l'expression n'est pas significativement modifiée dans les HIOs RFX6<sup>-/-</sup> (RFX6<sup>-/-</sup> vs CTRL : U=16, p=0,10) (**Figure 63I**), contrairement à ce qui a été précédemment montré dans les modèles murins. Quant à l'expression de *SST*, une tendance à l'augmentation est observée dans les HIOs mutants (RFX6<sup>-/-</sup> vs CTRL : U=1, p=0,071) (**Figure 63J**). Au contraire, les analyses sur les souris Rfx6<sup>ΔAdInt</sup> et embryons Rfx6<sup>-/-</sup> ont mis en évidence une diminution des transcrits *Sst* (**Figure 62H-I**) (Piccand et al., manuscrit en préparation), ce qui suggère un effet propre au modèle *in vitro* humain. Chez la souris, la biosynthèse de la sérotonine (5-HT) dans l'intestin est régulée par l'expression des facteurs de transcription *Nkx2.2* et *Lmx1a* (Gross et al., 2016). Dans le modèle d'entéroïdes dérivés de cryptes intestinales Rfx6<sup>ΔAdInt</sup>, l'expression de *Lmx1a* est significativement diminuée, ce qui est en accord avec les baisses associées de *Tph1* et *ChgA*. Chez les souris Rfx6<sup>ΔAdInt</sup>, le niveau d'expression des gènes codant ces deux facteurs de transcription est augmenté, en lien avec l'augmentation de *Tph1* et *ChgA*. C'est pourquoi leur expression a été analysée dans les HIOs RFX6<sup>-/-</sup> et a révélé une augmentation non significative de *NKX2.2* (RFX6<sup>-/-</sup> vs CTRL : U=13, p=0,18) (**Figure 63G**) et une augmentation significative de *LMX1A* (RFX6<sup>-/-</sup> vs CTRL : U=4, p=0,019) (**Figure 63H**). Pourtant, ces augmentations ne sont pas associées à une augmentation de *TPH1* et *CHGA* dans les HIOs RFX6<sup>-/-</sup>. Ainsi, il se pourrait que la régulation transcriptionnelle de la biosynthèse de sérotonine ne soit pas conservée entre les modèles murins et notre modèle *in vitro* humain.

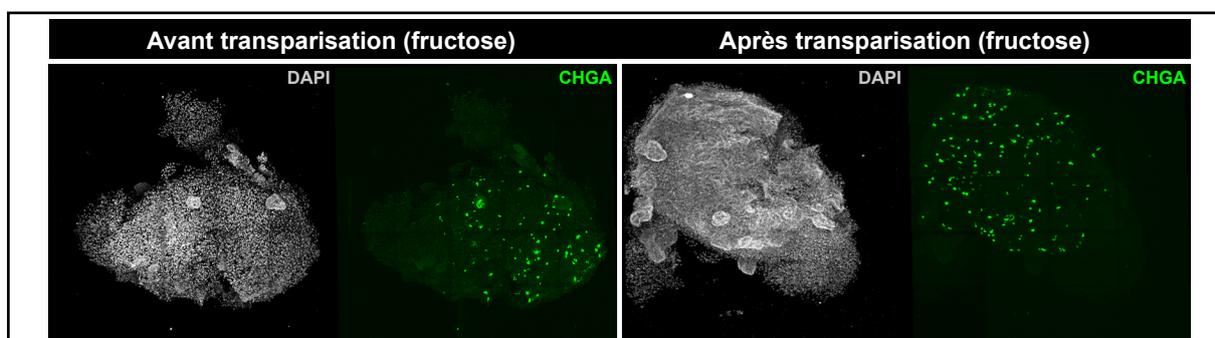
Comme pour les RT-qPCR sur les HIOs NGN3<sup>-/-</sup>, les gènes codant des facteurs de transcription connus pour être impliqués dans la cascade de régulation transcriptionnelle de la différenciation entéroendocrine chez la souris sont faiblement voire nullement exprimés. Parmi ceux détectés, *NGN3* est surexprimé mais cette différence d'expression n'est pas significative (RFX6<sup>-/-</sup> vs CTRL : U=25, p=0,20) (**Figure 63K**), alors que *ARX* est significativement surexprimé dans les HIOs RFX6<sup>-/-</sup> (RFX6<sup>-/-</sup> vs CTRL : U=2, p=0,0006) (**Figure 63L**). En revanche, *Arx* a une expression plus faible chez les mutants dans les modèles murins et *Ngn3* est augmenté de manière significative dans les modèles murins *in vivo* et non significative dans les entéroïdes. En conclusion de ces analyses par RT-qPCR, les HIOs semblent formés de tissus de nature épithéliale et intestinale et expriment des marqueurs des différents types cellulaires de l'épithélium, comme cela a été montré dans le cadre des analyses des HIOs NGN3<sup>-/-</sup>. Les clones 1, 14 et 18 différenciés en HIOs semblent bien déficients pour *RFX6* et cette délétion a l'air d'affecter la formation de certaines cellules entéroendocrines. Les gènes codant certaines hormones et facteurs de transcription ne sont pas détectés dans notre modèle, possiblement à cause d'un manque d'effectif d'échantillons et de quantité de matériel et/ou de l'immaturation des HIOs générés.

Toutefois, la surexpression de *ARX* et la surexpression de *NKX2.2* et *LMX1A* non associée à une surexpression similaire de *CHGA* et *TPH1*, pourrait traduire d'un rôle différent de ces facteurs de transcription chez la souris et chez l'homme. Ces observations, en contradiction avec les effets observés chez la souris, restent à confirmer par immunofluorescence et grâce à l'approche par RNA-seq.

## 2.2. Immunofluorescences *in toto*

Après l'analyse des ARNm des HIOs *RFX6*<sup>-/-</sup> et comparaison avec les modèles murins, des immunofluorescences des marqueurs entéroendocrines ont été réalisées sur des HIOs entiers (*in toto*), afin de confirmer nos observations au niveau de l'expression protéique. De la même façon que les immunofluorescences sur les HIOs *NGN3*<sup>-/-</sup>, les marquages des hormones (GLP1, GIP et 5-HT) n'ont pas fonctionné (non montré). De plus, il nous a été impossible de réaliser des marquages *RFX6*, malgré plusieurs tentatives. Ceci peut s'expliquer par un protocole inadapté à l'anticorps utilisé ou bien par une absence d'expression de *RFX6* dans les HIOs et ce, quel que soit leur génotype. Il sera alors supposé que *RFX6* présente le même patron d'expression dans les HIOs que ce qui a été montré chez la souris (**Figure 62A-G**).

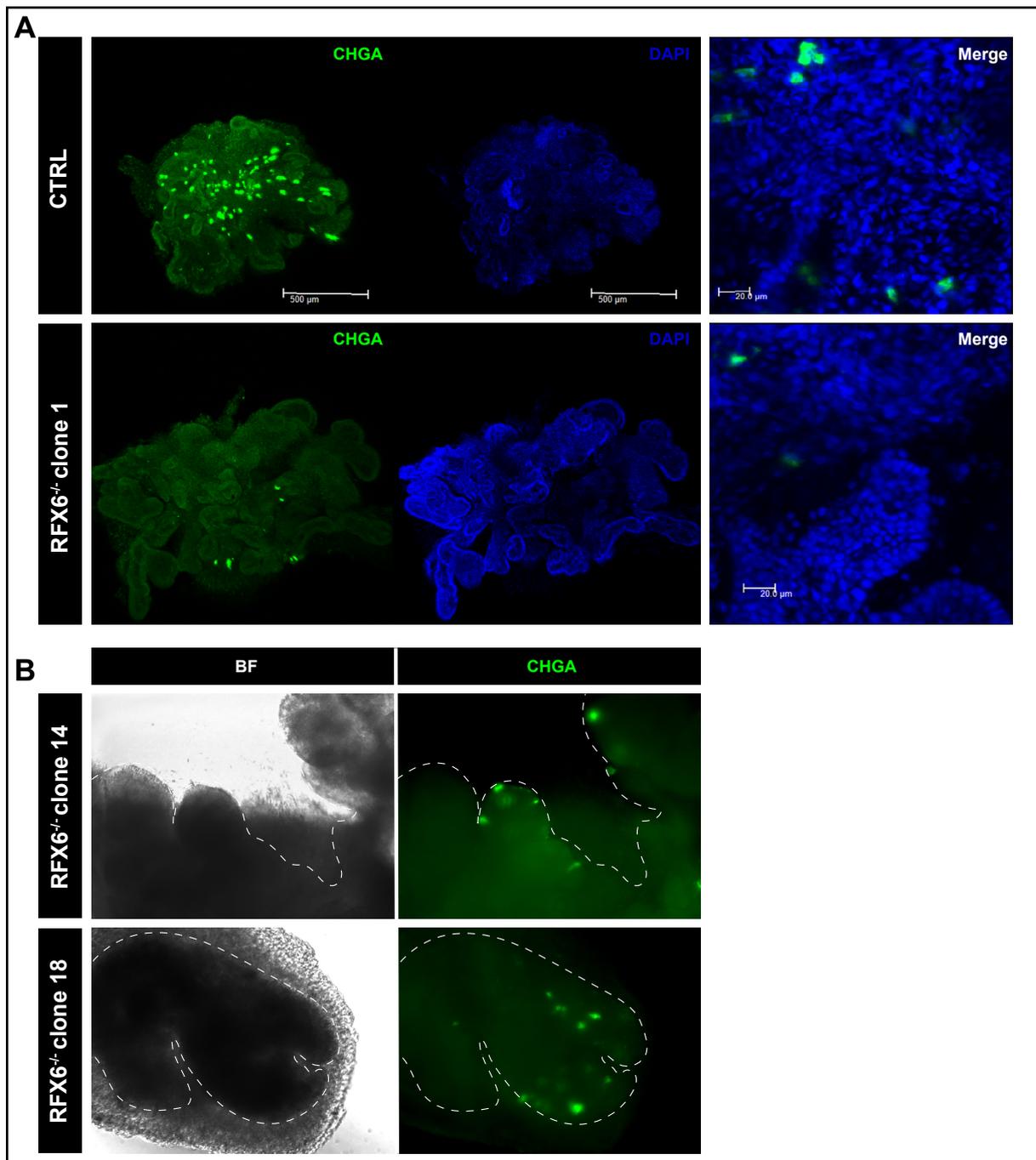
Néanmoins, les cellules entéroendocrines ont pu être marquées par l'expression de *CHGA* dans les HIOs *RFX6*<sup>-/-</sup> et CTRL avant d'être transparisés pour être imagés dans leur intégralité. Etant donné que les HIOs sont très hétérogènes en taille et en forme, il est parfois difficile de les imager en intégralité. En effet, on se heurte souvent à une zone opaque causée par la structure tridimensionnelle volumineuse des HIOs, qui peut empêcher une observation fiable du marquage. Pour contrecarrer ce problème, la technique de transparisation appelée *SeeDB* (Ke et al., 2013) a été utilisée sur les HIOs *RFX6*<sup>-/-</sup> et CTRL.



**Figure 64. Transparisation au fructose des HIOs CTRL après immunofluorescence sur CHGA**

Les images ont été acquises avec un microscope biphotonique et représentent la projection des intensités maximales.

Panneau de gauche : HIO marqué (DAPI + CHGA) et imagé avant transparisation.  
 Panneau de droite : HIO marqué (DAPI + CHGA) et imagé après transparisation.



**Figure 65. Immunofluorescences du marqueur entéroendocrine CHGA sur les HIOs entiers RFX6<sup>-/-</sup> et CTRL au stade J28**

(A) Images prises au macroconfocal du marquage CHGA sur un HIO CTRL entier représentatif (n=4) et un HIO RFX6<sup>-/-</sup> dérivé du clone 1 représentatif (n=2). Les images représentent la projection des intensités maximales. (B) Images prises au microscope à épifluorescence du marquage CHGA sur un HIO RFX6<sup>-/-</sup> issu du clone 14 et un issu du clone 18 (n=2 par clone). Les pointillés délimitent la structure des HIOs.

Ce protocole a initialement été décrit sur des cerveaux de souris fixés mais s'applique à divers échantillons et tissus. Il consiste à incorporer des solutions de fructose à concentration croissante dans l'échantillon. Les solutions sont composées uniquement de fructose et de  $\alpha$ -thioglycérol qui prennent la place des acides gras présents dans l'échantillon et le rendent transparent. Le  $\alpha$ -thioglycérol permet d'éviter la réaction de Maillard causant une autofluorescence et une coloration brune (Hou et al., 2015). C'est une méthode somme toute simple, compatible avec l'utilisation de fluorochromes et ne modifiant pas la morphologie du tissu. Ce protocole reste le plus économique et le moins nocif mais est plutôt adapté pour les échantillons de plus petite taille en comparaison aux méthodes modernes plus sophistiquées qui permettent la transparençation d'organes entiers de souris comme le cerveau (CLARITY, DISCO, etc.) (Wan et al., 2018). L'opacité des HIOs a ainsi pu être outrepassée de manière partielle, voire totale pour certains HIOs, et cela nous a donc permis d'imager à travers les HIOs entiers. Effectivement, comme en témoigne le marquage des cellules entéroendocrines (CHGA<sup>+</sup>) et des noyaux (DAPI) sur des HIOs CTRL, la morphologie tridimensionnelle des HIOs apparaît de façon bien plus détaillée, notamment pour les structures s'apparentant aux cryptes intestinales (**Figure 64**). De plus, les cellules CHGA<sup>+</sup> semblent plus abondantes après transparençation, ce qui laisse penser que l'absence de transparençation pourrait être source d'erreur dans la quantification de cellules positives. La transparençation des HIOs est donc une technique robuste qui, à terme, nous permettra une quantification précise et rapide des cellules marquées. Cependant, il reste des optimisations à réaliser, car certains HIOs sont de taille trop importante pour être entièrement transparents à la suite de ce traitement.

Des images prises au macroconfocal d'un HIO RFX6<sup>-/-</sup> issu du clone 1 et d'un HIO CTRL sont présentées en **figure 65A** et sont représentatives de l'ensemble des HIOs observés (CTRL : n=4 ; RFX6<sup>-/-</sup> clone 1 : n=2). Ici encore, la transparençation permet une meilleure interprétation des détails morphologiques des HIOs et le marquage CHGA met en évidence un nombre de cellules entéroendocrines qui semble être largement accru dans les HIOs CTRL en comparaison aux HIOs mutants (**Figure 65A**). Cette observation est également faite sur les HIOs RFX6<sup>-/-</sup> dérivés des clones 14 et 18, comme en atteste les images prises au microscope à épifluorescence. Les images apparaissent floues, car ce microscope n'est pas adapté à l'observation tridimensionnelle des tissus (**Figure 65B**). Contrairement aux HIOs NGN3<sup>-/-</sup> issus des clones 2 et 10, il persiste des cellules CHGA<sup>+</sup> dans les HIOs RFX6<sup>-/-</sup> et qui semblent notamment plus nombreuses que ce qui était observé dans le cas des HIOs NGN3<sup>-/-</sup> dérivés du clone 49Cd. RFX6 ne semble donc pas essentiel à la différenciation de tous les sous-types de cellules entéroendocrines, mais pourrait être important pour la différenciation de certains d'entre eux dans notre modèle *in vitro* humain.

### 2.3. Conclusion des analyses par RT-qPCR et immunofluorescences

Les premières conclusions sont semblables à celle faites dans l'analyse préliminaire des HIOs *NGN3*<sup>-/-</sup> dans la mesure où les HIOs ont été générés de la même manière et correspondent à du tissu intestinal exprimant des marqueurs de cellules épithéliales intestinales (*CDX2*, *FABP2*, *LYZ* et *MUC2*). Ces marqueurs sont exprimés à des niveaux similaires et les différences d'expression d'autres gènes sont donc vraisemblablement causées par la mutation génétique *RFX6*<sup>-/-</sup> et non à des problèmes dans les expériences de différenciation. Ces analyses par RT-qPCR nous permettent d'affirmer que les HIOs différenciés à partir des trois clones différents sont bien inactivés pour *RFX6*, puisqu'aucun transcrite n'est détecté. De plus, l'expression de l'ARNm de *CHGA* est diminuée dans ces HIOs, ce qui coïncide avec la diminution de cellules entéroendocrines mise en évidence par le marquage de *CHGA*. Bien que la majorité des hormones n'aient pu être détectées dans les HIOs mutants et CTRL, le niveau d'expression de *SST* et *TPH1* a pu être quantifié par RT-qPCR et a révélé une expression inchangée de *TPH1* et une augmentation presque significative de *SST* dans les HIOs *RFX6*<sup>-/-</sup>. Une augmentation de *TPH1* et *CHGA* aurait pu être attendue puisque *Nkx2.2* et *Lmx1a* contrôlent la régulation de la biosynthèse de sérotonine chez la souris (Gross et al., 2016) et dans notre cas, l'expression de *LMX1A* est significativement accrue dans les HIOs *RFX6*<sup>-/-</sup> et l'expression de *NKX2.2* présente également une augmentation non significative. Bien que l'expression de *TPH1* soit inchangée, la diminution de l'expression de *CHGA* associée à l'augmentation de l'expression de *LMX1A* et *NKX2.2*, laisse supposer une régulation différente de la biosynthèse de sérotonine chez la souris et dans les HIOs.

Bien que non significativement, l'expression de *NGN3* semble augmentée dans les HIOs *RFX6*<sup>-/-</sup>, comme c'est le cas dans les modèles murins *in vivo*. Des expériences de ChIP et de transactivation ont révélé une répression de l'activité du promoteur de *Ngn3* par *Rfx6* chez la souris (Piccand et al., manuscrit en préparation). *Rfx6* cible directement *Ngn3* et la répression de *Ngn3* qui en découle est requise pour la différenciation entéroendocrine. Chez les souris *Rfx6*<sup>ΔAdInt</sup>, *Ngn3* n'est plus réprimé par *Rfx6* et ceci induit un nombre accru de progéniteurs *Ngn3*<sup>+</sup>. Cette activité répressive de *Rfx6* sur *Ngn3* semble potentiellement conservée dans nos HIOs. Par contre, une forte augmentation du niveau d'expression de *ARX* a été mise en évidence dans les HIOs *RFX6*<sup>-/-</sup>, contrairement aux résultats obtenus dans les modèles murins pour lesquels une absence ou une forte diminution des transcrits *Arx* est observée (**Figure 62**) (Piccand et al., manuscrit en préparation). Ceci indique possiblement un rôle différent de *RFX6* sur *ARX* chez l'homme. Nos observations sont toutefois à confirmer puisque cela reste des résultats préliminaires sur un nombre d'échantillon limité.

### 3. Séquençage des ARNm (RNA-seq)

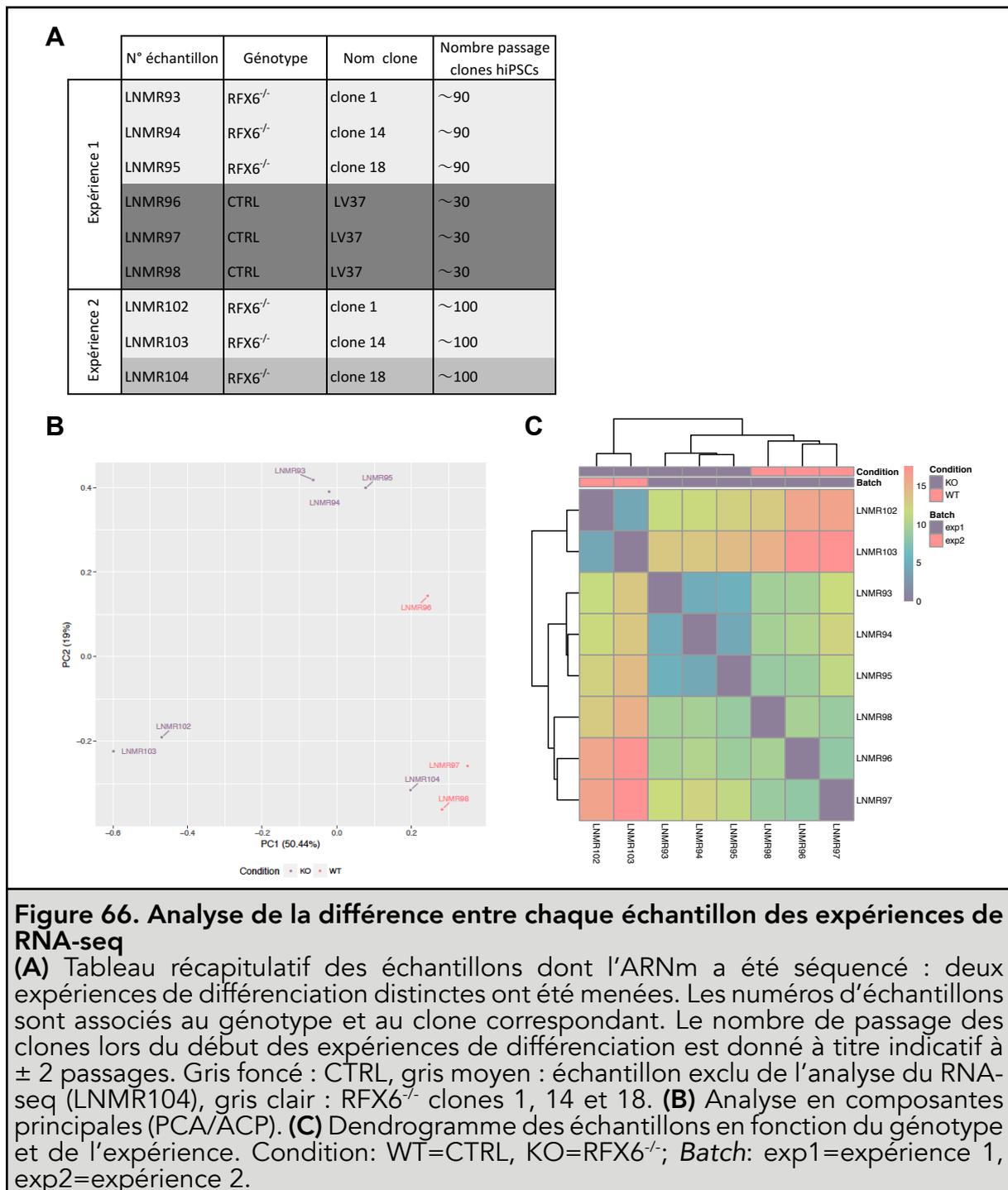
Afin d'étudier le transcriptome des HIOs RFX6<sup>-/-</sup> plus en détails et de confirmer nos observations préliminaires, une expérience de RNA-seq a été réalisée et les gènes différentiellement exprimés ont été analysés. Chaque échantillon est constitué d'une quinzaine d'organoïdes rassemblés. Les HIOs RFX6<sup>-/-</sup> dérivés de chacun des clones sont représentés par deux échantillons issus de deux différenciations indépendantes. Les HIOs CTRL sont représentés par trois échantillons qui sont issus d'une même différenciation (expérience 1) (**Figure 66A**).

Les données de RNA-seq ont été analysées par la bio-informaticienne de l'équipe, Constance Vagne. Les détails de la méthode d'analyse sont disponibles dans la section dédiée du matériel et méthodes (**voir 3.3**).

#### 3.1. Différence entre chaque échantillon

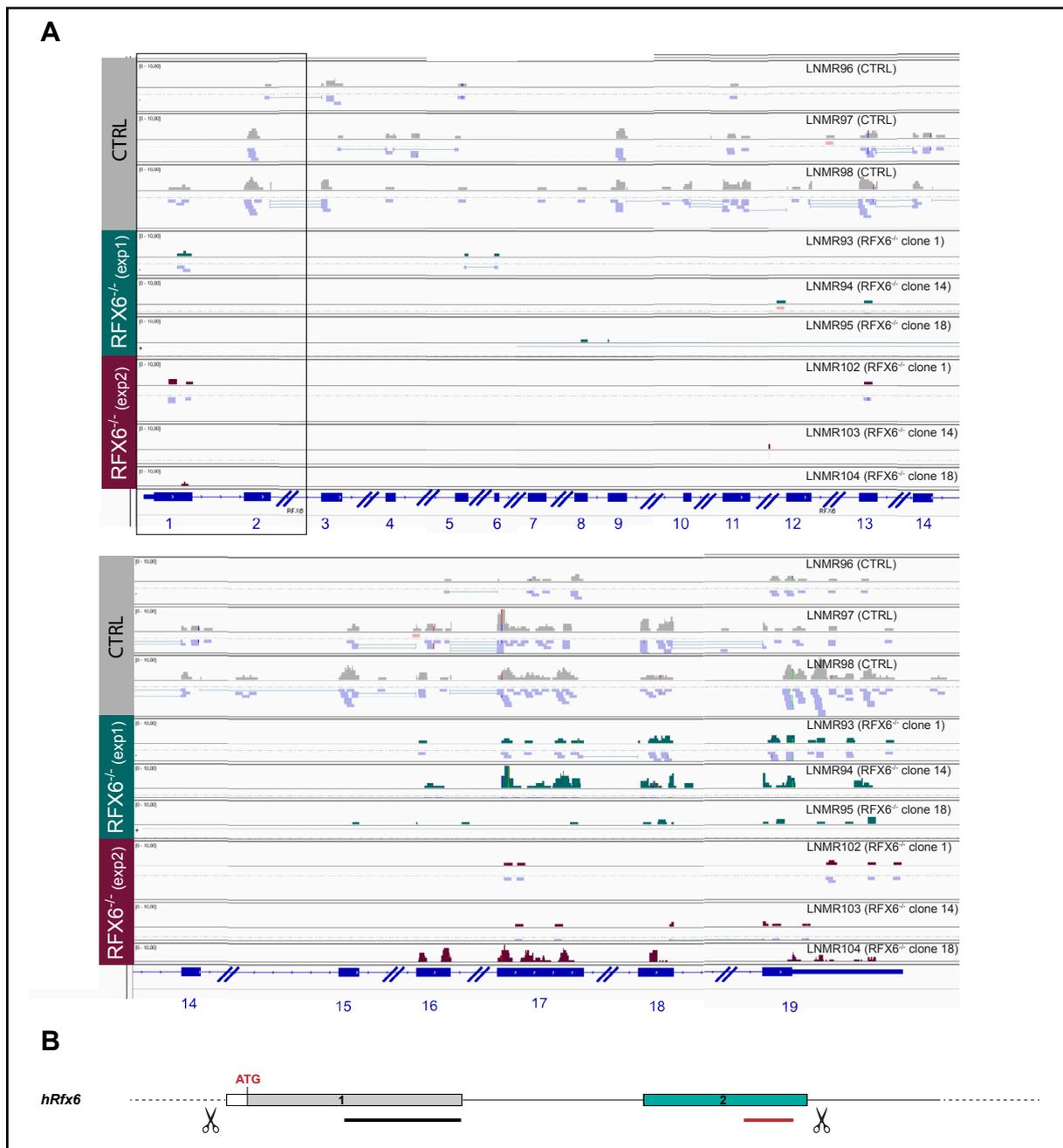
Tout d'abord, une analyse de la distance entre chaque échantillon, basée sur les niveaux d'expression de l'ensemble des gènes, a été réalisée (**Figure 66B-C**). L'analyse en composantes principales (PCA/ACP) met en évidence trois principaux regroupements, l'un avec les HIOs RFX6<sup>-/-</sup> de l'expérience 1, un deuxième avec les HIOs RFX6<sup>-/-</sup> dérivés des clones 1 et 14 de l'expérience 2 et un troisième avec les HIOs CTRL de l'expérience 1, bien que l'un des échantillons soit quelque peu éloigné des autres (LNMR96) (**Figure 66B**). Cependant, l'échantillon LNMR104 (HIOs RFX6<sup>-/-</sup> dérivés du clone 18) qui devrait théoriquement former un groupe avec les autres HIOs RFX6<sup>-/-</sup> de l'expérience 2, est associé aux HIOs CTRL et a donc été exclu des analyses.

L'hétérogénéité des échantillons est également représentée par un dendrogramme qui permet d'identifier leur regroupement selon l'expérience et le génotype (**Figure 66C**). A priori, il semble y avoir principalement un regroupement des échantillons selon l'expérience. En effet, les HIOs RFX6<sup>-/-</sup> et CTRL de l'expérience 1 sont associés, de même que les HIOs RFX6<sup>-/-</sup> de l'expérience 2 le sont entre eux. Ce *batch effect* (effet de l'expérience) a donc été pris en compte dans la normalisation des données de RNA-seq. Cette fois, et contrairement à l'analyse en composantes principales, l'échantillon LNMR96 (CTRL) est associé aux autres HIOs CTRL et n'a ainsi pas été jugé assez différent pour être exclu de l'analyse. Ces regroupements semblent cohérents puisque les mutants forment des groupes entre eux et les CTRL également.



### 3.2. Alignements des reads obtenus sur la séquence d'ADN codant RFX6

Dans le but de confirmer le génotype des HIOs RFX6<sup>-/-</sup> issus des clones 1, 14 et 18, les reads obtenus ont tout d'abord été alignés sur le génome de référence humain Hg38. Les reads s'alignant sur la séquence d'ADN codant RFX6 ont été identifiés et leurs séquences extraites afin de les localiser précisément sur le gène (**Figure 67**). Ceci nous permet de détecter d'éventuels reads s'alignant sur la séquence codante de RFX6 qui est supposée être délétée dans les HIOs RFX6<sup>-/-</sup>.



**Figure 67. Alignement des reads sur RFX6**

(A) Reads alignés sur les différents exons du gène *RFX6* pour chaque échantillon. L'encadré indique la région déletée comprenant l'exon 1 et 2. (B) Alignement des reads obtenus dans les échantillons HIOs *RFX6*<sup>-/-</sup> sur les exons 1 et 2 théoriquement déletés, après extraction de leurs séquences précises. La barre noire représente la localisation des reads présents dans les HIOs *RFX6*<sup>-/-</sup> dérivés du clone 1 (LNMR93 et 102) et l'échantillon LNMR104 (exclu de l'analyse). La barre rouge représente la localisation des reads présents dans l'échantillon LNMR104 (exclu de l'analyse). La partie grisée représente une partie de la séquence codante, la partie en bleu représente la séquence codant une partie du domaine *winged helix* et les ciseaux délimitent la séquence déletée dans les *RFX6*<sup>-/-</sup>.

D'après le génotypage initial des clones *RFX6*<sup>-/-</sup>, la délétion induite comprend les exons 1 et 2, incluant le codon d'initiation ATG. L'alignement créé à partir du logiciel IGV (*Integrative Genomics Viewer*, Robinson et al. 2011) montre une absence de *reads* alignés sur ces deux exons pour les échantillons LNMR94, 95 et 103 (clones 14 et 18). Par contre, les échantillons LNMR93 et 102 (clone 1) présentent des *reads* s'alignant sur l'exon 1 mais en très faible quantité. Seuls les échantillons CTRL présentent des *reads* sur les deux exons (**Figure 67A**). Les séquences des *reads* s'alignant sur l'exon 1 et une partie de l'exon 2 ont été extraites pour les échantillons *RFX6*<sup>-/-</sup>.

Dans les échantillons LNMR93 et 102, deux et trois *reads* s'alignant sur l'exon 1 sont respectivement retrouvés. L'échantillon LNMR104 présente des *reads* sur les deux exons, validant notre décision de ne pas l'inclure dans les analyses des données de RNA-seq (**Figure 67B**). Comme la quantité de *reads* s'alignant sur l'exon 1 de *RFX6* est négligeable dans les échantillons LNMR93 et 102, une contamination avec des échantillons contrôles est envisageable, ce qui ne remet pas en question le génotype des clones *RFX6*<sup>-/-</sup>, considérés comme déficients pour *RFX6*. L'analyse du RNA-seq de ces échantillons devrait donc être représentative de l'effet de l'inactivation de *RFX6* dans notre modèle humain *in vitro*.

### 3.3. Nature épithéliale et intestinale des HIOs générés

Bien que les analyses précédentes par RT-qPCR et immunofluorescences sur les HIOs *NGN3*<sup>-/-</sup> et *RFX6*<sup>-/-</sup>, ainsi que l'analyse des données de RNA-seq des HIOs *NGN3*<sup>-/-</sup>, indiquent la présence de marqueurs intestinaux dans les tissus générés, l'expression de marqueurs des différentes populations cellulaires de l'épithélium intestinal a également été analysée d'après l'expérience de RNA-seq des HIOs *RFX6*<sup>-/-</sup>. L'expression de ces gènes a été comparée entre les HIOs *RFX6*<sup>-/-</sup> et les HIOs CTRL afin de déceler les éventuelles conséquences de la mutation sur la différenciation épithéliale intestinale (**Figure 68**).

Les marqueurs de la population de cellules souches CBC *LGR5*, *OLFM4* et *ASCL2* sont exprimés à un niveau similaire dans les HIOs *RFX6*<sup>-/-</sup> et CTRL. Le marqueur de l'autre population de cellules souches « +4 » (LRC) *BMI* n'est également pas différenciellement exprimé. Par contre, *TERT*, autre marqueur de ces cellules, est significativement sous-exprimé dans les HIOs *RFX6*<sup>-/-</sup>. Globalement, l'inactivation de *RFX6* ne semble pas provoquer d'altération majeure dans la production et la maintenance des cellules souches.

L'expression des gènes *LYZ*, *DLL4* et *XBP1*, marqueurs des cellules de Paneth, est inchangée, alors que *DLL1* présente une tendance à l'augmentation dans les HIOs *RFX6*<sup>-/-</sup>. *DLL1* est exprimé dans les cellules de Paneth mais également par de rares cellules progénitrices descendant des cellules souches CBC *LGR5*<sup>+</sup> (Van Es et al.,

2012). Les cellules de Paneth ne semblent donc globalement pas être affectées par l'absence de *RFX6*, cependant, il se pourrait qu'elle affecte des cellules progénitrices localisées dans le compartiment des cellules en voie de différenciation.

Les cellules à mucus ne semblent pas non plus affectées puisque l'expression des marqueurs *MUC2*, *AGR2* et *KIT* est similaire dans les HIOs mutants et CTRL. Seul *TFF3* est significativement sous-exprimé dans les HIOs *RFX6*<sup>-/-</sup>, ce qui est surprenant puisqu'il est co-exprimé avec *MUC2* dans les cellules à mucus et que l'expression de ce dernier est inchangée (Aamann et al., 2014). L'absence de *RFX6* ne semble vraisemblablement pas altérer la formation des cellules à mucus. Ces observations sont en accord avec les résultats obtenus par RT-qPCR.

Parmi les marqueurs d'entérocytes, les gènes *SLC26A6*, *SLC9A1* et les marqueurs d'entérocytes de l'intestin proximal *CBR1* et *EPHX2* ont une expression inchangée entre les HIOs *RFX6*<sup>-/-</sup> et CTRL. Par contre, *SLC2A2* codant le transporteur de glucose GLUT2 et le marqueur des villosités *VIL1* ont une expression moindre dans les HIOs *RFX6*<sup>-/-</sup>, mais ceci n'est pas nécessairement dû à une altération dans la formation des entérocytes puisque ce ne sont pas les seules cellules épithéliales à exprimer ces deux gènes.

En effet, la villine marque toutes les microvillosités de la bordure en brosse et le transporteur GLUT2 est également exprimé à la surface des cellules entéroendocrines (Ussar, 2017 ; Mace et al., 2012). Le marqueur distal d'entérocytes *MEP1A* est également significativement sous-exprimé dans les HIOs mutants mais il reste à déterminer si la diminution de ce marqueur provient d'une nature moins distale du tissu formé dans les HIOs *RFX6*<sup>-/-</sup>, comme c'est le cas dans les HIOs *NGN3*<sup>-/-</sup>, ou d'une diminution du nombre d'entérocytes au niveau distal.

Concernant les marqueurs de cellules « en touffes », *LRMP*, *PTGS2* et *HOPX* sont exprimés à des niveaux similaires dans les HIOs *RFX6*<sup>-/-</sup> et CTRL, alors que *DCLK1* est significativement surexprimé et *EGFR* légèrement sous-exprimé dans les HIOs *RFX6*<sup>-/-</sup>. Il semble que *DCLK1* ne soit pas forcément un marqueur spécifique des cellules en touffes mais est aussi impliqué dans l'état quiescent et pluripotent des cellules et marquent notamment les cellules souches tumorales caractérisées par une prolifération aberrante (Chandrakesan et al., 2015 ; Nakanishi et al., 2013).

En résumé, de façon générale et malgré des dérégulations de l'expression de certains marqueurs, il ne semble pas qu'un de ces types cellulaires en particulier soit affecté par la délétion de *RFX6*. Ces observations sont en accord avec les modèles murins puisque la différenciation des cellules de Paneth, des cellules à mucus et des entérocytes n'est pas affectée chez les souris *Rfx6*<sup>ΔAdInt</sup> (Piccand et al., manuscrit en préparation).

	Gènes	Log2 FC	P-valeur ajustée	Moyenne CTRL	Moyenne RFX6 <sup>-/-</sup>
Cellules souches	LGR5	-0,13	8,14E-01	1391	1721
	OLFM4	0,06	8,57E-01	405	421
	ASCL2	-1,09	9,25E-02	55	26
	BMI1	-0,15	8,13E-01	587	761
	<b>TERT</b>	<b>-1,75</b>	<b>1,83E-03</b>	<b>117</b>	<b>30</b>
Cellules de Paneth	LYZ	-0,22	9,02E-01	16971	9439
	<b>DLL1</b>	<b>0,68</b>	<b>5,56E-02</b>	<b>675</b>	<b>1342</b>
	DLL4	0,16	7,41E-01	110	109
	XBP1	0,15	5,96E-01	7228	6178
Cellules à mucus	MUC2	0,43	7,76E-01	678	557
	AGR2	-0,53	4,03E-01	39158	19485
	KIT	-0,68	9,55E-02	490	398
	SPEDF	ND	ND	ND	ND
	<b>TFF3</b>	<b>-1,81</b>	<b>2,44E-03</b>	<b>14541</b>	<b>2867</b>
Entérocytes	<b>SLC2A2</b>	<b>-2,54</b>	<b>6,75E-02</b>	<b>402</b>	<b>46</b>
	SLC26A6	-0,69	1,39E-01	1280	699
	SLC9A1	0,3	2,87E-01	1481	1472
	<b>VIL1</b>	<b>-1,39</b>	<b>2,33E-03</b>	<b>12013</b>	<b>3213</b>
	CBR1	-0,34	2,09E-01	570	538
	EPHX2	-0,52	1,61E-01	1439	755
	<b>MEP1A</b>	<b>-2,72</b>	<b>1,65E-02</b>	<b>5432</b>	<b>594</b>
Cellules en "touffes"	LRMP	0,13	9,29E-01	18	17
	<b>DCLK1</b>	<b>1,41</b>	<b>3,23E-05</b>	<b>422</b>	<b>2331</b>
	PTGS2	0,53	2,46E-01	325	312
	HOPX	-0,08	9,41E-01	250	242
	<b>EGFR</b>	<b>-0,64</b>	<b>4,36E-02</b>	<b>4754</b>	<b>2841</b>

**Figure 68. Tableau récapitulatif de l'expression de marqueurs spécifiques aux sous-types cellulaires de l'épithélium intestinal**

Les gènes sont regroupés par type cellulaire. Bleu foncé : lignage absorbant, bleu moyen : lignage sécréteur, bleu clair : autres types cellulaires.

Le Log2(FC) (Fold Change) représente la variation d'expression entre HIOs CTRL et HIOs RFX6<sup>-/-</sup>. Une variation d'expression est jugée significative si la p-valeur ajustée est inférieure à 0,05. La moyenne des reads normalisés est indiquée pour les CTRL et les mutants RFX6<sup>-/-</sup>. En rouge sont indiqués les gènes significativement sous-exprimés dans les HIOs RFX6<sup>-/-</sup>, en bleu: gènes significativement surexprimés, en orange: tendance à l'augmentation ou diminution de l'expression.

### 3.4. Etude de l'expression de marqueurs généraux de cellules entéroendocrines

Bien qu'il soit difficile de considérer ces cellules comme une population homogène, certains marqueurs ont été décrits comme étant exprimés par l'ensemble des sous-types entéroendocrines. Les marqueurs analysés dans cette partie sont les mêmes que dans le cadre de l'analyse des HIOs NGN3<sup>-/-</sup> et sont tirés de la littérature récente (Haber et al., 2017 ; Workman et al., 2016). La majorité des marqueurs des cellules entéroendocrines semble avoir une expression modifiée par l'inactivation de RFX6

(**Figure 69**). En effet, le marqueur générique de cellules entéroendocrines *CHGA*, *REG4*, ainsi que la convertase de prohormone *PCSK9* sont significativement sous-exprimés dans les HIOs *RFX6*<sup>-/-</sup>. A contrario, un certain nombre de marqueurs sont surexprimés chez les mutants, comme les gènes codant les protéines vésiculaires *CHGB*, *SCG2*, *SCG3*, *SCG5*, *VGF*, *SYP* et les gènes codant les convertases des prohormones *PCSK2* et *PCSK1N*. Ceci peut s'expliquer par le fait que ce ne sont pas forcément des marqueurs généraux de cellules entéroendocrines mais ils peuvent être plus ou moins exprimés dans certains sous-types entéroendocrines.

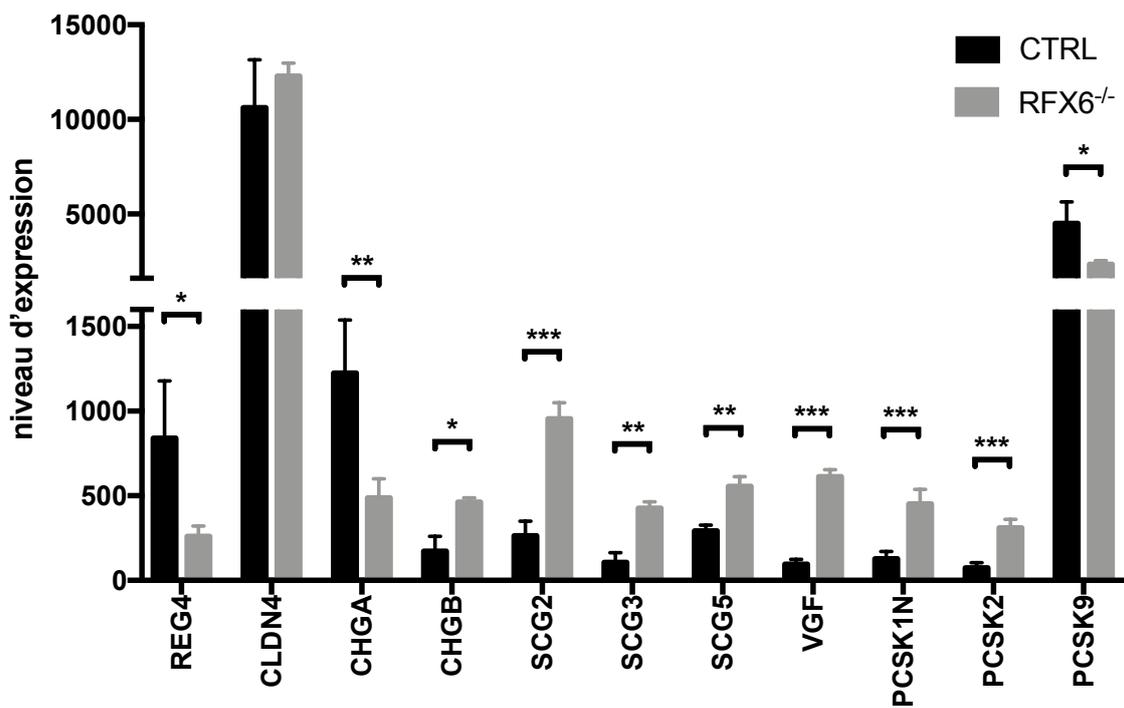
Bien qu'il soit difficile d'établir une conclusion à ces résultats a priori contradictoires, ces observations semblent être en faveur d'un remaniement des sous-types entéroendocrines dû à l'absence de *RFX6*. Il est donc vraisemblable que l'absence de *RFX6* ait un impact sur la différenciation et/ou la fonction des cellules entéroendocrines dans notre modèle *in vitro*. Cette hypothèse sera explorée par l'étude de l'expression des gènes codant les différents produits de sécrétion (ou leur enzyme de biosynthèse) des cellules entéroendocrines et des facteurs de transcription régulant leur différenciation chez la souris.

### 3.5. Impact de l'inactivation de *RFX6* sur la différenciation entéroendocrine

#### 3.5.1. Expression des gènes codant les hormones ou leurs enzymes de biosynthèse exprimées par les cellules entéroendocrines

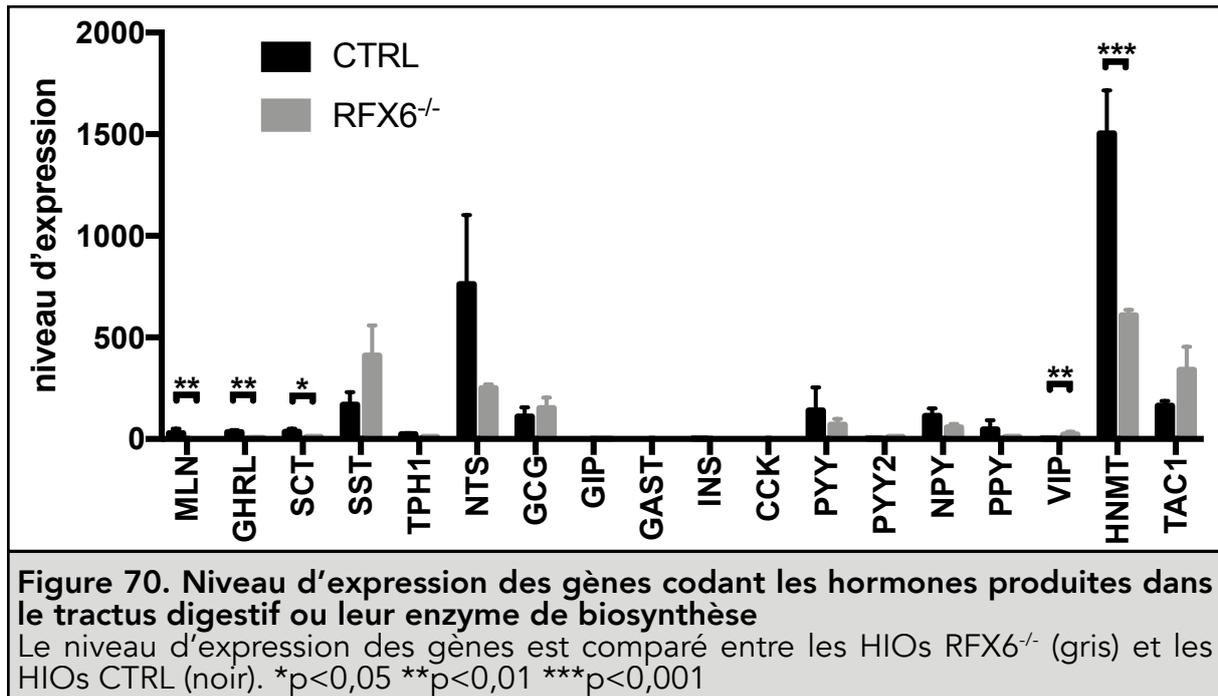
L'expression des marqueurs entéroendocrines étant grandement affectée dans les HIOs *RFX6*<sup>-/-</sup>, l'expression des gènes codant les hormones a été analysée afin de déterminer quels sous-types entéroendocrines sont altérés par l'inactivation de *RFX6* (**Figure 70**). De façon similaire aux analyses des HIOs *NGN3*<sup>-/-</sup>, les gènes *GIP*, *GAST*, *INS* et *CCK* ne sont à nouveau pas détectés dans les HIOs. Il est donc possible de conclure que ces hormones ne sont tout simplement pas exprimées dans les HIOs générés dans nos conditions expérimentales. Bien que les gènes *SCT*, *GHRL*, *MLN*, *GCG*, *PPY*, *TPH1*, *PYY*, *NPY* *PYY2* et *VIP* ne soient que faiblement exprimés, une diminution de l'expression de *MLN*, *GHRL* et *SCT* a pu être mise en évidence dans les HIOs mutants, ainsi qu'une augmentation de *VIP*. Les niveaux d'expression de *SST*, *NTS*, *HNMT* et *TAC1* sont importants dans les HIOs et parmi eux, seul le gène codant l'enzyme de biosynthèse de l'histamine *HNMT* est significativement sous-exprimé dans les HIOs *RFX6*<sup>-/-</sup>. De plus, une diminution non-significative de *NTS* et une augmentation presque significative de l'expression de *TAC1* ( $p=0,06$ ) et *SST* ( $p=0,064$ ) est observée chez les mutants. Ceci est en accord avec la tendance à l'augmentation du niveau de transcrit *SST* montrée en RT-qPCR.

	Gènes	Log2 FC	P-valeur ajustée	Moyenne CTRL	Moyenne RFX6 <sup>-/-</sup>
Cellules entéroendocrines	CHGA	-1,33	3,93E-03	1224	635
	CHGB	1,42	2,13E-02	173	918
	REG4	-1,69	2,69E-02	840	242
	CLDN4	0,21	7,10E-01	10614	7813
	SYP	2,09	3,99E-02	211	1554
	SCG2	1,86	8,16E-06	264	1603
	SCG3	2,02	2,04E-03	106	877
	SCG5	0,93	1,87E-03	292	795
	PCSK1N	1,83	4,07E-04	127	741
	PCSK1	1,03	1,26E-01	146	543
	PCSK2	2,09	1,32E-04	73	486
	PCSK9	-0,94	2,04E-02	4504	1843
	VGf	2,68	4,55E-13	96	697



**Figure 69. Expression de marqueurs généraux de cellules entéroendocrines**  
 Panneau de la partie haute: tableau listant des marqueurs de cellules entéroendocrines. Le Log2(FC) (Fold Change) représente la variation d'expression entre HIOs CTRL et HIOs RFX6<sup>-/-</sup>. Une variation d'expression est jugée significative si la p-valeur ajustée est inférieure à 0,05. La moyenne des reads normalisés est indiquée pour les CTRL et les mutants RFX6<sup>-/-</sup>. En rouge sont indiqués les gènes significativement sous-exprimés dans les HIOs RFX6<sup>-/-</sup> et en bleu sont indiqués les gènes significativement surexprimés dans les HIOs RFX6<sup>-/-</sup>. Panneau de la partie basse: histogramme représentant l'expression des marqueurs entéroendocrines dans les HIOs RFX6<sup>-/-</sup> (gris) et CTRL (noir). \*p<0,05 \*\*p<0,01 \*\*\*p<0,001

Au vu de l'ensemble de ces résultats, l'inactivation de *RFX6* semble avoir un impact sur les cellules entéroendocrines d'après les modifications de l'expression des gènes codant les hormones, mais l'étude doit être approfondie pour mettre en évidence un effet de l'inactivation de *RFX6* spécifiquement sur un ou plusieurs sous-types entéroendocrines. Dans ce but, l'expression de facteurs de transcription régulant la différenciation entéroendocrine chez la souris a été analysée dans les HIOs, afin de déterminer les voies de différenciation affectées et de caractériser plus en détail le rôle de *RFX6* dans la différenciation épithéliale intestinale chez l'homme.



### 3.5.2. Expression des gènes codant des facteurs de transcription régulant la différenciation entéroendocrine

Rfx6 est inscrit dans une cascade de facteurs de transcription, en aval de Ngn3, régulant la différenciation de sous-types entéroendocrines chez la souris (Piccand et al., manuscrit en préparation). Il est ainsi attendu que cette régulation transcriptionnelle soit perturbée par l'inactivation de *RFX6* dans les HIOs. Les facteurs de transcription impliqués chez la souris ont donc été analysés dans nos HIOs RFX6<sup>-/-</sup> (**Figure 71**). *ATOH1/MATH1* est exprimé dans des progéniteurs précoces des lignées sécrétrices et semble surexprimé dans les HIOs RFX6<sup>-/-</sup>. Au contraire, *HES1* qui est exprimé dans les progéniteurs des entérocytes, est significativement sous-exprimé chez les mutants. Il semblerait donc que l'inactivation de *RFX6* favorise la différenciation des cellules souches en progéniteurs des lignées sécrétrices plutôt qu'en progéniteurs des entérocytes. Cependant, une altération dans la différenciation de la majorité de ces lignages cellulaires n'a pas été mise en évidence par l'analyse des marqueurs des

différents types cellulaires de l'épithélium et suggère ainsi une implication de RFX6 restreinte à la spécification des progéniteurs et/ou à la différenciation des cellules entéroendocrines (**Figures 68 et 69**). Une telle fonction de RFX6 dans la spécification des cellules souches en progéniteurs des lignages sécréteurs ou absorbants n'a pas été décrite dans les modèles murins, où l'expression des marqueurs de ces progéniteurs n'est pas affectée (**Figure 62**).

Dans l'intestin de la souris adulte, *Rfx6* est exprimé dans les progéniteurs entéroendocrines *Ngn3*<sup>+</sup>, puis son expression est maintenue dans les cellules sécrétrices d'hormones (**Figure 62B-G**). Plusieurs facteurs de transcription agissent en amont et/ou en aval de *Rfx6* pour réguler la différenciation des sous-types entéroendocrines mais les mécanismes de signalisation précis ne sont pas bien connus et la position exacte de *Rfx6* dans cette cascade de régulation reste indéterminée. L'équipe a mis en évidence une augmentation des transcrits *Ngn3* chez les souris *Rfx6*<sup>ΔAdInt</sup>, due à une accumulation des progéniteurs *Ngn3*<sup>+</sup> dans les cryptes, liée à la perte de *Rfx6*.

Gènes	Log2 FC	P-valeur ajustée	Moyenne CTRL	Moyenne RFX6 <sup>-/-</sup>
<b>ATOH1/MATH1</b>	1,86	5,48E-02	7	26
<b>HES1</b>	-1,13	5,28E-04	2389	1128
<b>NEUROD1</b>	2,11	6,22E-03	32	325
<b>ARX</b>	2,22	6,65E-03	10	176
<b>NKX2-2</b>	1,99	5,04E-01	6,2	132
<b>PAX6</b>	0,83	6,01E-01	58	487
<b>PAX4</b>	0,61	ND	1,6	1,6
<b>PDX1</b>	-0,87	5,14E-01	1031	387
<b>FOXA1</b>	-0,27	5,96E-01	2743	1947
<b>FOXA2</b>	-0,09	8,61E-01	932	698
<b>NGN3</b>	1,4	2,95E-01	3,5	28
<b>INSM1</b>	2,33	1,91E-08	65	600
<b>ISL1</b>	0,36	6,45E-01	261	759
<b>LMX1A</b>	2,96	2,73E-04	7,4	87

**Figure 71. Tableau récapitulatif de l'expression de facteurs de transcription régulant la différenciation des cellules de l'épithélium intestinal**

Le Log2(FC) (Fold Change) représente la variation d'expression entre HIOs CTRL et HIOs RFX6<sup>-/-</sup>. Une variation d'expression est jugée significative si la p-valeur ajustée est inférieure à 0,05. En rouge est indiqué un gène significativement sous-exprimé dans les HIOs RFX6<sup>-/-</sup>, en orange est indiqué un gène pour lequel l'analyse statistique révèle une tendance à l'augmentation de l'expression et en bleu sont indiqués les gènes significativement surexprimés. ND : non déterminé, car l'expression du gène est insuffisante pour l'analyse statistique.

Rfx6 est nécessaire en aval de Ngn3 pour induire les programmes de différenciation des sous-types entéroendocrines par des mécanismes impliquant la répression directe de *Ngn3* chez la souris (Piccand et al., manuscrit en préparation). Dans nos HIOs RFX6<sup>-/-</sup>, le niveau d'expression de *NGN3* est augmenté de manière non significative. Cette observation peut s'expliquer par un effectif d'échantillon trop réduit pour mettre en évidence une réelle différence, une expression trop faible au sein des HIOs ou bien un rôle ou une interaction différente entre les deux facteurs de transcription chez l'homme et chez la souris. Dans les HIOs RFX6<sup>-/-</sup>, l'expression de *NEUROD1*, principale cible de *NGN3* et impliqué dans la différenciation des cellules sécrétrices de CCK et de SCT, est augmentée de manière significative, ce qui favorise l'hypothèse d'une augmentation de *NGN3* masquée par l'un des paramètres cités. *NeuroD1* est également surexprimé dans le côlon des souris Rfx6<sup>ΔAdInt</sup> mais pas dans la partie plus proximale de l'intestin (Piccand et al., manuscrit en préparation). Il est donc plausible que les tissus formés *in vitro* soient de nature plus distale ou bien que cette régulation ne soit pas restreinte au côlon chez l'homme. Le niveau des transcrits *ARX* est significativement et fortement augmenté dans les HIOs RFX6<sup>-/-</sup>, alors qu'il est diminué dans les modèles murins. Par contre, *Pax4/PAX4*, qui est augmenté chez la souris Rfx6<sup>ΔAdInt</sup>, n'est pas suffisamment exprimé dans les HIOs pour déceler une différence dans son expression. Ainsi, il est envisageable que la balance *ARX/PAX4*, qui régule la différenciation de certains sous-types entéroendocrines (Beucher et al., 2012), soit inversée dans les HIO RFX6<sup>-/-</sup> en comparaison au souris Rfx6<sup>ΔAdInt</sup>.

*Insm1* est impliqué dans la différenciation des cellules sécrétrices de Nts et de substance P et de la même manière que *Neurod1*, son expression n'est pas affectée dans le jéjunum mais légèrement augmentée dans le côlon des souris Rfx6<sup>ΔAdInt</sup> (Piccand et al., manuscrit en préparation). Dans nos HIOs RFX6<sup>-/-</sup>, *INSM1* est surexprimé, ce qui pourrait renforcer l'hypothèse de la formation de tissu plus distal ou celle d'une fonction de *INSM1* et *NEUROD1* non restreinte au côlon chez l'homme. Chez les souris déficientes pour *Rfx6*, la différenciation de toutes les cellules entéroendocrines est affectée, hormis les cellules entérochromaffines productrices de sérotonine. L'expression insuffisante de *TPH1* dans les HIOs ne nous permet pas d'établir une analogie entre la souris et l'homme. Comme observé par RT-qPCR, l'expression de *LMX1A* est augmentée dans les HIOs RFX6<sup>-/-</sup> et une augmentation non significative de *NKX2.2* est également observée. Ces deux facteurs de transcription régulent la biosynthèse de sérotonine dans les cellules entérochromaffines chez la souris (Gross et al., 2016) et bien que cela ne puisse être vérifié, il serait donc théoriquement attendu que l'augmentation de l'expression de ces facteurs de transcription provoque une expression accrue de sérotonine. Mais tout comme dans les analyses par RT-qPCR, l'expression de *CHGA*, marqueur entéroendocrine préférentiel des cellules entérochromaffines, est diminuée, suggérant plutôt une diminution de sérotonine.

Les autres facteurs de transcription impliqués dans la différenciation des sous-types de cellules sécrétrices d'hormones *PDX1*, *PAX6*, *FOXA1*, *FOXA2* et *ISL1* ne sont pas différenciellement exprimés entre les HIOs *RFX6*<sup>-/-</sup> et CTRL. Les observations faites en RT-qPCR ont été confirmées par l'expérience de RNA-seq mais comme mentionné précédemment, le nombre d'échantillon limitant la détection de potentiels effets de la mutation d'un point de vue statistique, il faudrait réitérer l'expérience pour augmenter l'effectif.

### 3.6. Comparaison des variations d'expression des gènes d'intérêt entre les HIOs *RFX6*<sup>-/-</sup> et les modèles murins

Les résultats obtenus dans les HIOs pour l'ensemble des gènes d'intérêt, c'est-à-dire les gènes codant les hormones ou leur enzyme de biosynthèse et les gènes codant des facteurs de transcription impliqués dans la régulation de la différenciation entéroendocrine, ont été comparés à ceux générés précédemment dans les différents modèles murins (**Figure 72**). Il y a une relative hétérogénéité dans les variations d'expression génique selon les modèles étudiés. En effet, en fonction des gènes considérés, les variations d'expression peuvent être caractéristiques de l'organisme (souris/homme), du contexte (*in vivo/in vitro*) ou encore de la nature du tissu (fœtal/adulte). *Arx/ARX* par exemple, est significativement sous-exprimé dans les modèles murins et surexprimé dans les HIOs *RFX6*<sup>-/-</sup>. L'expression de *ChgA/CHGA* est augmentée chez les souris *Rfx6*<sup>ΔAdInt</sup> ou inchangée chez les embryons *Rfx6*<sup>-/-</sup>, alors qu'elle est significativement diminuée dans les modèles *in vitro*. Concernant *Tac1/TAC1*, une tendance à la diminution d'expression est observée dans les modèles d'origine fœtal (embryons E18 et HIOs), tandis que son expression n'est pas modulée dans les autres modèles. L'expression de certains gènes est modulée de manière similaire entre tous les modèles, comme c'est le cas pour *Ghrl/GHRL* et *Sct/SCT*. D'autre part, certaines modulations semblent propres à notre modèle humain, comme l'augmentation de *SST*, *VIP*, *ARX*, *ATOH1* et *INSM1*, ainsi que la diminution de *MLN*, *HNMT* et *HES1*. Ces spécificités pourraient être attribuées à un rôle spécifique de *RFX6* chez l'homme mais il faudrait confirmer ces observations par d'autres approches complémentaires. Globalement, la dérégulation de l'expression d'un nombre important de gènes d'intérêt dans les différents modèles étudiés suggère un rôle crucial de *RFX6* dans la différenciation des sous-types entéroendocrines chez l'homme, comme chez la souris. Néanmoins, l'hétérogénéité des résultats obtenus par l'analyse des HIOs ne permet pas de conclure quant à sa fonction exacte dans ces processus. Malgré la conservation d'une part importante des mécanismes régissant le développement et la différenciation intestinale entre la souris et l'homme, il est probable que des différences fines existent entre ces deux espèces et se révèlent en partie au travers de nos analyses.

	Souris <i>in vivo</i>		Souris <i>in vitro</i>	Humain <i>in vitro</i>
	Rfx6 <sup>ΔAdInt</sup>	Rfx6 <sup>-/-</sup> E18	Entéroïdes Rfx6 <sup>-/-</sup>	HIOs RFX6 <sup>-/-</sup>
CHGA	↗	=	↘	↘
TPH1	↗	↗	↘	=
SST	↘	↘	N/A	↗ (p=0,064)
CCK	↘	↘	↘	N/D
NTS	↘	↘	↘	=
GIP	↘	↘	=	N/D
GHRL	↘	↘	N/A	↘
MLN	N/D	N/D	N/A	↘
GCG	↘	↘	N/A	=
PYY	↘	=	N/A	=
HNMT	=	=	N/A	↘
VIP	=	=	N/A	↗
TAC1	=	↗ (p=0,063)	N/A	↗ (p=0,06)
SCT	↘	↘	=	↘
NKX2.2	↗	=	N/A	=
LMX1A	↗	↗	↘	↗
ARX	↘	↘	↘	↗
NGN3	↗	↗	=	=
ATOH1	=	=	=	↗ (p=0,054)
HES1	=	=	=	↘
NEUROD1	↗	=	N/A	↗
PAX4	↗	=	↗	N/D
INSM1	=	=	N/A	↗
PDX1	=	↘	N/A	=
PAX6	↘	=	↘	=
FOXA1	=	↗	=	=
FOXA2	=	↗	=	=
ISL1	↘	=	↘	=

**Figure 72 . Tableau comparatif des phénotypes entéroendocrines observés dans différents modèles déficients pour *Rfx6/RFX6***

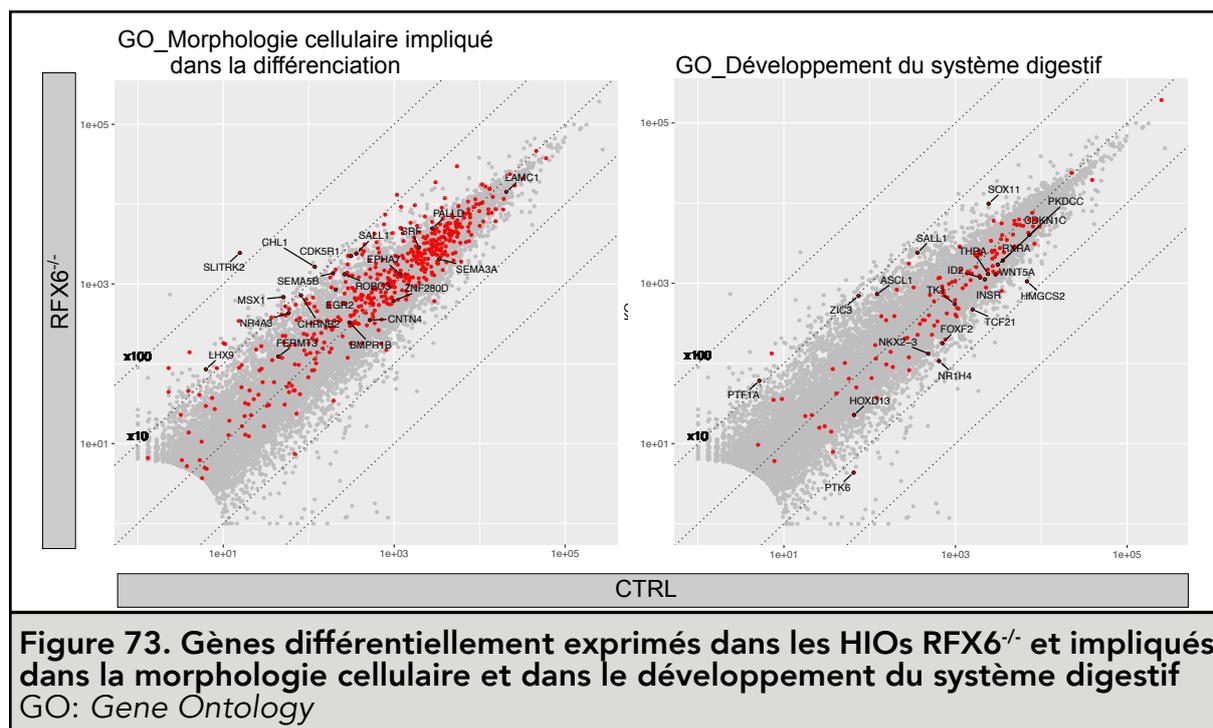
Les modèles murins *in vivo* regroupent les souris Rfx6<sup>-/-</sup> à E18 et les souris Rfx6<sup>ΔAdInt</sup> et les entéroïdes Rfx6<sup>-/-</sup> constituent le modèle murin *in vitro*. Notre modèle *in vitro* HIOs RFX6<sup>-/-</sup> représente le seul modèle humain existant. Les gènes comparés consistent en les gènes codant des hormones ou leur enzyme de biosynthèse et les gènes codant des facteurs de transcription régulant la différenciation des cellules entéroendocrines. N/A: non analysé, N/D: non détecté, ↘: expression diminuée chez les mutants (rouge), ↗: expression augmentée chez les mutants (vert), =: expression inchangée.

### 3.7. Analyses fonctionnelles des résultats de RNA-seq

Afin de caractériser plus en détails les conséquences de l'inactivation de *RFX6*, les gènes les plus différenciellement exprimés ont été étudiés (**Annexe 4**), notamment par des analyses fonctionnelles. Celles-ci permettent de déterminer si ces gènes sont impliqués dans des voies ou fonctions communes. Les informations relatives aux méthodes d'analyse sont disponibles dans la section dédiée du matériel et méthodes (**voir 3.3.**). Parmi les groupes de gènes identifiés lors de l'analyse fonctionnelle sont retrouvés «Morphogenèse cellulaire impliquée dans la différenciation», «Développement du système digestif», «Métabolisme des lipides», «Métabolisme du tryptophane», «Glycolyse et gluconéogenèse», «Activité canaux cationiques» et «Signalisation synaptique».

#### 3.7.1. Gènes impliqués dans le développement, la morphogénèse et la différenciation de l'intestin

Dans le cas de la pathologie de Mitchell-Riley causée par des mutations de *RFX6*, le diabète néonatal est associé à de sévères malformations du tractus gastro-intestinal, comme l'atrésie, la sténose et la malrotation intestinale (Smith et al., 2010). De plus, une accumulation de progéniteurs exprimant *Ngn3* est retrouvée dans les cryptes des souris *Rfx6*<sup>-/-</sup>, témoignant d'une prolifération accrue au niveau de l'iléon et du côlon (Piccand et al., manuscrit en préparation). *RFX6* étant exprimé précocement au cours du développement du tractus digestif dans l'endoderme intestinal, il n'est pas surprenant d'observer de tels défauts anatomiques en l'absence de ce facteur de transcription.



Dans ce contexte, il est intéressant de voir que les groupes fonctionnels «développement du système digestif» et «morphogenèse cellulaire impliquée dans la différenciation» apparaissent lors de l'analyse fonctionnelle des gènes différentiellement exprimés (**Figure 73**).

Parmi ces gènes impliqués dans le développement intestinal et sous-exprimés dans les HIOs RFX6<sup>-/-</sup> figure *PKDCC*, qui est exprimé pendant l'embryogenèse mais aussi dans certains tissus adultes. Les embryons de souris *Pkdcc*<sup>-/-</sup> ont des anomalies morphologiques diverses, dont un rétrécissement de l'intestin (Imuta et al., 2009). C'est également le cas pour les souris mutantes pour *Cdkn1c*, gène codant une protéine qui cause l'arrêt du cycle cellulaire en phase G1, qui présentent des anomalies du tube digestif allant jusqu'à l'absence des segments intestinaux jéjunum et iléon (Yan et al., 1997) et dont l'expression est significativement diminuée dans les HIOs RFX6<sup>-/-</sup>. En rapport avec l'augmentation du nombre de progéniteurs Ngn3<sup>+</sup> associée à une augmentation de la prolifération au sein des cryptes dans le modèle murin *Rfx6*<sup>-/-</sup>, il est intéressant d'observer la présence d'une diminution de *NKX2.3* dans les HIOs RFX6<sup>-/-</sup>, puisque ce gène est exprimé dans le mésenchyme de l'intestin grêle et sa délétion chez la souris provoque une hyperprolifération des cellules dans les cryptes intestinales (Pabst et al., 1999). D'autres gènes de ce groupe fonctionnel sont surexprimés, tel que le *serum response factor (SRF)*, qui est un facteur de transcription régulant l'expression de plusieurs gènes impliqués dans la prolifération, migration, différenciation et apoptose cellulaires. En effet, une surexpression de ce facteur, notamment dans l'épithélium gastrique, entraîne une prolifération cellulaire accrue (Modak et al., 2010). Ces observations suggèrent un défaut dans la prolifération cellulaire similaire au phénotype des souris *Rfx6*<sup>-/-</sup>, ainsi que de potentielles anomalies morphologiques dans les HIOs RFX6<sup>-/-</sup> en lien avec celles observées au niveau du tube digestif des patients souffrant du syndrome de Mitchell-Riley.

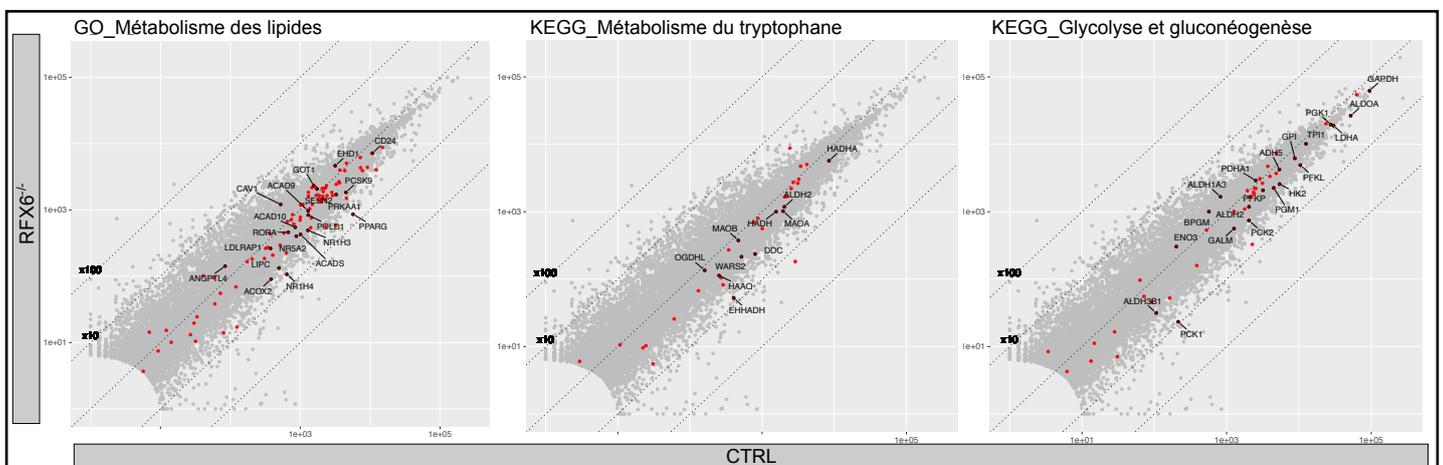
Il serait alors intéressant de réaliser des analyses histologiques des HIOs, ainsi que des analyses de la prolifération cellulaire, par immunomarquage sur le marqueur de prolifération Ki67, par exemple.

### 3.7.2. Gènes impliqués dans le métabolisme et la physiologie des cellules épithéliales intestinales

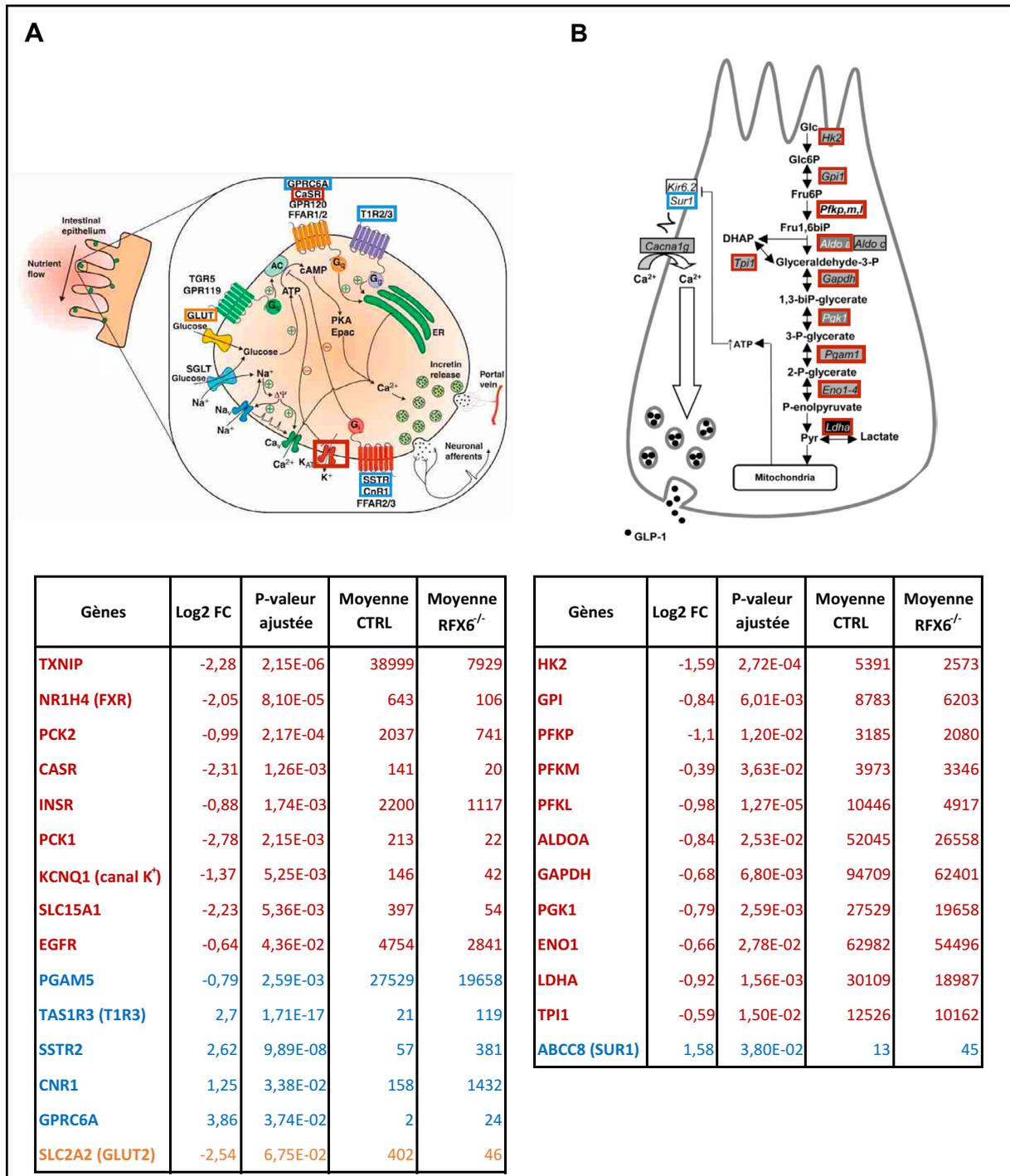
Le maintien de l'expression de *Rfx6* à l'âge adulte dans les cellules endocrines laisse présager d'un rôle dans la régulation de la fonction physiologique de ces cellules. Dans les cellules  $\beta$  du pancréas chez la souris, *Rfx6* contrôle l'homéostasie du glucose et maintient l'état différencié des cellules en activant la voie de sécrétion de l'insuline (Piccand et al., 2014 ; Chandra et al., 2014). De plus, les souris *Rfx6* <sup>$\Delta$ Adint</sup> présentent une malabsorption des lipides et une homéostasie du glucose perturbée, probablement causées par l'absence presque totale des sous-types de cellules

entéroendocrines (Piccand et al., manuscrit en préparation). Bien que le rôle de Rfx6 dans les cellules entéroendocrines ne soit pas décrit à l'heure actuelle, il est envisageable qu'il soit également impliqué dans des fonctions métaboliques similaires à ce qui est observé dans les cellules  $\beta$ . Dans ces conditions, il est intéressant de noter que les groupes fonctionnels «glycolyse et gluconéogenèse», «métabolisme des lipides» et «métabolisme du tryptophane» émergent de l'analyse fonctionnelle des gènes différentiellement exprimés dans les HIOs RFX6<sup>-/-</sup> (**Figure 74**).

L'analyse fonctionnelle met en avant des gènes engagés dans la régulation du métabolisme du tryptophane. Le tryptophane est un précurseur de la sérotonine et de la mélatonine qui sont toutes deux synthétisées dans les mitochondries des cellules entérochromaffines de l'épithélium intestinal (Keszthelyi et al., 2009). Cette voie métabolique implique différentes enzymes catalysant un certain nombre de réactions chimiques aboutissant à la biosynthèse de sérotonine et mélatonine (Hanson et al., 2014). Dans nos HIOs RFX6<sup>-/-</sup>, la majorité des gènes codant ces enzymes sont sous-exprimés, comme par exemple *WARS2*, *DDC*, *MAOA*, *MAOB* et *ALDH2*. De plus, les cellules entérochromaffine-like (ECL) localisées dans l'épithélium gastrique produisent de l'histamine et *HNMT*, gène codant l'enzyme de biosynthèse, est également sous-exprimé dans les HIOs RFX6<sup>-/-</sup>. Ces observations suggèrent une diminution du nombre de cellules productrices de sérotonine (EC) et productrices d'histamine (ECL), ou une diminution de l'activité de biosynthèse et de sécrétion de celles-ci. Ceci indique également que les HIOs sont notamment formés de tissu épithélial d'origine gastrique. L'hypothèse d'un nombre réduit de cellules EC et ECL est d'autant plus crédible qu'une baisse d'expression de *CHGA*, marqueur de ces cellules, est observée dans les HIOs RFX6<sup>-/-</sup>. L'observation d'une baisse de *HNMT* est également intéressante, puisqu'elle suggère une implication de RFX6 dans la différenciation du lignage endocrine gastrique.



**Figure 74. Gènes différentiellement exprimés dans les HIOs RFX6<sup>-/-</sup> et impliqués dans le métabolisme des lipides, le métabolisme du tryptophane et la glycolyse et gluconéogenèse**  
 GO: Gene Ontology, KEGG: Kyoto Encyclopedia of Genes and Genomes



**Figure 75. Gènes différentiellement exprimés dans les HIOs RFX6<sup>-/-</sup> et impliqués dans le métabolisme du glucose et la sécrétion hormonale par les cellules entéroendocrines**

(A) Modèle des mécanismes moléculaires impliqués dans la sécrétion des hormones incrétones par les cellules entéroendocrines (Adapté de Ezcurra et al., 2013) (haut), tableau présentant l'expression de gènes codant des constituants impliqués dans ces mécanismes (bas).

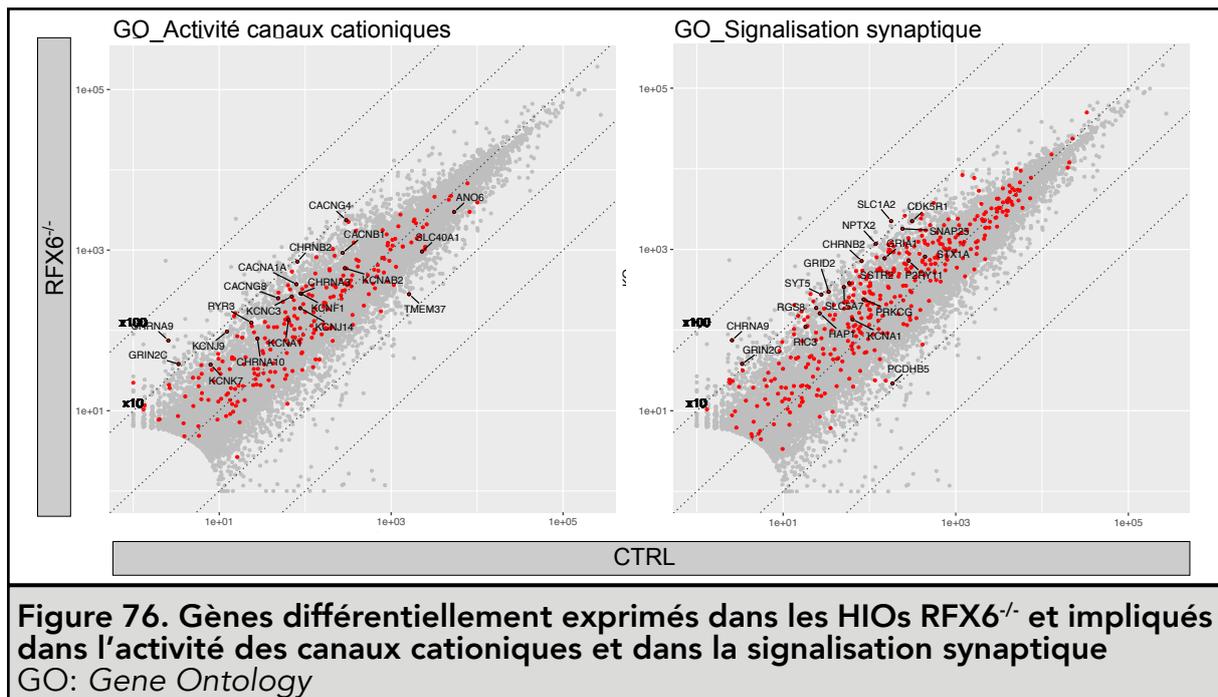
(B) Machinerie moléculaire impliquée dans la glycolyse (Adapté de Trabelsi et al., 2015) (haut), tableau présentant l'expression de gènes codant les enzymes impliquées dans la glycolyse (bas).

Rouge: gènes/constituants/enzymes sous-exprimés, bleu: gènes/constituants/enzymes surexprimés, orange: gène/constituant présentant une tendance à la diminution de l'expression.

La présence du groupe fonctionnel «glycolyse et gluconéogenèse» peut paraître surprenante puisque la gluconéogenèse est classiquement attribuée aux cellules hépatiques et musculaires. Pourtant, bien que cela soit très peu décrit, l'intestin contribue à l'homéostasie du glucose par sa grande capacité glycolytique et par la gluconéogenèse intestinale, notamment en cas de jeûne quand la production de glucose hépatique est limitée (Penhoat et al., 2014 ; Mithieux, 2009). Ainsi, il est envisageable que dans les cellules entéroendocrines différenciées, RFX6 soit impliqué dans la régulation du métabolisme du glucose, comme c'est le cas dans la cellule  $\beta$  pancréatique, notamment par un rôle dans la régulation de la synthèse des hormones incrétines (GLP1 et GIP). En effet, le glucose induit l'expression du gène codant le proglucagon par la voie de la glycolyse dans les cellules L, promouvant la production de GLP1. De plus, la glycolyse augmente la concentration d'ATP cytosolique et induit la sécrétion de GLP1 (Trabelsi et al., 2015). De nombreux gènes impliqués dans ces régulations du métabolisme et de la réponse au glucose sont différentiellement exprimés dans les HIOs RFX6<sup>-/-</sup>, comme illustré en **figure 75**. En effet, des gènes codant des enzymes impliqués dans la glycolyse et/ou la néoglucogenèse sont différentiellement exprimés dans les HIOs RFX6<sup>-/-</sup>. C'est notamment le cas de *PGK1* qui code l'enzyme glycolytique phosphoglycérate kinase catalysant la conversion du 1,3-diphosphoglycérate en 3-phosphoglycérate, dont l'expression est diminuée. *PCK1* et *PCK2*, gènes codant les isoformes de l'enzyme phosphoénolpyruvate carboxykinase, qui catalyse la réaction limitante de la gluconéogenèse, sont tous deux sous-exprimés dans les HIOs RFX6<sup>-/-</sup>. D'autre part, des gènes codant des récepteurs et transporteurs notamment localisés à la surface des cellules entéroendocrines et impliqués dans les mécanismes de sécrétion de GLP1 en réponse au glucose s'avèrent différentiellement régulés (**Figure 75**).

### 3.7.3. Gènes codant des éléments initialement attribués au système nerveux

De manière surprenante, l'analyse fonctionnelle dégage un certain nombre de gènes codant pour des protéines constitutives des canaux ioniques, des vésicules exocytiques et des protéines notamment impliquées dans la signalisation synaptique, qui sont autant de fonctions généralement attribuées au système nerveux (**Figure 76**). Néanmoins, l'expression de certains de ces gènes a également été mise en évidence dans le tractus digestif. C'est notamment le cas de *HAP1* qui, en plus de son rôle dans le cerveau, est impliqué dans le transport vésiculaire dans certaines cellules endocrines, particulièrement celles dont les hormones produites sont dérivées des acides aminés, telles les cellules EC (Liao et al., 2005). La synaptotagmine 5 (*SYT5*) est également une protéine régulant l'exocytose vésiculaire à la fois dans les neurones et les cellules endocrines (Dlugosz et al., 2014). La syntaxine 1 (*STX1A*) est exprimée dans l'épithélium intestinal et dans le cerveau et joue un rôle essentiel dans l'exocytose des hormones incrétines telles que le GLP1

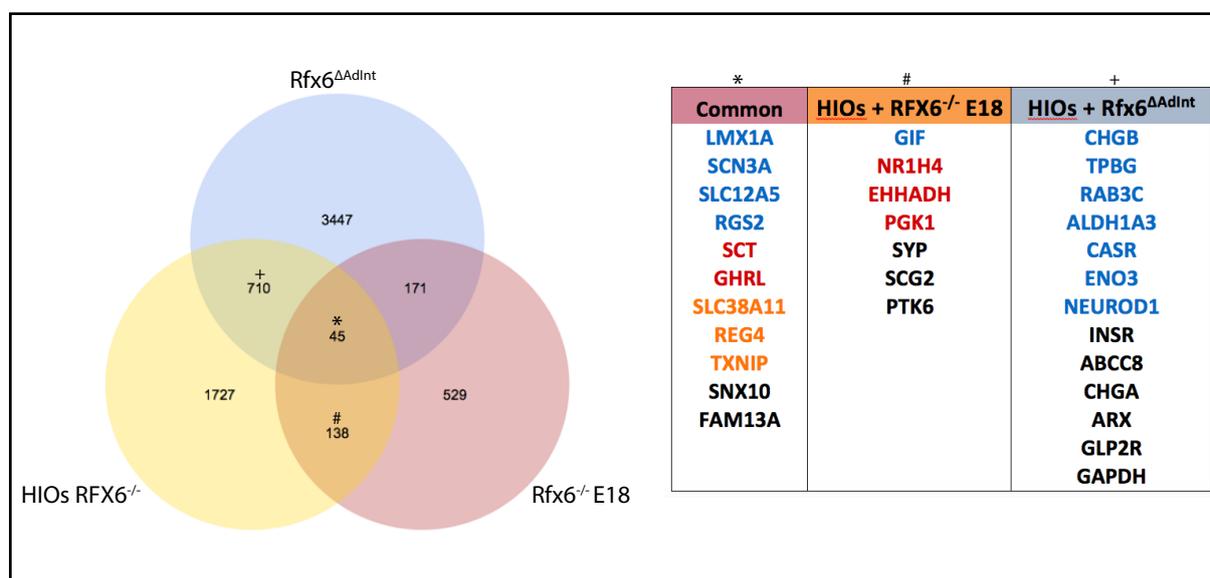


Dans le diabète de type 2 par exemple, l'expression de cette protéine est diminuée dans les îlots pancréatiques provoquant l'insuffisance sécrétoire de l'insuline (Liang et al., 2017). En plus de ces protéines impliquées dans la régulation du trafic vésiculaire, l'expression d'un nombre important de canaux ioniques est augmentée dans les HIOs mutants. Effectivement, des constituants des canaux calciques, sodiques et potassiques connus pour leur expression dans le système nerveux sont aussi exprimés dans les cellules de l'épithélium intestinal où les équilibres ioniques régulent un certain nombre de fonctions cellulaires dont l'exocytose des granules de sécrétion contenant les hormones. Bien que les cellules entéroendocrines soient excitables électriquement, il est généralement pensé que ces cellules communiquent indirectement avec les fibres nerveuses via la sécrétion d'hormone. Pourtant, il a récemment été montré que certaines cellules entéroendocrines possèdent des extensions de leur surface basolatérale, appelés neuropodes, qui pourraient servir d'intermédiaire dans la communication entre ces cellules et le système nerveux innervant l'intestin grêle et le côlon (**Figure 15**) (Bohorquez et al., 2015). En résumé, bien que surprenante de prime abord, la présence de groupes de gènes attribués au système nerveux renforce l'idée des conséquences de l'inactivation de *RFX6* sur les cellules entéroendocrines.

L'analyse transcriptomique globale des HIOs  $RFX6^{-/-}$  donne un grand nombre d'indications quant à des fonctions possibles de *RFX6*, mais dont il est difficile de tirer des conclusions directes.

### 3.7.4. Analyse comparative des données transcriptomiques issues des différents modèles d'inactivation de *Rfx6*/*RFX6*

Afin d'identifier les pistes les plus pertinentes, les données issues de l'analyse des HIOs *RFX6*<sup>-/-</sup> ont été comparées à deux jeux de données de transcriptomique générés par l'équipe dans des modèles murins déficients pour *Rfx6* : analyse sur puces Agilent d'intestin grêle d'embryon de souris E18.5 *Rfx6*<sup>-/-</sup> et analyse de RNA-seq sur l'iléon de souris adulte dont la délétion de *Rfx6* a été spécifiquement induite dans l'intestin adulte (*Rfx6*<sup>ΔAdInt</sup>). L'équipe a tout d'abord étudié l'inactivation de *Rfx6* grâce à un modèle de souris *Rfx6*<sup>-/-</sup> qui ont été analysées au stade embryonnaire E18, car des analyses de stades ultérieurs n'étaient pas possible, étant donné que les souriceaux ne survivaient pas à la mutation. C'est pour cette raison qu'un modèle de souris avec une délétion de *Rfx6* inductible au tamoxifène a été étudié dans un second temps. Ainsi, les gènes différentiellement exprimés dans les trois analyses ont été confrontés et le nombre de gènes régulés communs aux différentes expériences est indiqué dans un diagramme de Venn et des gènes d'intérêt particulier sont listés dans un tableau récapitulatif (**Figure 77**). Il y a 45 gènes différentiellement exprimés en communs entre les trois jeux de données, dont la liste est disponible en **annexe 5**. Parmi eux, 11 gènes d'intérêt sont listés dans le tableau de la **figure 77**.



**Figure 77. Comparaison des gènes différentiellement exprimés entre les différents modèles d'inactivation de *Rfx6*/*RFX6***

Panneau de gauche: Diagramme de Venn représentant le nombre de gènes différentiellement exprimés communs aux différents modèles. Panneau de droite: Tableau listant des gènes d'intérêt différentiellement exprimés communs aux trois modèles (rose/\*), communs aux HIOs *RFX6*<sup>-/-</sup> et souris *Rfx6*<sup>-/-</sup> E18 (orange/#) et commun aux HIOs *RFX6*<sup>-/-</sup> et souris *Rfx6*<sup>ΔAdInt</sup> (bleu/+). En bleu sont représentés les gènes surexprimés en commun, en rouge: les gènes sous-exprimés en commun, en orange: les gènes dont la modulation d'expression est similaire dans les modèles murins mais pas dans les HIOs, en noir: les gènes dont la modulation d'expression est différente entre les différents modèles.

L'expression de *GHRL* et *SCT*, par exemple, est diminuée dans les trois modèles, suggérant ainsi un rôle incontestable de *Rfx6/RFX6* dans la différenciation des sous-types entéroendocrines sécréteurs de ghréline et de sécrétine et/ou dans l'expression des gènes codant ces hormones. *Lmx1a/LMX1A* est surexprimé dans les trois modèles et peut donc être une cible directe réprimée par *Rfx6/RFX6*, qu'il serait intéressant de confirmer. Le marqueur entéroendocrine *Reg4/REG4* est surexprimé dans les modèles murins et sous-exprimé dans les HIOs *RFX6<sup>-/-</sup>*, indiquant probablement un remaniement des sous-types entéroendocrines différent suite à l'inactivation de *Rfx6/RFX6* chez la souris et chez l'homme. *Scn3a/SCN3A* est lui, surexprimé dans les trois modèles, et code les canaux sodiques *NaV1.3* qui sont importants dans l'excitabilité des cellules entérochromaffines et dans la sécrétion de sérotonine (Strege et al., 2017), traduisant là encore d'un possible rôle de *RFX6* dans la différenciation des cellules entérochromaffines et/ou dans la sécrétion de sérotonine. *Slc12a5/SLC12A5*, qui est différentiellement exprimé dans chacun des modèles, code le co-transporteur neuronal *K<sup>+</sup>Cl<sup>-</sup> 2*, ayant pour fonction de moduler la sécrétion d'insuline dans les cellules  $\beta$  (Kursan et al., 2017) et qui est également exprimé dans d'autres cellules endocrines, laissant imaginer là encore, un rôle dans la sécrétion d'autres hormones dans les cellules entéroendocrines.

Au vu de la présence de ces gènes différentiellement exprimés communément aux trois modèles étudiés, il semble prudent d'affirmer que *RFX6* joue un rôle dans la différenciation entéroendocrine comme c'est le cas chez la souris, et/ou dans la fonction entéroendocrine chez l'homme. Parmi les gènes communément régulés dans les HIOs et les souris *Rfx6<sup>-/-</sup>* E18, on retrouve les marqueurs de cellules entéroendocrines *SYP/Syp* et *SCG2/Scg2* qui sont légèrement sous-exprimés chez les souris *Rfx6<sup>-/-</sup>* E18, alors qu'une augmentation importante est retrouvée dans les HIOs. C'est également le cas dans l'analyse comparative des HIOs et des souris *Rfx6<sup>ΔAdInt</sup>*, dans laquelle le marqueur entéroendocrine *CHGB/ChgB* est communément surexprimé, alors que *CHGA/ChgA* n'est sous-exprimé que dans les HIOs. Ces observations sont également en faveur d'un remaniement des sous-types entéroendocrines différents chez l'homme et la souris, provoqué par l'absence de *RFX6/Rfx6*. De manière intéressante, *NEUROD1/NeuroD1* est surexprimé à la fois dans les HIOs *RFX6<sup>-/-</sup>* et les souris *Rfx6<sup>ΔAdInt</sup>*. Il y a donc un réel impact de l'inactivation de *RFX6* sur les cellules entéroendocrines, même si de façon surprenante, les modulations de l'expression des différents marqueurs sont souvent opposées entre les différents modèles. Une explication possible serait que les sous-types affectés par la mutation soient différents entre l'homme et la souris et que certains marqueurs généraux des cellules entéroendocrines soient en réalité plus spécifiques de certains sous-types et permettent donc de mettre ces différences en évidence. La différence de contexte (*in vitro/in vivo* et foetal/adulte) peut également être la cause de telles divergences.

En conclusion, plusieurs gènes sont communément affectés par l'inactivation de *RFX6/Rfx6* dans les différents modèles et bien que l'expression de certains d'entre eux ne soit pas régulée dans le même sens dans les modèles murins et les HIOs, il semblerait que certaines cibles de *RFX6/Rfx6* soient conservées d'une espèce à l'autre. D'après l'ensemble des gènes dont l'expression est modulée par l'inactivation de *RFX6/Rfx6* dans les différents modèles, les pistes évoquées quant à un rôle de *RFX6* dans la différenciation entéroendocrine et dans la fonction de ces cellules semblent pertinentes. Toutefois, il est important de noter que les données comparées sont à la fois des expériences de RNA-seq et des données de puces à ADN (Agilent), ce qui peut biaiser le nombre de gènes différentiellement régulés en commun dans certaines comparaisons. De plus, les modèles murins sont des modèles *in vivo*, à l'inverse de notre système qui est un modèle humain *in vitro*, dépourvu de contexte physiologique (circulation sanguine, système nerveux, microbiote, système immunitaire, alimentation) qui peut grandement influencer le transcriptome et ainsi, expliquer certaines des différences observées entre les modèles.

### 3.8. Conclusions et points de discussion relevés

En conclusion, les HIOs *RFX6*<sup>-/-</sup> et CTRL expriment des marqueurs d'épithélium intestinal à des niveaux similaires, attestant de la nature des tissus générés après différenciation *in vitro* des hiPSCs. Les différents clones *RFX6*<sup>-/-</sup> générés semblent bien déficients pour *RFX6* et donc propices à l'étude du rôle de ce gène dans le développement intestinal. Un résultat majeur de cette expérience est que l'inactivation de *RFX6* induit des modifications importantes de l'expression d'un grand nombre de marqueurs entéroendocrines. Ces changements dans la régulation transcriptionnelle semblent relativement spécifiques aux cellules entéroendocrines, bien que le faible niveau d'expression de la plupart des hormones nous empêche de déterminer les types cellulaires précis affectés par la mutation. L'expression de *CHGA*, par exemple, est drastiquement réduite chez les mutants et est associée à la diminution de *REG4*. Effectivement, il a été montré que *REG4* et *CHGA* co-localisent avec la sérotonine dans tous le tractus gastro-intestinal (Heiskala et al., 2013). La faible expression de *TPH1* ne nous permet pas de conclure quant à une diminution de sérotonine ou du nombre de cellules entérochromaffines sécrétrices de sérotonine. Par contre, un certain nombre de gènes impliqués dans le métabolisme du tryptophane, précurseur de la sérotonine et de la mélatonine, sont sous-exprimés dans les HIOs mutants. De plus, le gène codant l'enzyme de biosynthèse de l'histamine *HNMT*, produite par les cellules ECL de l'épithélium gastrique, est fortement sous-exprimé. Il semble donc que l'absence de *RFX6* pourrait avoir un impact particulier sur les cellules EC et ECL et que du tissu gastrique soit notamment formé dans nos HIOs.

Etrangement, beaucoup de marqueurs entéroendocrines (granines, convertases de pro-hormone...) ainsi que des facteurs de transcription impliqués dans la régulation transcriptionnelle de certains sous-types entéroendocrines sont surexprimés dans les mutants, indiquant une augmentation potentielle du nombre de cellules de ces sous-types entéroendocrines ou de la production de leurs hormones associées. En accord avec la souris *Rfx6*<sup>ΔAdInt</sup>, l'augmentation de l'expression de *INSM1*, *NEUROD1* et *NGN3* est également observée, suggérant une certaine conservation de l'interaction de RFX6 avec ces facteurs de transcription entre l'homme et la souris. A contrario, *ARX* est drastiquement surexprimé dans les HIOs *RFX6*<sup>-/-</sup> tandis qu'il est absent chez les souris mutantes. Il se pourrait donc que la « balance » *ARX/PAX4* ne soit pas régulée par RFX6 de la même manière chez l'homme et chez la souris. Pourtant, des expériences de RT-qPCR sur des HIOs dont l'expression de *ARX* a été altérée grâce à des vecteurs lentiviraux exprimant des shRNAs dirigés contre ce gène, ont démontré un rôle conservé de *ARX* chez l'homme et la souris (Du et al., 2012). Evidemment, nos résultats ne sont qu'hypothétiques et il faudra amplifier l'effectif des échantillons, ainsi que confirmer ceci par d'autres techniques complémentaires à l'expérience de RNA-seq (Immunofluorescences, RT-qPCR...) afin de décortiquer le rôle précis de RFX6 dans la différenciation entéroendocrine chez l'homme.

L'expression de RFX6 est conservée dans les cellules entéroendocrines chez la souris à l'âge adulte et ceci pourrait expliquer un rôle dans la fonction de ces cellules, comme c'est le cas dans les cellules  $\beta$  pancréatiques. L'analyse fonctionnelle des résultats de RNA-seq semble mettre en évidence un rôle de RFX6 dans la régulation du métabolisme, en particulier du glucose, important pour la sécrétion des hormones. Toutefois, il est important de rappeler que les cellules entéroendocrines ne représentent que 1% des cellules de l'épithélium intestinal et qu'il est donc difficile de détecter des variations d'expression génique. Ainsi, l'ampleur de l'altération du métabolisme du glucose observée suggère que ces voies métaboliques sont affectées dans plusieurs types cellulaires. Or, cela semble peu probable au vu de l'expression restreinte de *Rfx6* dans les cellules entéroendocrines chez la souris. Il est alors indispensable de réussir à identifier la localisation de son expression au sein des HIOs.

D'après nos expériences, RFX6 semble donc essentiel à la différenciation des cellules entéroendocrines puisque son absence provoque un remaniement des sous-types entéroendocrines, partiellement conservé entre l'homme et la souris et avec des spécificités apparentes dans notre modèle *in vitro*. Un rôle dans la fonction des cellules entéroendocrines est aussi envisageable puisque la délétion de *RFX6* provoque une perturbation de certaines voies métaboliques, dont celle du glucose, notamment importante à la sécrétion des hormones incrélines.

# **DISCUSSION**

# Discussion

## 1. Discussion des résultats expérimentaux

### 1.1. Génération des modèles de perte de fonction dans les hiPSCs : difficultés et limites

#### *Récapitulatif des clones générés par les deux stratégies CRISPR/Cas9*

En résumé, trois lignées cellulaires respectivement déficientes pour les gènes *NGN3*, *RFX6* et *ARX* ont pu être générées par la technologie CRISPR/Cas9, tandis que l'inactivation de *PAX4* n'a pas fonctionné. Les raisons pour lesquelles aucun clone *PAX4*<sup>-/-</sup> n'a pu être obtenu ne sont pas clairement identifiées. La différence majeure entre ce gène et les trois autres réside dans sa complexité, en effet, il en existe plusieurs isoformes (**Figure 41D**), ce qui peut compliquer la mise en place de la stratégie CRISPR/Cas9. Une autre hypothèse serait que l'absence de *PAX4* soit létale pour les hiPSCs, bien que cela n'ait jamais été mis en évidence. Il est néanmoins intéressant de noter qu'aucune lignée de cellules souches avec perte de fonction de *PAX4* n'a été décrite dans la littérature, posant la question de la faisabilité de cette approche, d'autant plus qu'à l'inverse, plusieurs études de gain de fonction de *PAX4* ont été réalisées dans des cellules souches embryonnaires de souris et humaines (Gage et al., 2014 ; Blyszczuk et al., 2003).

Le clone hétérozygote *ARX*<sup>+/-</sup> (clone 2), le clone *ARX*<sup>-/-</sup> (clone 3), ainsi que les clones *RFX6*<sup>-/-</sup> (clones 1, 14 et 18) et les clones *NGN3*<sup>-/-</sup> (clones 2 et 10) ont été obtenus par l'approche CRISPR/Cas9 visant la délétion de larges portions des gènes par l'utilisation de la Cas9 sauvage. Seul le clone *NGN3*<sup>-/-</sup> 49Cd, dont la perte de fonction semble partielle, a été généré avec l'approche CRISPR/Cas9 de double-*nickase*. Ainsi, dans le cadre de nos expériences, la stratégie provoquant la délétion d'exons entiers des gènes s'est avérée plus efficace que l'approche de double-*nickase* et a rendu l'identification de clones mutants plus aisée.

#### *Limites de la technologie CRISPR/Cas9*

Bien qu'il ait été décrit que les effets non-ciblés de la technique CRISPR/Cas9 étaient plus limités avec l'approche de double-*nickase* qu'avec la stratégie utilisant la Cas9 sauvage, des mutations à des loci non-ciblés sont envisageables dans tous les cas (Ran et al., 2013). C'est donc avec précaution que les clones générés doivent être

analysés. Une publication récente a mis en évidence les conséquences importantes des mutations non-ciblées, qui confèrent de réelles limites à ce système (Kosicki et al., 2018). En effet, en plus des mutations aléatoires dans le génome précédemment reportées (Boroviak et al., 2017 ; Canver et al., 2014), des mutations non souhaitées à proximité du locus cible ont été décrites. Ces dommages à l'ADN pourraient avoir d'importantes conséquences sur l'intégrité du génome. Dans ce contexte, il serait judicieux de procéder à une caractérisation complète du génome des clones mutants générés avec CRISPR/Cas9, notamment en analysant leur caryotype et en réalisant un séquençage complet, afin de s'assurer de l'absence d'effets non-ciblés potentiellement délétères.

## 1.2. Phénotype des HIOs *NGN3*<sup>-/-</sup>

### ***Les HIOs NGN3<sup>-/-</sup> ne présentent pas de cellules CHGA<sup>+</sup>***

La délétion de *NGN3* semble avoir un impact sur la différenciation entéroendocrine puisqu'aucune cellule *CHGA*<sup>+</sup> n'est détectée, et une diminution de plus de 50% du niveau de transcrite est observée. Bien qu'il soit surprenant de constater une absence totale de cellules *CHGA*<sup>+</sup> alors que son ARNm semble encore exprimé, il a déjà été mis en évidence que le niveau d'expression des ARNm n'est pas forcément corrélé avec l'expression protéique des hormones (Beucher et al., 2012 ; Mortensen et al., 2003). L'absence de cellules *CHGA*<sup>+</sup> est en accord avec l'étude sur les HIOs exprimant des shRNAs dirigés contre *NGN3*, et dans laquelle ils concluent à une absence totale de cellule entéroendocrine (Spence et al., 2011). De plus, une absence de cellules entéroendocrines a également été notée chez la souris déficiente pour *Ngn3*, coïncidant avec la perte d'expression des hormones intestinales (Mellitzer et al., 2010 ; Jenny et al., 2002; Lee et al., 2002). Le modèle murin présentant une délétion de *Ngn3* spécifiquement dans l'intestin (*Ngn3*<sup>Δint</sup>) reproduit partiellement la pathologie humaine causée par des mutations ponctuelles dans la séquence codant le domaine de liaison à l'ADN de *NGN3*, qui se caractérisent par la perte totale des cellules entéroendocrines menant à une malabsorption intestinale et une diarrhée chronique (Hancili et al., 2018; German-Diaz et al., 2017; Sayar et al., 2013; Pinney et al., 2011; Rubio-Cabezas et al., 2011 ; Cortina et al., 2007; Wang et al., 2006). L'absence de cellules *CHGA*<sup>+</sup> et la diminution d'expression des marqueurs entéroendocrines *SCG2*, *VGF* et *CHGA* dans les HIOs *NGN3*<sup>-/-</sup> suggèrent un impact de l'inactivation de *NGN3* sur les cellules entéroendocrines et est en accord avec les observations faites sur les biopsies des patients atteints de diarrhées malabsorptives congénitales causées par une mutation dans *NGN3*, montrant l'absence totale de cellules entéroendocrines sécrétrices d'hormones.

### **L'expression des facteurs de transcription et hormones entéroendocrines ne semble pas affectée dans les HIOs NGN3<sup>-/-</sup>**

Etrangement, dans notre modèle *in vitro*, l'expression des gènes codant les hormones, du moins ceux qui ont pu être détectés, ne semble pas affectée par la mutation de *NGN3*. Il est également surprenant de constater que l'expression des gènes codant des facteurs de transcription impliqués dans la régulation de la différenciation entéroendocrine n'est pas affectée dans les HIOs NGN3<sup>-/-</sup>. Une hypothèse envisagée est que le niveau d'expression des gènes codant les hormones soit trop faible pour déceler une différence d'expression entre les mutants et les contrôles, mais ceci n'est pas vrai dans le cas des facteurs de transcription, dont l'expression est largement détectée dans les HIOs. Puisque le niveau des transcrits *CHGA* n'est pas complètement nul mais qu'aucune cellule *CHGA*<sup>+</sup> n'est détectée, on peut envisager un schéma similaire concernant les gènes codant les hormones qui sont détectés, à savoir qu'elles ne sont pas produites mais que les cellules contiennent tout de même les transcrits. En effet, il a été montré dans des études sur des cellules entéroendocrines isolées, qu'une cellule entéroendocrine exprimant une ou deux hormones est capable d'exprimer jusqu'à six transcrits d'hormones différentes (Egerod et al., 2012). Pour évaluer cette hypothèse, il sera indispensable de réussir à imager l'expression hormonale par immunomarquage, ce que nous n'avons pas réussi à faire à l'heure actuelle dans nos conditions expérimentales. Il est envisageable que les anticorps utilisés ne soient pas assez sensibles ou que le protocole ne soit pas optimal. Ainsi, il serait intéressant de tester des anticorps différents et/ou d'appliquer des techniques d'amplification de signal comme le TSA (*Tyramide Signal Amplification*) par exemple, afin de détecter l'expression hormonale dans nos HIOs.

### **Implication de NGN3 dans le lignage sécréteur ?**

D'après les analyses par RT-qPCR, il semblait que toutes les lignées sécrétrices soient affectées par la mutation (diminution significative de l'expression de *CHGA*, *LYZ* et *MUC2*), ce qui est contredit par les données de RNA-seq puisque l'expression des marqueurs des cellules de Paneth et des cellules à mucus n'est pas altérée par la mutation. Ces différences peuvent être liées au fait que tous les échantillons analysés par séquençage ne sont pas ceux étudiés par RT-qPCR et/ou qu'une des approches soit plus sensible en terme de détection du niveau d'expression des ARNm que l'autre. Il serait alors intéressant d'analyser l'expression de ces marqueurs par immunofluorescence et de quantifier le nombre de cellules des différentes lignées cellulaires pour déterminer si *NGN3* a un rôle dans la différenciation du lignage sécréteur ou non. Toutefois, au vu de la littérature, un tel phénotype n'est pas attendu. En effet, les modèles murins déficients pour *Ngn3* et les patients présentant des mutations de *NGN3* ne présentent globalement pas d'altérations de ces types cellulaires.

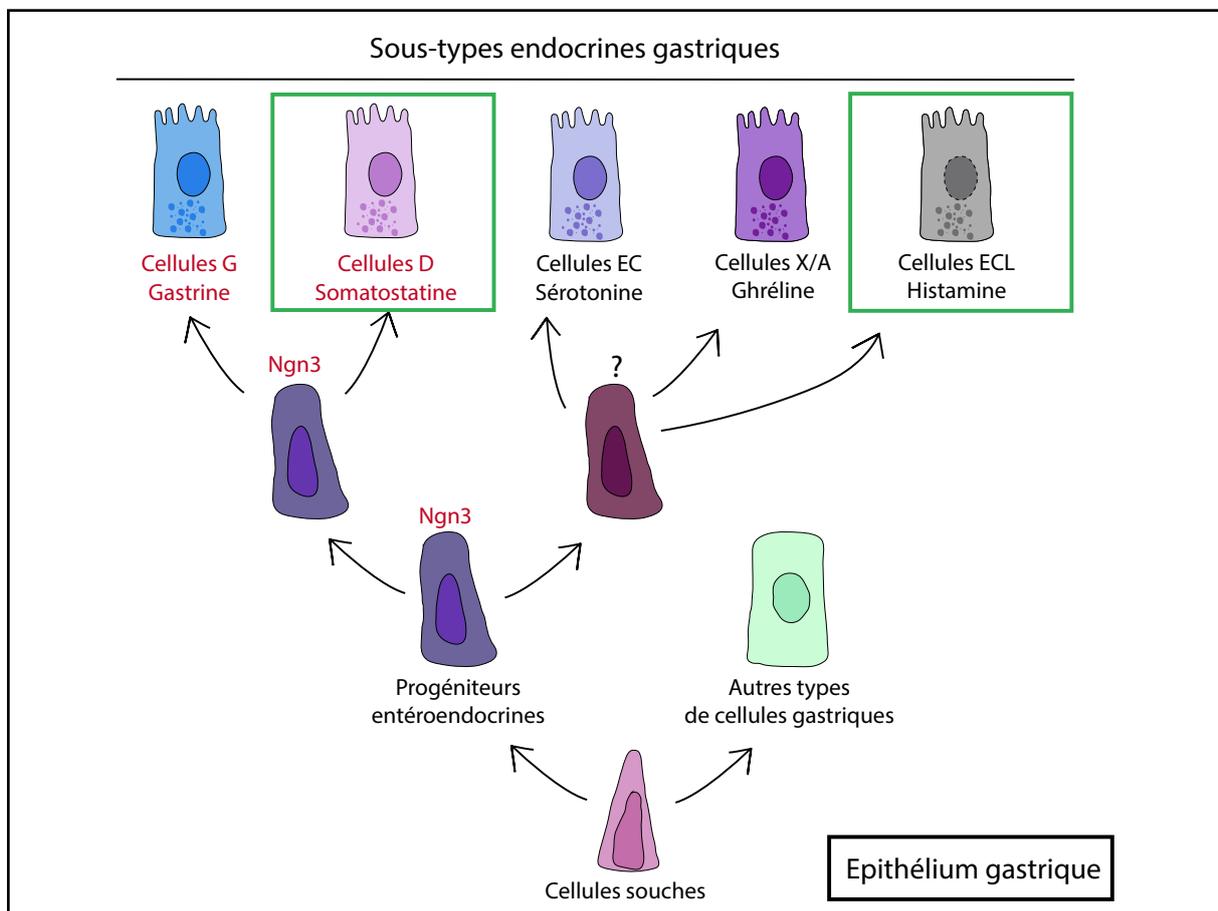
### **Compensation partielle de l'absence de NGN3 ?**

Contrairement à la souris, où *Ngn3* contrôle la différenciation endocrine dans le pancréas (Gradwohl et al., 2000), *Ngn3* n'est pas exprimé dans le pancréas du poisson-zèbre et le contrôle de cette différenciation est assuré par les facteurs de transcription *Ascl1b* et *NeuroD1* (Flasse et al., 2013), tous deux exprimés dans nos HIOs. *Ngn3* n'est donc pas l'unique déterminant des lignées endocrines chez les vertébrés et il peut alors être envisagé que d'autres facteurs de transcription puissent réguler la différenciation des cellules entéroendocrines chez l'homme et ainsi, partiellement compenser la perte de *NGN3*. De plus, il a été montré que les facteurs de transcription à bHLH *Ngn3* et *NeuroD1* possèdent une redondance de fonction (Huang et al., 2000) qui pourrait ainsi expliquer le maintien de l'activation des cascades de régulation transcriptionnelle en aval de *NGN3* malgré l'inactivation de celui-ci dans nos HIOs. Cette hypothèse semble toutefois peu probable au vu du phénotype présent chez les patients porteurs de mutation dans *NGN3*.

### **Les clones *NGN3*<sup>-/-</sup> se différencient en tissu épithélial intestinal de nature proximale plutôt que distale**

L'hypothèse privilégiée pour expliquer l'absence du phénotype attendu est en faveur d'une différenciation en tissu de nature plus proximale (épithélium gastrique et/ou intestinal proximal) des clones *NGN3*<sup>-/-</sup> en comparaison aux contrôles. Premièrement, une augmentation presque significative de *PDX1* ( $p=0,052$ ) est observée chez les mutants, gène qui est plutôt exprimé dans les parties proximales du tube digestif : le pancréas, l'estomac et le duodénum (Guz et al., 1995). De plus, d'après les analyses de l'expérience de RNA-seq, un certain nombre de gènes parmi les plus significativement sous-exprimés appartiennent à la famille de gènes *HOX* qui jouent un rôle dans le développement du tractus digestif, notamment dans la mise en place de l'axe proximo-distal (Zorn et Wells, 2009). Dans notre cas, les gènes *HOX* sous-exprimés (*HOXA10*, *A11*, *A13*, *D10*, *D11* et *D13*) semblent correspondre à ceux exprimés dans la partie distale du tube digestif (gros intestin, côlon, rectum), favorisant potentiellement le développement des parties proximales (Mùnera et al., 2017). L'analyse comparative des marqueurs proximaux et distaux a permis de valider cette observation (**Figure 59**). La nature proximale des tissus générés dans les HIOs *NGN3*<sup>-/-</sup> pourrait ainsi expliquer l'absence du phénotype attendu, puisque *Ngn3* n'est indispensable qu'à la différenciation des cellules G sécrétrices de gastrine et D sécrétrices de somatostatine dans l'estomac des souris (Jenny et al., 2002). Effectivement, dans les HIOs, comme chez la souris, l'expression de *HNMT* (gène codant l'enzyme de biosynthèse de l'histamine) n'est pas affectée par l'inactivation de *NGN3*. Toutefois, l'expression de *SST* est inchangée dans les HIOs mutants, ce

qui suggère tout de même une fonction différente de *NGN3* chez l'homme, bien qu'il soit encore nécessaire de vérifier l'expression de la protéine associée (**Figure 78**). L'hypothèse de la formation de tissu plus proximal exige une confirmation de l'expression des marqueurs proximaux observés en RNA-seq, par RT-qPCR et immunofluorescence. Même si ces observations venaient à être confirmées par les autres approches, un lien de causalité avec la mutation dans *NGN3* resterait à établir. Pour cela, il serait nécessaire de rétablir la fonction de *NGN3* dans les clones hiPSCs *NGN3*<sup>-/-</sup> afin de vérifier que cela annule le phénotype « proximal » des HIOs générés. Toutefois, comme cette observation est propre aux HIOs *NGN3*<sup>-/-</sup>, la plupart des marqueurs proximaux et distaux ne présentant pas de variation d'expression dans les HIOs *RFX6*<sup>-/-</sup> (données non montrées), l'hypothèse d'un effet du génotype *NGN3*<sup>-/-</sup> sur la régionalisation du tissu formé, reste probable.



**Figure 78. Rôle de *Ngn3* dans la différenciation endocrine de l'épithélium gastrique chez la souris**

*Ngn3* est nécessaire à la différenciation des cellules G sécrétrices de gastrine et D sécrétrices de somatostatine dans l'épithélium gastrique de souris, alors que la différenciation des cellules à sérotonine, histamine et ghréline semble dépendante d'un autre facteur encore inconnu. A contrario, l'inactivation de *NGN3* dans les HIOs ne semble pas altérer la différenciation des cellules à somatostatine (encadré vert). Encadrés verts: expression de *SST* et *HNMT* (enzyme de biosynthèse d'histamine) inchangés. (Adapté de Jenny et al., 2002 ; Lee et al., 2002)

**Le nombre de passage des hiPSCs, un potentiel biais expérimental ?**

Bien qu'aucune donnée ne permette d'affirmer que la qualité de la différenciation dépende du nombre de passage des cellules de départ, il est important de noter que les HIOs NGN3<sup>-/-</sup> ont été générés avec des clones d'un nombre de passage plus élevé (P60) que les HIOs CTRL (P30). L'accumulation des passages peut provoquer des altérations du caryotype des hiPSCs et en affecter leur pluripotence. L'expression de marqueurs de pluripotence a été vérifiée par RT-qPCR sur les clones NGN3<sup>-/-</sup> et RFX6<sup>-/-</sup> et aucune différence d'expression n'a été observée avec les hiPSCs contrôles. Le nombre de passage des cellules ne semble donc pas influencer sur l'expression de ces marqueurs. Seul le caryotype des hiPSCs contrôles a été analysé et il faudrait également analyser celui des clones mutants. Dans ce contexte, en vue d'analyses futures, de nouveaux clones NGN3<sup>-/-</sup> ont été générés à partir de cellules d'un faible nombre de passage (P25), ainsi que dans une autre lignée cellulaire afin de vérifier la reproductibilité de nos observations.

**Prolifération cellulaire aberrante et anomalies morphologiques associées à la mutation ?**

De manière intéressante, l'analyse fonctionnelle des gènes différentiellement exprimés dans les HIOs NGN3<sup>-/-</sup>, révèle la présence de gènes impliqués dans les processus de développement et morphogenèse de l'intestin et dans la prolifération cellulaire. L'altération de leur expression provoque des anomalies morphologiques du tube digestif et une prolifération cellulaire accrue que l'on peut associer au phénotype observé chez les souris déficientes pour *Ngn3* spécifiquement dans l'intestin, qui possèdent des cryptes élargies dues à une prolifération aberrante mais également des villosités plus courtes à cause d'une augmentation du renouvellement cellulaire (Mellitzer et al., 2010). De plus, l'intestin grêle de certains patients possédant une mutation dans *NGN3* présente une atrophie des villosités et une hyperplasie des cryptes (Pinney et al., 2011). Il serait donc intéressant d'étudier l'expression de marqueurs de prolifération, comme Ki67, mais également d'étudier la structure histologique des HIOs NGN3<sup>-/-</sup>.

De façon générale, les altérations observées sont moindres en comparaison à ce qui était attendu au vu de la littérature existante. Il ne peut être exclu que ces différences soient dues à un certain nombre de limitations techniques et inhérentes au modèle qui seront discutées par la suite. Une limite évidente de cette étude est le faible effectif d'échantillons, notamment des contrôles sauvages et la possible présence de tissus de nature autre qu'intestinale (tissu gastrique, mésenchyme, etc.) qui pourraient masquer de potentiels variations d'expression des gènes au sein du tissu intestinal.

Enfin, nous ne disposons malheureusement pas de données de séquençage d'ARN sur d'autres modèles inactivés pour *NGN3* qui permettraient d'établir une comparaison de gènes différentiellement exprimés.

### 1.3. Potentiels rôles de *RFX6* dans les cellules intestinales chez l'homme

#### **La régulation de la différenciation entéroendocrine est affectée par la perte de *RFX6***

L'expression de la plupart des marqueurs généraux de cellules entéroendocrines est affectée suite à la délétion de *RFX6*. Comme pour les analyses des HIOs *NGN3*<sup>-/-</sup>, il est frustrant de voir que l'expression de beaucoup de gènes codant les hormones ou leur enzyme de biosynthèse est très faible, empêchant ainsi de déterminer avec précision les sous-types entéroendocrines affectés par la mutation.

Néanmoins, dans les HIOs *RFX6*<sup>-/-</sup>, l'expression de *HNMT*, *MLN*, *GHRL*, *SCT* et dans une moindre mesure *NTS* est diminuée par rapport aux HIOs CTRL. En revanche l'expression de *VIP* est augmentée, ainsi que celle de *TAC1* ( $p=0,06$ ) et *SST* ( $p=0,064$ ). Concernant la régulation transcriptionnelle de la différenciation entéroendocrine, l'expression de *NEUROD1*, *ARX*, *INSM1* et *LMX1A* est augmentée chez les mutants. Bien qu'il soit difficile de mettre en relation ces observations avec la littérature existante, il semble évident que la régulation de la différenciation entéroendocrine soit affectée par la mutation. L'ensemble des gènes d'intérêt affecté par la perte de *RFX6* est représenté sur le schéma récapitulatif en **figure 79**.

#### ***RFX6* semble affecter la différenciation et/ou fonction des cellules entérochromaffines et entérochromaffines-like**

La diminution de l'expression de gènes codant des enzymes impliquées dans le métabolisme du tryptophane, précurseur de la sérotonine, suggère un impact particulier de la mutation de *RFX6* sur les cellules entérochromaffines (EC) productrices de sérotonine. Cette hypothèse est renforcée par la diminution de l'expression de *CHGA* et la baisse non-significative de *TPH1*, observée seulement par RT-qPCR, le niveau d'expression étant très faible dans les expériences de RNA-seq. Ces résultats sont inattendus puisque dans le modèle murin *in vivo* l'ensemble des sous-types entéroendocrines est affecté par la mutation de *Rfx6*, à l'exception des cellules EC. Bien que ces résultats soient préliminaires, ils semblent correspondre aux observations faites dans l'analyse du modèle murin *in vitro*, qui présentent une diminution de *Tph1* et de *ChgA* (**Figure 72**). En regard de ces observations, il est surprenant d'observer une augmentation de l'expression de *LMX1A*, ainsi qu'une augmentation non-significative de *NKX2.2*. En effet, ces facteurs de transcription régulent la différenciation des cellules EC et leur niveau d'expression est positivement corrélé avec l'expression de *Tph1* et *ChgA* chez la souris (Gross et al., 2016).

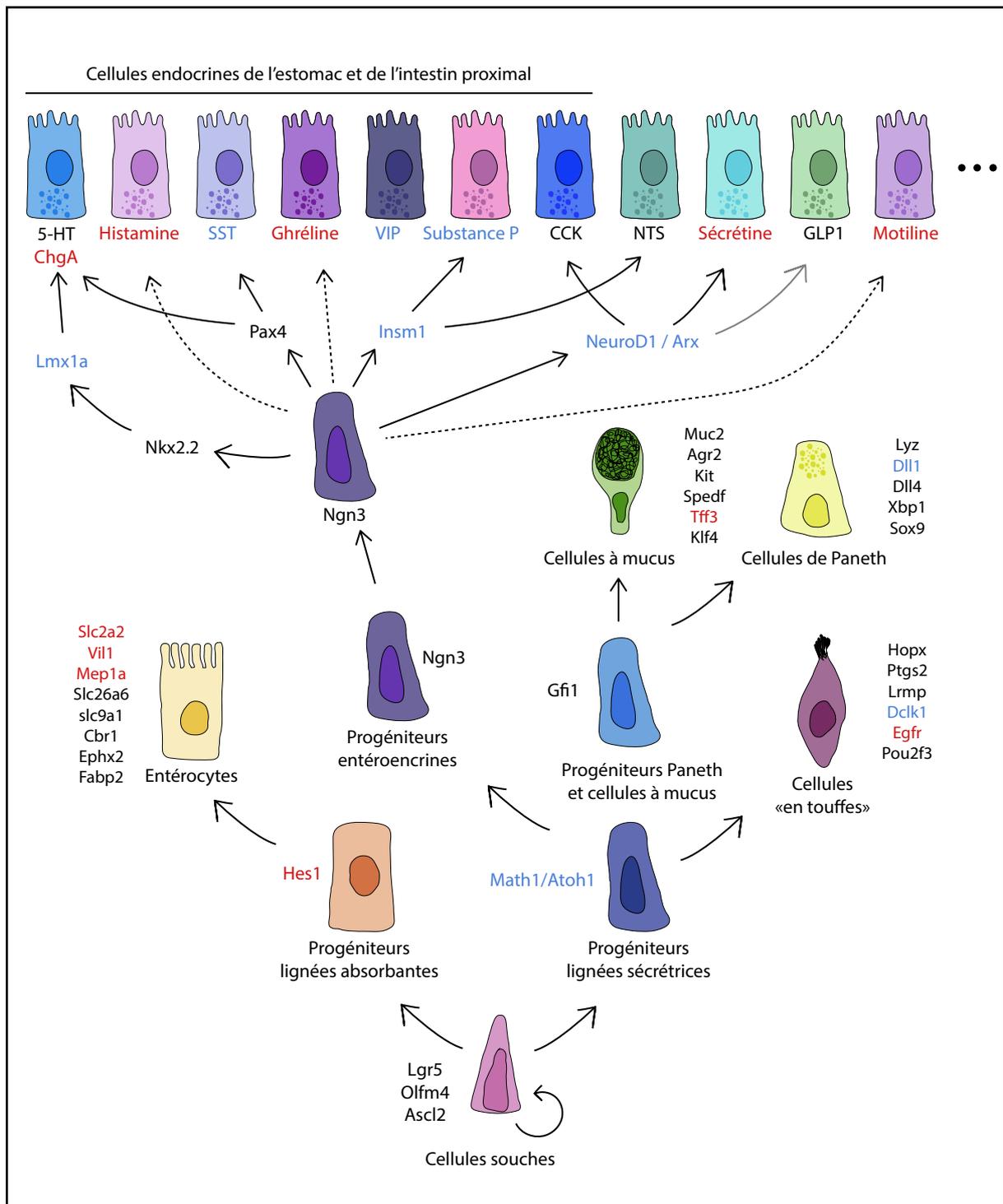
D'autre part, la forte augmentation de l'expression de l'enzyme de biosynthèse de l'histamine *HNMT* indique un effet de la mutation sur la formation des cellules ECL. Ces cellules étant normalement retrouvées dans l'épithélium gastrique, il semblerait que les HIOs soient notamment formés de tissu épithélial gastrique. Pourtant, à l'inverse des clones *NGN3<sup>-/-</sup>*, les clones *RFX6<sup>-/-</sup>* ne se différencient pas en tissu de nature plus proximale que les hiPSCs contrôle puisque les marqueurs proximaux et distaux ont une expression similaire entre les génotypes (non montré). Cette observation est donc plutôt en faveur de la présence d'un mélange de tissu épithélial mixte (estomac, intestin grêle et gros intestin) dans les HIOs formés dans nos conditions expérimentales. De plus, la possible incidence de la mutation sur la différenciation des cellules ECL indiquerait un rôle de *RFX6* dans la différenciation des cellules endocrines de l'épithélium gastrique.

### ***Le rôle précis de RFX6 dans la différenciation entéroendocrine reste à déterminer***

La fonction de *Rfx6* dans la différenciation entéroendocrine n'est pas décrite dans la littérature et seule son implication dans la régulation de l'expression et de la sécrétion de GIP par les cellules K a été mise en évidence dans une lignée cellulaire intestinale de souris *STC-1* (Suzuki et al., 2012). Le gène codant GIP n'étant pas détecté dans notre modèle, il nous est impossible de confirmer cette interaction dans les HIOs. Il semblerait tout de même y avoir des fonctions conservées de *Rfx6/RFX6* entre l'homme et la souris puisque les gènes codant les hormones GHRL et SCT, par exemple, ont une expression diminuée par la mutation dans les modèles murins et les HIOs. Au vu des analyses des HIOs *RFX6<sup>-/-</sup>*, la différenciation entéroendocrine est incontestablement affectée par la mutation. Malheureusement, il n'est pas possible à l'heure actuelle de connaître le rôle exact de *RFX6* dans ce processus chez l'homme. Comme mentionné précédemment, cela nécessite des expériences supplémentaires de marquages des hormones par immunofluorescence par exemple, mais également d'augmenter les effectifs analysés pour consolider les données acquises. Il est aussi primordial de réussir à mettre au point un protocole d'immunomarquage permettant la visualisation de l'expression de *RFX6* au sein des HIOs, afin de ne pas se limiter à une transposition des données obtenues chez la souris. Les limites de ces analyses relatives au modèle d'étude seront évoquées dans une partie dédiée.

### ***Implication de RFX6 dans la formation des progéniteurs sécréteurs et absorbants ?***

De façon inattendue, la différenciation des progéniteurs du lignage sécréteur et du lignage absorbant paraît affectée puisque l'expression de *MATH1/ATOH1*, marqueur des progéniteurs de toutes les lignées sécrétrices est augmentée et l'expression de *HES1*, marqueur des progéniteurs des entérocytes est diminuée.



**Figure 79. Schéma récapitulatif des conséquences de l'inactivation de *RFX6* sur la différenciation endocrine au sein des HIOs**

L'inactivation de *RFX6* dans les HIOs affecte les niveaux d'expression de nombreux gènes codant des facteurs de transcription impliqués dans la différenciation des cellules endocrines productrices d'hormones, qui sont indiqués de façon non-exhaustive sur ce schéma. Les cellules endocrines de l'estomac et/ou de l'intestin proximal sont indiquées.

Bleu: gènes surexprimés dans les HIOs *RFX6*<sup>-/-</sup>, rouge: gènes sous-exprimés, noir: non différenciellement exprimés; flèches en pointillés: voies de différenciation dont les facteurs de transcription n'ont volontairement pas été représentés.

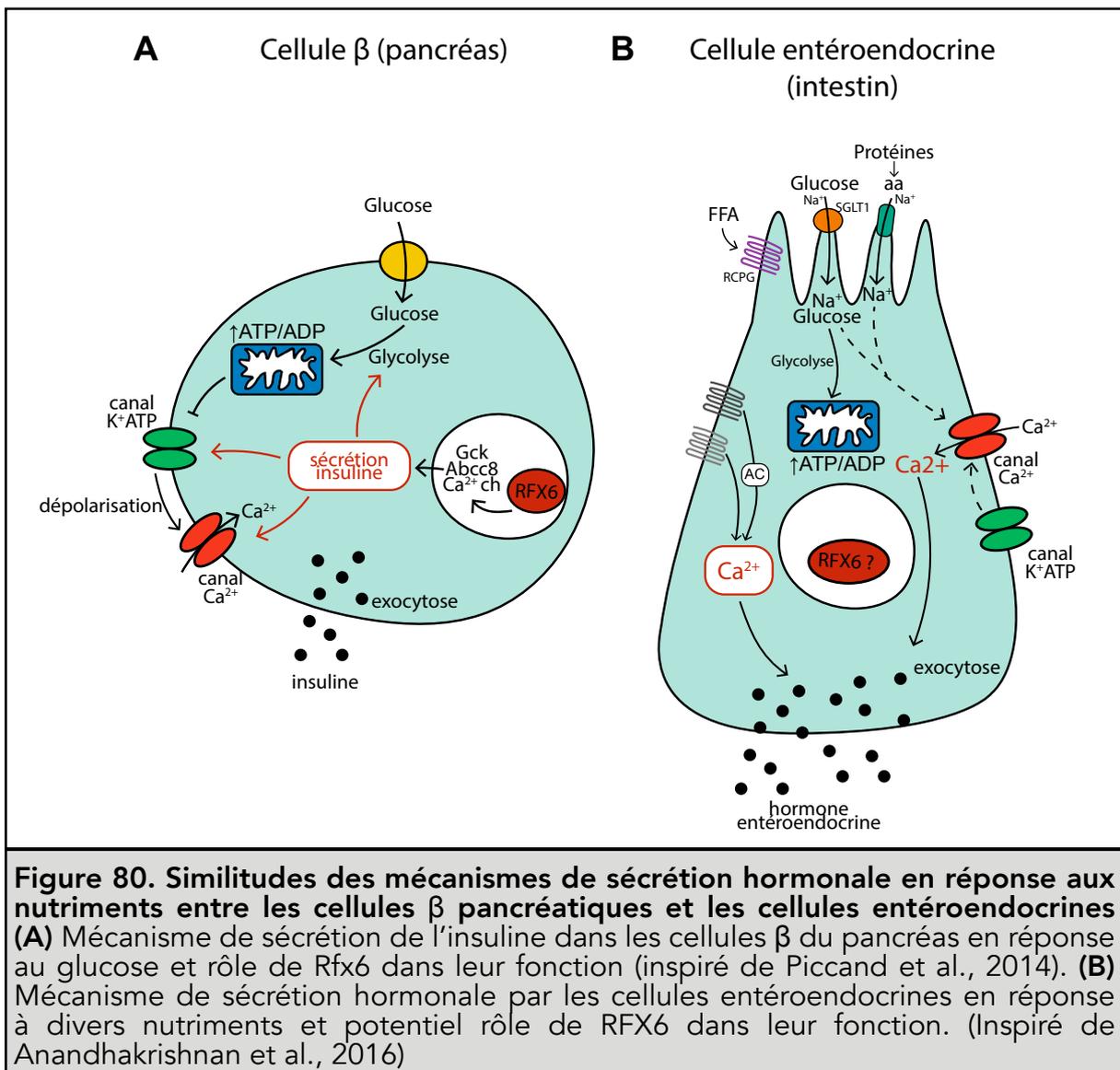
Cependant, une altération dans la différenciation de ces lignages cellulaires n'a pas été mise en évidence par l'analyse des marqueurs des différents types cellulaires de l'épithélium, bien que certains marqueurs d'entérocytes soient significativement sous-exprimés. Ceci suggère une implication de RFX6 restreinte à la spécification de ces progéniteurs. Une telle fonction de Rfx6 n'a pas été décrite dans les modèles murins, où l'expression des marqueurs des progéniteurs sécréteurs et absorbants n'est pas affectée. Cependant, il est intrigant de voir que dans les cryptes intestinales de souris, l'expression de Rfx6 n'est pas restreinte aux cellules Ngn3<sup>+</sup>, suggérant une possible expression de Rfx6 dans d'autres progéniteurs (**Figure 62B-D**). Ainsi, il faudrait confirmer nos observations par des expériences additionnelles permettant, par exemple, de quantifier le nombre de progéniteurs MATH1<sup>+</sup> et HES1<sup>+</sup> après immunomarquage, et déterminer une éventuelle co-localisation avec RFX6.

### **Vers un rôle de RFX6 dans la fonction des cellules entéroendocrines ?**

Pendant l'embryogenèse chez la souris, Rfx6 est exprimé dans l'endoderme intestinal au stade embryonnaire E9,5, puis son expression est restreinte aux progéniteurs pancréatiques et endocrines, et il est ensuite exprimé dans les cellules endocrines matures de l'embryon mais également à l'âge adulte (Soyer et al., 2010). Ce patron d'expression lui confère une implication dans le développement intestinal, qui a été étudié dans un modèle de souris déficiente pour Rfx6 (Piccand et al., manuscrit en préparation) et dont les résultats ont servi de support de comparaison et de discussion dans l'analyse des HIOs RFX6<sup>-/-</sup>, mais suggère également un rôle au sein des cellules entéroendocrines différenciées. Les analyses fonctionnelles des données de l'expérience de RNA-seq sont en faveur d'un tel rôle et plus spécifiquement dans la régulation du métabolisme énergétique. En effet, des gènes dérégulés dans les HIOs RFX6<sup>-/-</sup> sont impliqués dans le métabolisme des lipides, ce qui peut être mis en lien avec la malabsorption des lipides identifiée chez les souris Rfx6<sup>ΔAdInt</sup> (Piccand et al, manuscrit en préparation), ainsi que dans des cas humains du syndrome de Mitchell-Riley (Smith et al., 2010).

De plus, plusieurs gènes impliqués dans le métabolisme du glucose (glycolyse et gluconéogenèse intestinale) et associés à la régulation de la sécrétion d'hormones incrétones ont une expression affectée par la mutation de RFX6. Cette hypothèse est d'autant plus probable qu'il a récemment été montré par l'équipe, que Rfx6 joue un rôle crucial dans le contrôle de l'homéostasie du glucose et le maintien de l'état différencié des cellules β du pancréas par l'activation de la sécrétion d'insuline chez la souris (Piccand et al., 2014). Les cellules entéroendocrines et les cellules β possèdent un nombre important de constituants et de mécanismes cellulaires communs permettant d'aboutir à l'exocytose des granules sécrétoires contenant les hormones (**Figure 80**). On peut alors imaginer que RFX6 ait un rôle similaire dans la fonction des

cellules entéroendocrines, en plus de son rôle dans la différenciation de ces cellules. Cependant, il est important de noter que les variations dans l'expression génique de la machinerie métabolique observées, sont sûrement le résultat d'une altération commune à plusieurs types cellulaires puisque celles-ci sont massives au regard de la faible proportion théorique de cellules entéroendocrines dans les HIOs. Cette idée est crédible en regard de la variation d'expression des marqueurs des progéniteurs des lignages sécréteurs et absorbants, et la diminution de certains des marqueurs d'entérocytes. Cette hypothèse renforce la nécessité de localiser l'expression de RFX6 dans les HIOs. Comme il a été démontré que les HIOs ont la capacité de sécréter des hormones en réponse aux nutriments (Zietek et al., 2015), il serait pertinent de réaliser des tests ELISA de sécrétion hormonale en réponse au glucose dans les HIOs, afin d'explorer la fonction de RFX6 dans la physiologie des cellules entéroendocrines. Cependant, cette approche nécessite l'obtention de HIOs matures qui produisent l'ensemble des hormones entéroendocrines.



### **La mutation de *RFX6* est associée à des anomalies morphologiques du tube digestif**

L'analyse fonctionnelle des gènes différentiellement exprimés dans les HIOs *RFX6*<sup>-/-</sup>, révèle la présence de gènes impliqués dans des processus de morphogenèse intestinale, dont la déficience est associée à des anomalies intestinales morphologiques et cellulaires. Ces observations sont particulièrement intéressantes puisque dans le cas de la pathologie de Mitchell-Riley causée par des mutations de *RFX6*, le diabète néonatal est associé à de sévères malformations du tractus gastro-intestinal, comme l'atrésie, la sténose et la malrotation intestinale (Smith et al., 2010). *RFX6* étant exprimé précocement au cours du développement du tractus digestif dans l'endoderme intestinal, il n'est pas surprenant d'observer de tels défauts anatomiques intestinaux en l'absence de ce facteur de transcription. Chez les patients souffrant du syndrome de Mitchell-Riley, ces malformations sont parfois associées à des hétérotopies gastriques, consistant en la présence de tissu gastrique dans l'intestin grêle (Skopkova et al., 2016 ; Sansbury et al., 2015). Des marqueurs de tissu gastrique sont retrouvés dans les HIOs quel que soit le génotype et leur expression ne varie pas dans les HIOs *RFX6*<sup>-/-</sup>, suggérant plutôt une différenciation *in vitro* d'un tissu qui ne correspond pas strictement à de l'intestin grêle mais à un mélange de tissu proximal et distal.

### **Des marqueurs de tissus nerveux sont détectés dans les HIOs et semblent affecter par la mutation de *RFX6***

De façon intéressante, l'analyse fonctionnelle dégage un certain nombre de gènes codant pour des protéines constituant des canaux ioniques, des vésicules exocytiques et des protéines notamment impliquées dans la signalisation synaptique, qui sont autant de fonctions généralement attribuées au système nerveux. Ceci peut néanmoins traduire la présence et la modification par la mutation de *RFX6* de la machinerie moléculaire permettant aux cellules entéroendocrines de sécréter les hormones par exocytose des vésicules mais également de communiquer avec le système nerveux. Bien que le système nerveux entérique soit théoriquement absent dans les HIOs, il a récemment été montré que certaines cellules entéroendocrines possèdent des extensions de leur surface basolatérale, appelés neuropodes, qui pourraient servir d'intermédiaire dans la communication entre ces cellules et le système nerveux (**Figure 15**). De plus, par leur lien étroit avec le système nerveux entérique, ces cellules possèdent les éléments nécessaires à la neurotransmission en exprimant notamment des gènes codant des protéines pré et post-synaptiques (Bohórquez et al., 2015). Les travaux récents de Brokhman et al. (2018) ont mis en évidence que le système nerveux entérique se développe à partir des crêtes neurales mais également de l'endoderme. Ainsi, il est possible que du tissu nerveux soit involontairement formé

lors de nos expériences de différenciation *in vitro*. Cette hypothèse renforce l'idée que les HIOs sont constitués de plusieurs tissus d'origine endodermique (estomac, intestin proximal et distal et tissu nerveux ?).

### **Conclusion et perspectives quant au rôle de RFX6 dans l'intestin**

L'analyse transcriptomique globale des HIOs RFX6<sup>-/-</sup> donne un grand nombre d'indications quant à des fonctions possibles de RFX6, mais dont il est difficile de tirer des conclusions directes. La comparaison de ces données avec celles générées dans les modèles murins *in vivo* a permis d'identifier les pistes les plus pertinentes en déterminant les gènes communément régulés par Rfx6/RFX6 dans les différents modèles. Parmi les gènes dont l'expression est affectée à la fois dans les HIOs et un ou plusieurs modèles murins *in vivo*, sont retrouvés les marqueurs entéroendocrines CHGA, CHGB et REG4, les gènes codant les hormones ghréline et sécrétine, ainsi que les gènes codant les facteurs de transcription LMX1A, ARX et NEUROD1. Au vu de ces résultats, il semble manifeste que la différenciation entéroendocrine soit affectée par l'inactivation de RFX6. Néanmoins, son rôle exact au sein de la cascade de régulation de la différenciation entéroendocrine reste à déterminer par des études ultérieures.

Par ailleurs, on retrouve également des gènes comme TXNIP, NR1H4, PGK1, ABCC8 et CASR, codant des récepteurs, enzymes et effecteurs, notamment impliqués dans la physiologie des cellules entéroendocrines. Ainsi, en plus de son rôle dans la différenciation entéroendocrine, il semblerait que RFX6 ait un rôle dans la fonction de ces cellules, ce qui coïncide avec la persistance de son expression dans les cellules entéroendocrines matures chez la souris.

Cependant, les modèles murins sont des modèles *in vivo*, à l'inverse de notre système qui est un modèle humain *in vitro*, dépourvu de contexte physiologique, qui peut grandement influencer le transcriptome et ainsi expliquer certaines des différences observées entre les modèles.

Afin de caractériser le rôle de RFX6 chez l'homme et notamment les spécificités de la différenciation intestinale humaine, il faudra dans un premier temps confirmer l'ensemble de ces observations, mais également les reproduire sur des HIOs plus matures, comportant l'ensemble des sous-types entéroendocrines et exprimant toutes les hormones. Aussi, il est essentiel de caractériser le patron d'expression spatio-temporel de RFX6 au sein des HIOs, par des approches d'immunomarquages ou de traçage cellulaire, afin de confirmer le maintien de son expression dans les cellules entéroendocrines différenciées. De plus, pour identifier les cibles directes de RFX6 parmi les gènes différentiellement exprimés dans les mutants, des expériences de ChIP pourraient être menées, de même que des expériences de transactivation afin de déterminer si RFX6 régule leur transcription.

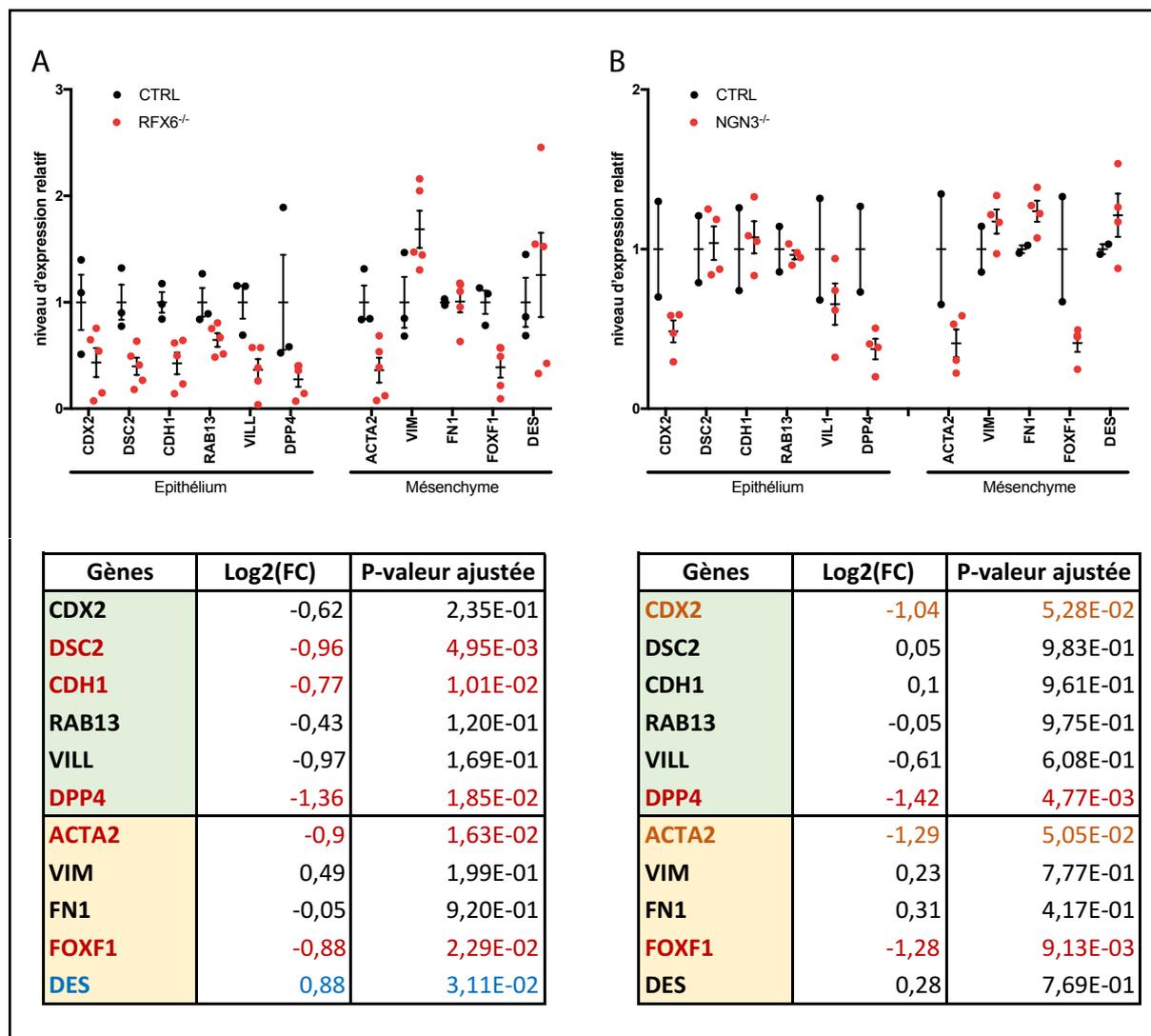
## 2. Limites du modèle HIOs

### 2.1. Variabilité de la proportion épithélium/mésenchyme au sein des échantillons

Bien que l'analyse transcriptomique des HIOs nous ait permis de révéler des différences dans l'expression des gènes dans les HIOs mutants *NGN3<sup>-/-</sup>* et *RFX6<sup>-/-</sup>* en comparaison aux HIOs CTRL, elle a également pointé du doigt une certaine variabilité de l'expression des marqueurs des tissus formés au sein des expériences de différenciation *in vitro*. En effet, les RT-qPCR ont mis en évidence une hétérogénéité dans l'expression des marqueurs d'épithélium intestinal au sein des HIOs de même génotype, qui a ensuite été retrouvée dans les analyses de RNA-seq. L'expression de marqueurs d'épithélium a été analysée plus en détail et démontre une variabilité d'expression entre les différents échantillons (**Figure 81**). Les HIOs sont entourés d'un mésenchyme qui contient des myo-fibroblastes, des cellules endothéliales et des cellules musculaires lisses (Watson et al., 2014), et qui est déterminant dans la morphogénèse, l'homéostasie et l'identité régionale de l'intestin (Kedinger et al., 1998). Une variabilité d'expression de marqueurs décrits comme étant spécifiquement exprimés dans le mésenchyme a également été mise en évidence entre les différents échantillons. Outre la variabilité inter-échantillons, un certain nombre de marqueurs épithéliaux et mésenchymateux sont différentiellement exprimés entre les HIOs mutants et CTRL (**Figure 81**). Il est donc probable que l'efficacité de différenciation entre les clones mutants et CTRL ne soit pas similaire, pouvant ainsi biaiser l'analyse sous-jacente mais également influencer l'identité régionale et le développement de l'endoderme.

### 2.2. Variabilité dans la morphologie des HIOs

L'hétérogénéité décrite par l'analyse de l'expression des gènes est également traduite par des variations morphologiques (taille, forme, densité) entre les HIOs. Des HIOs de même génotype peuvent adopter différentes formes, avec des protubérances plus ou moins marquées, qui sont forcément influencées par leur composition cellulaire (**Figure 82**). Même si chaque échantillon analysé est composé d'une quinzaine de HIOs, il semblerait qu'un tel effectif ne suffise pas à neutraliser cette variabilité, qui pourrait masquer des différences induites par le génotype. Cette variabilité est vraisemblablement inhérente au protocole de différenciation *in vitro*, dont l'efficacité dépend de différents paramètres, comme l'aspect, la densité et la qualité des hiPSCs lors de l'initiation du protocole, mais également du *timing* des changements de milieu ou encore des propriétés intrinsèques des réactifs utilisés.



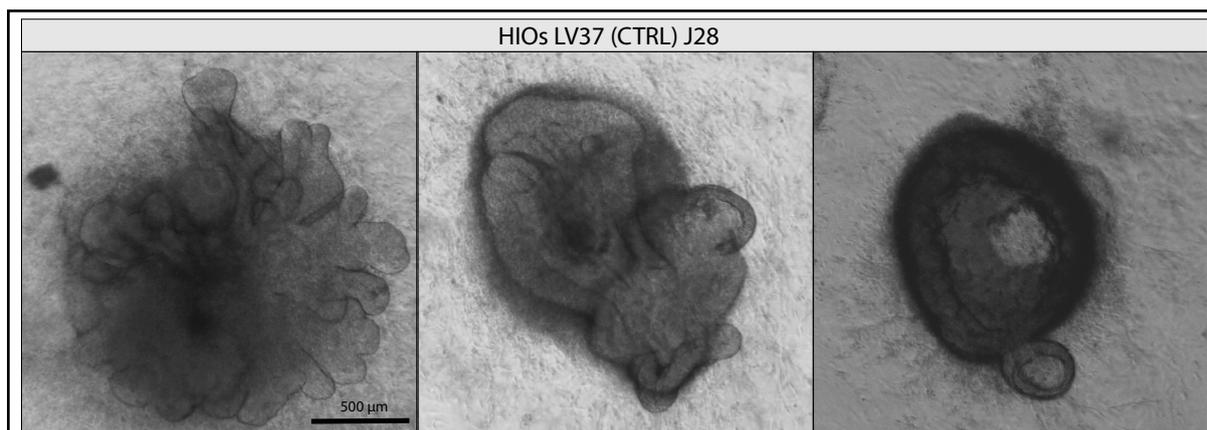
### Figure 81. Variabilité de la proportion d'épithélium et de mésenchyme entre les différents échantillons

(A-B) Analyse des marqueurs d'épithélium et de mésenchyme dans les échantillons d'après les expériences de RNA-seq: représentation graphique illustrant la variabilité de ces marqueurs dans les différents échantillons et tableau représentant les variations d'expression de ces marqueurs entre les mutants et CTRL: (A) HIOs RFX6<sup>-/-</sup> (B) HIOs NGN3<sup>-/-</sup>.

Rouge : gènes significativement sous-exprimés dans les mutants, bleu : gènes significativement surexprimés dans les mutants, orange : gènes à tendance différenciellement exprimés.

(Marqueurs issus de: McCracken et al., 2014; Bondow et al., 2012; Mani et al., 2009; Nokes et al., 2008; Funakoshi et al., 2008; Darmoul et al., 1994)

Malgré l'application stricte de toutes les étapes du protocole visant à limiter autant que possible la variabilité, celle-ci semble inévitable et il serait alors intéressant de développer des outils bio-informatiques de normalisation permettant d'en tenir compte lors de l'analyse transcriptomique de ce type d'échantillons. Un autre moyen pour limiter les conséquences de cette variabilité est de travailler avec des effectifs importants.



**Figure 82. Hétérogénéité morphologique des HIOs**

Images de microscopie optique à lumière transmise représentant la variabilité de la forme, taille et aspect des HIOs CTRL au stade J28. Cette variabilité est également observée chez les HIOs mutants. L'échelle est représentée par la barre noire correspondant à 500  $\mu\text{m}$ .

### 2.3. Effectifs limités et rareté des cellules entéroendocrines

Dans le cadre de notre étude par RNA-seq, les effectifs étudiés étaient très limités, entre 2 et 5 échantillons par condition (mutants vs CTRL) et il serait donc nécessaire de les augmenter, afin de limiter le poids de l'hétérogénéité des HIOs dans l'analyse statistique. Il est probable qu'avec des effectifs plus importants, un grand nombre de gènes qui n'apparaissent pas différenciellement exprimés dans notre analyse, le soit en réalité. Une autre limite de nos analyses est représentée par le fait que nos cellules d'intérêt soient une population rare au sein de l'épithélium intestinal. En effet, les cellules entéroendocrines ne représentent que 1% des cellules épithéliales, ce qui ne constitue qu'un très faible nombre de cellules au sein des HIOs. Il est donc difficile de déceler un effet des mutations sur cette population précise de cellule et il ne peut être exclu que l'absence des effets attendus sur la différenciation entéroendocrine soit due à la faible quantité de cellules entéroendocrines dans les échantillons.

Dans le but de pallier au très faible nombre de cellules entéroendocrines présentes dans les HIOs et ainsi faciliter leur étude, diverses approches ont récemment été développées. En effet, N. Petersen a montré une augmentation du nombre de cellules L sécrétrices de GLP1 dans des entéroïdes murins et organoïdes humains *ex vivo*, en réponse à l'application soit d'acides gras à chaînes courtes (SCFA), soit par le blocage de la signalisation Notch par la dibenzazépine (inhibiteur de la gamma sécrétase) ou encore par utilisation d'inhibiteur de ROCK (Y27632) (Petersen et al., 2018 ; 2015 ; 2014). Basak et al. (2017) ont démontré que l'induction de la quiescence des cellules Lgr5 par blocage de la voie EGF et MAPK induisait la différenciation en cellules entéroendocrines, permettant ainsi de générer des entéroïdes enrichis en tous les sous-types entéroendocrines. Enfin, Sinagoga et al. (2018) ont mis en évidence qu'une surexpression induite de *NGN3* pendant 8h dans les HIOs, permet

la formation d'organoïdes contenant jusqu'à 25% de cellules entéroendocrines, sécrétant l'ensemble des hormones et répondant à l'injection de glucose. Ces nouvelles découvertes pourraient être appliquées lors de prochaines expériences de différenciation de nos clones mutants, afin de mettre en évidence les variations géniques propres à la population de cellules entéroendocrines. Afin d'étudier spécifiquement cette population en s'affranchissant des autres types cellulaires épithéliaux, il est possible d'avoir recours à des méthodes de tri cellulaire sur les HIOs. Les cellules entéroendocrines peuvent être triées grâce à l'expression de marqueurs de surface spécifiques, tels que CLDN4 (Nagatake et al., 2014). Une autre méthode consisterait à modifier génétiquement les hiPSCs, afin de marquer les cellules entéroendocrines avec un fluorochrome, dont l'expression permettra de réaliser un tri cellulaire au FACS.

## 2.4. Maturité des HIOs

L'analyse a uniquement été réalisée sur des HIOs du stade J28, or il a été mis en évidence que l'expression de certains gènes est variable selon le stade de croissance étudié (**Figure 39B**). Ainsi, afin d'avoir une vision plus complète de l'impact des mutations générées sur le programme transcriptomique au cours du développement, il serait nécessaire d'élargir l'étude à d'autres stades de croissance des HIOs. D'autant plus qu'il a été montré que les HIOs gagnent progressivement en maturité, exprimant ainsi davantage de marqueurs de cellules différenciées de l'épithélium intestinal avec une compartimentalisation crypte-villosité plus marquée (Spence et al., 2011). Le stade précoce choisi pourrait donc expliquer l'absence d'expression des gènes codant certaines hormones. Pour pallier à cette limite du modèle d'étude, il est possible de transplanter les HIOs générés sous la capsule rénale de souris immunodéficientes, afin de permettre une maturation et expansion des tissus formés *in vitro*. En effet, les HIOs ainsi maturés, sont plus volumineux et expriment l'ensemble des hormones entéroendocrines (Watson et al., 2014).

## 2.5. Considérations méthodologiques concernant la culture cellulaire des hiPSCs

Les hiPSCs utilisées pour générer les clones mutants sont les LV37, qui sont dérivées d'une lignée de fibroblastes CPRE2 reprogrammée à l'aide de lentivirus portant un gène monocistronique codant le cocktail de Thomson OCT4, NANOG, SOX2 et LIN28 (ONSL). Cette lignée CPRE2 a été reprogrammée de différentes manières pour obtenir une reprogrammation la plus efficace possible et l'utilisation d'une cassette polycistronique portée par des rétrovirus s'est avérée être la combinaison la plus performante (Jung et al., 2014). Ainsi, la lignée LV37 qui a été utilisée dans nos expériences ne semble pas être la plus appropriée à la différenciation *in vitro* et il serait

alors intéressant d'établir des clones dans une lignée qui a été reprogrammée de la manière la plus efficace. Si la différenciation *in vitro* peut être affectée par la nature des cellules utilisées, il se pourrait également que le nombre de passage de celles-ci puisse affecter l'efficacité de la différenciation. En effet, lors d'une communication personnelle avec le Dr. Mahé, il a été suggéré que le nombre de passage des hiPSCs ne devait excéder 60, sous peine d'affecter la différenciation dirigée des cellules. Dans notre cas, les clones NGN3<sup>-/-</sup> sont à environ 60 passages et les clones RFX6<sup>-/-</sup> à 90 passages, alors que les hiPSCs CTRL sont à 30 passages. En plus des mutations volontairement induites, il se pourrait donc que le nombre de passage élevé des mutants puisse affecter la différenciation et créer des différences dans l'expression des gènes qui ne seraient pas dues à l'inactivation des facteurs de transcription. Effectivement, on peut penser que les manipulations répétées des cellules aient un impact sur le génome des hiPSCs et notamment sur leur caryotype. Il est vrai que le caryotype des LV37 a été analysé et jugé normal, mais ceci sur des cellules d'un nombre de passage faible (P25) et n'ayant subi aucune manipulation de transfection. Il faudrait donc aussi analyser le caryotype des clones mutants afin de déterminer si les manipulations répétées de culture cellulaire et la transfection des cellules n'ont eu aucun impact. En conclusion, pour éviter tout effet qui pourrait parasiter les analyses transcriptomiques, il est nécessaire de réitérer les expériences avec des clones mutants générés à partir de cellules plus « jeunes » et d'inclure des contrôles isotypiques, c'est-à-dire du même nombre de passage et qui aurait subi les mêmes manipulations de transfection que les mutants. C'est dans cette optique que de nouveaux clones NGN3<sup>-/-</sup> ont été obtenus en transfectant des hiPSCs à P20.

### 3. Les HIOs, un modèle pertinent ?

Les limites techniques décrites ci-avant sont propres à nos expériences et analyses mais on peut également se poser la question de la pertinence et de la prédictivité du modèle de différenciation *in vitro* à partir de cellules souches pluripotentes. Il permet certes de récapituler les grandes étapes du développement intestinal embryonnaire humain mais avec un manque de contexte physiologique certain. Un organe fonctionnel dans un organisme n'est pas simplement un groupe de cellules isolé mais est en interaction constante avec son environnement. Dans l'intestin, ces interactions se font notamment avec le système nerveux entérique (ENS), le microbiote, les glandes de Brunner, le système vasculaire et le système immunitaire (Dickson et al., 2018 ; Carabotti et al., 2015). Ce sont autant d'éléments qui manquent dans les HIOs générés et qui influencent, sans nuls doutes, la fidélité des résultats. Il a par exemple été montré que les cellules entéroendocrines sont étroitement liées aux neurones du système nerveux entérique par des prolongements cytoplasmiques appelés neuropodes, permettant la transmission synaptique entre les cellules entéroendocrines et les neurones (Bohorquez et al., 2015).

De plus, les preuves de l'importance primordiale du microbiote intestinal dans le fonctionnement normal et pathologique du corps humain, sont de nos jours irréfutables. Effectivement, les dysbioses provoquées par des altérations dans la diversité ou la structure du microbiote intestinal peuvent affecter les activités métaboliques, menant à des pathologies comme l'obésité et le diabète (Faith et al., 2013). Il est également admis que des modifications dans le régime alimentaire peuvent drastiquement affecter la différenciation et la fonction des cellules entéroendocrines, puisque celles-ci sont hautement sensibles aux nutriments (Richards et al., 2016). Ces changements à long terme n'influencent pas seulement la transcription des hormones au sein des cellules entéroendocrines mais modifient également le nombre des sous-types cellulaires intestinaux (Ritze et al., 2015). Ces découvertes, parmi d'autres, permettent de rendre compte de l'importance des tissus et facteurs environnants et mettent en lumière l'absence de ceux-ci dans des modèles *in vitro* tels que les HIOs.

Pour pallier à ce manque et améliorer ce modèle d'étude, l'équipe de J. Wells qui a établi le protocole de différenciation des hiPSCs en HIOs *in vitro*, a généré des HIOs innervés (HIO+ENS) en combinant des cellules de la crête neurale vagale dérivées des hiPSCs avec des sphéroïdes intestinales (Workman et al., 2016). La maturation *in vivo* des HIO+ENS leur permet de développer des caractéristiques similaires au système nerveux entérique humain, comme les faisceaux neuronaux alignés sur les couches musculaires stratifiées et une activité de type péristaltique suite à une stimulation (Workman et al., 2016).

Récemment, de nouveaux systèmes *in vitro* d'intestin sur puce utilisant la culture cellulaire en microfluidique et qui modélisent la structure, la fonction, la physiologie et potentiellement les pathologies intestinales humaines ont été développés. Ces systèmes permettent de simuler l'interaction entre le tissu intestinal et le microbiote, ainsi que le système immunitaire. Par contre, il ne permet pas l'étude du développement embryonnaire puisque les cellules intestinales utilisées sont des cellules épithéliales différenciées (Bein et al., 2018).

En plus de l'absence des tissus avoisinants, les HIOs sont considérés comme foëtaux et contrairement aux entéroïdes, ils ne reproduisent pas la spécificité des segments intestinaux mais forment une population mixte de cellule de nature proximale et distale (Spence et al., 2011). Or, des études ont montré une répartition dépendante de l'axe proximo-distal des différents sous-types entéroendocrines (Adriaenssens et al., 2018), ce qui a été confirmé par des travaux réalisés par des membres de l'équipe (A. Beucher et F. Blot). Par ailleurs, la nature des types cellulaires présents au sein des HIOs et la reproductibilité de leur production n'est pas avérée. C'est pourquoi des expériences de RNA-seq et RT-qPCR par l'approche en cellules uniques (*single cell*) se développent sur les HIOs, fournissant ainsi une plus grande sensibilité d'analyse de l'hétérogénéité cellulaire (Camp et al., 2017). Ces nouvelles approches, notamment entreprises au sein du laboratoire, permettront dans un futur proche une meilleure compréhension de la composition et des relations entre les types cellulaires dans les HIOs, afin d'étudier les mécanismes sous-tendant l'organogenèse chez l'homme.

# **BIBLIOGRAPHIE**

# Bibliographie

---

## A

---

Abou-Saleh, H., Zouein, F.A., El-Yazbi, A., Sanoudou, D., Raynaud, C., Rao, C., Pintus, G., Dehaini, H., and Eid, A.H. (2018). The march of pluripotent stem cells in cardiovascular regenerative medicine. *Stem Cell Research & Therapy* 9.

Abud, H.E., Watson, N., and Heath, J.K. (2005). Growth of intestinal epithelium in organ culture is dependent on EGF signalling. *Experimental Cell Research* 303, 252–262.

Adriaenssens, A.E., Reimann, F., and Gribble, F.M. (2018). Distribution and Stimulus Secretion Coupling of Enteroendocrine Cells along the Intestinal Tract. In *Comprehensive Physiology*, D.M. Pollock, ed. (Hoboken, NJ, USA: John Wiley & Sons, Inc.), pp. 1603–1638.

Aken, B.L., Achuthan, P., Akanni, W., Amode, M.R., Bernsdorff, F., Bhai, J., Billis, K., Carvalho-Silva, D., Cummins, C., Clapham, P., et al. (2017). Ensembl 2017. *Nucleic Acids Research* 45, D635–D642.

Ali-Rachedi, A., Varndell, I.M., Adrian, T.E., Gapp, D.A., Van Noorden, S., Bloom, S.R., and Polak, J.M. (1984). Peptide YY (PYY) immunoreactivity is co-stored with glucagon-related immunoreactants in endocrine cells of the gut and pancreas. *Histochemistry* 80, 487–491.

Amsen, D., Antov, A., and Flavell, R.A. (2009). The different faces of Notch in T-helper-cell differentiation. *Nature Reviews Immunology* 9, 116–124.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169.

Anderson, G.J., and Darshan, D. (2008). Small-molecule dissection of BMP signaling. *Nature Chemical Biology* 4, 15–16.

Aranias, T., Grosfeld, A., Poitou, C., Omar, A.A., Le Gall, M., Miquel, S., Garbin, K., Ribeiro, A., Bouillot, J.-L., Bado, A., et al. (2015). Lipid-rich diet enhances L-cell density in obese subjects and in mice through improved L-cell differentiation. *Journal of Nutritional Science* 4.

Aumailley, M., and Verrando, P. (1993). Structure et pathologie des membranes basales. *médecine/sciences* 9, 926.

---

## B

---

Baggio, L.L., and Drucker, D.J. (2007). Biology of Incretins: GLP-1 and GIP. *Gastroenterology* 132, 2131–2157.

Bankaitis, E.D., Ha, A., Kuo, C.J., and Magness, S.T. (2018). Reserve Stem Cells in Intestinal Homeostasis and Injury. *Gastroenterology* 155, 1348–1361.

Barker, N. (2014). Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. *Nature Reviews Molecular Cell Biology* 15, 19–33.

Barker, N., van Es, J.H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P.J., et al. (2007). Identification of stem cells in small intestine and colon by marker gene *Lgr5*. *Nature* 449, 1003–1007.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science* 315, 1709–1712.

- Bartolomucci, A., Possenti, R., Mahata, S.K., Fischer-Colbrie, R., Loh, Y.P., and Salton, S.R.J. (2011). The Extended Granin Family: Structure, Function, and Biomedical Implications. *Endocrine Reviews* 32, 755–797.
- Basak, O., Beumer, J., Wiebrands, K., Seno, H., van Oudenaarden, A., and Clevers, H. (2017). Induced Quiescence of Lgr5+ Stem Cells in Intestinal Organoids Enables Differentiation of Hormone-Producing Enteroendocrine Cells. *Cell Stem Cell* 20, 177-190.e4.
- Bastide, P., Darido, C., Pannequin, J., Kist, R., Robine, S., Marty-Double, C., Bibeau, F., Scherer, G., Joubert, D., Hollande, F., et al. (2007). Sox9 regulates cell proliferation and is required for Paneth cell differentiation in the intestinal epithelium. *The Journal of Cell Biology* 178, 635–648.
- Battle, E., Henderson, J.T., Beghtel, H., Sancho, E., Huls, G., Meeldijk, J., Robertson, J., van de Wetering, M., Pawson, T., and Clevers, H. in the Intestinal Epithelium by Controlling the Expression of EphB/EphrinB. 13.
- Bayha, E., Jørgensen, M.C., Serup, P., and Grapin-Botton, A. (2009). Retinoic Acid Signaling Organizes Endodermal Organ Specification along the Entire Antero-Posterior Axis. *PLoS ONE* 4, e5845.
- Beck, F. (2002). Homeobox genes in gut development. *Gut* 51, 450–454.
- Beck, F., Tata, F., and Chawengsaksophak, K. (2000). Homeobox genes and gut development. *BioEssays* 22, 431–441.
- Bein, A., Shin, W., Jalili-Firoozinehad, S., Park, M.H., Sontheimer-Phelps, A., Tovaglieri, A., Chalkiadaki, A., Kim, H.J., and Ingber, D.E. (2018). Microfluidic Organ-on-a-Chip Models of Human Intestine. *Cellular and Molecular Gastroenterology and Hepatology* 5, 659–668.
- Bergström, J.H., Berg, K.A., Rodríguez-Piñero, A.M., Stecher, B., Johansson, M.E.V., and Hansson, G.C. (2014). AGR2, an Endoplasmic Reticulum Protein, Is Secreted into the Gastrointestinal Mucus. *PLoS ONE* 9, e104186.
- Beucher, A., Gjernes, E., Collin, C., Courtney, M., Meunier, A., Collombat, P., and Gradwohl, G. (2012). The Homeodomain-Containing Transcription Factors Arx and Pax4 Control Enteroendocrine Subtype Specification in Mice. *PLoS ONE* 7, e36449.
- Bienvenu, T. (2002). ARX, a novel Prd-class-homeobox gene highly expressed in the telencephalon, is mutated in X-linked mental retardation. *Human Molecular Genetics* 11, 981–991.
- Birchenough, G.M.H., Johansson, M.E., Gustafsson, J.K., Bergström, J.H., and Hansson, G.C. (2015). New developments in goblet cell mucus secretion and function. *Mucosal Immunology* 8, 712–719.
- Blyszczuk, P., Czyz, J., Kania, G., Wagner, M., Roll, U., St-Onge, L., and Wobus, A.M. (2003). Expression of Pax4 in embryonic stem cells promotes differentiation of nestin-positive progenitor and insulin-producing cells. *Proceedings of the National Academy of Sciences* 100, 998–1003.
- Bohórquez, D.V., Shahid, R.A., Erdmann, A., Kreger, A.M., Wang, Y., Calakos, N., Wang, F., and Liddle, R.A. (2015). Neuroepithelial circuit formed by innervation of sensory enteroendocrine cells. *Journal of Clinical Investigation* 125, 782–786.
- Bondow, B.J., Faber, M.L., Wojta, K.J., Walker, E., and Battle, M.A. (2012). E-cadherin is required for intestinal morphogenesis in the mouse. *Dev Biol* 371, 1–12.
- Boroviak, K., Fu, B., Yang, F., Doe, B., and Bradley, A. (2017). Revealing hidden complexities of genomic rearrangements generated with Cas9. *Scientific Reports* 7.
- Brokhman, I., Xu, J., Coles, B.L.K., Razavi, R., Engert, S., Lickert, H., Babona-Pilipos, R., Morshead, C.M., Sibley, E., Chen, C., et al. (2018). Dual embryonic origin of the mammalian enteric nervous system. *Developmental Biology*.

Brouns, S.J.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J.H., Snijders, A.P.L., Dickman, M.J., Makarova, K.S., Koonin, E.V., and van der Oost, J. (2008). Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Science* 321, 960–964.

Buczacki, S.J.A., Zecchini, H.I., Nicholson, A.M., Russell, R., Vermeulen, L., Kemp, R., and Winton, D.J. (2013). Intestinal label-retaining cells are secretory precursors expressing Lgr5. *Nature* 495, 65–69.

Buehler, D.P., Prasad, N., Byers, S.L., Rosebrock, J., Levy, S.E., Clark, T.A., and Southard-Smith, M. (2014). Su2025 Differential Expression of Phox2b Marks Distinct Progenitor Cell Populations That Differ in Developmental Potential and Gene Expression in the Fetal Mouse Enteric Nervous System. *Gastroenterology* 146, S-526.

Burrin, D.G., Stoll, B., Guan, X., Cui, L., Chang, X., and Hadsell, D. (2007). GLP-2 rapidly activates divergent intracellular signaling pathways involved in intestinal cell survival and proliferation in neonatal piglets. *American Journal of Physiology-Endocrinology and Metabolism* 292, E281–E291.

---

## C

---

Camp, J.G., Wollny, D., and Treutlein, B. (2018). Single-cell genomics to guide human stem cell and tissue engineering. *Nature Methods* 15, 661–667.

Canver, M.C., Bauer, D.E., Dass, A., Yien, Y.Y., Chung, J., Masuda, T., Maeda, T., Paw, B.H., and Orkin, S.H. (2014). Characterization of Genomic Deletion Efficiency Mediated by Clustered Regularly Interspaced Palindromic Repeats (CRISPR)/Cas9 Nuclease System in Mammalian Cells. *Journal of Biological Chemistry* 289, 21312–21324.

Carabotti, M., Scirocco, A., Maselli, M.A., and Severi, C. The gut-brain axis: interactions between enteric microbiota, central and enteric nervous systems. *Annals of Gastroenterology* 7.

Carmon, K.S., Gong, X., Lin, Q., Thomas, A., and Liu, Q. R-spondins function as ligands of the orphan receptors LGR4 and LGR5 to regulate Wnt/ $\beta$ -catenin signaling. *CELL BIOLOGY* 6.

Cervantes, S., Yamaguchi, T.P., and Hebrok, M. (2009). Wnt5a is essential for intestinal elongation in mice. *Developmental Biology* 326, 285–294.

Chandra, V., Albagli-Curiel, O., Hastoy, B., Piccand, J., Randriamampita, C., Vaillant, E., Cavé, H., Busiah, K., Froguel, P., Vaxillaire, M., et al. (2014). RFX6 Regulates Insulin Secretion by Modulating Ca<sup>2+</sup> Homeostasis in Human  $\beta$  Cells. *Cell Reports* 9, 2206–2218.

Chang, T.M., and Chey, W.Y. (2001). Neurohormonal control of exocrine pancreas. *Curr. Opin. Gastroenterol.* 17, 416–425.

Chen, Y.-H., and Pruett-Miller, S.M. (2018). Improving single-cell cloning workflow for gene editing in human pluripotent stem cells. *Stem Cell Research* 31, 186–192.

Chen, C., Fang, R., Davis, C., Maravelias, C., and Sibley, E. (2009). Pdx1 inactivation restricted to the intestinal epithelium in mice alters duodenal gene expression in enterocytes and enteroendocrine cells. *American Journal of Physiology-Gastrointestinal and Liver Physiology* 297, G1126–G1137.

Cheng, H. (1974a). Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine II. Mucous cells. *American Journal of Anatomy* 141, 481–501.

Cheng, H. (1974b). Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine IV. Paneth cells. *American Journal of Anatomy* 141, 521–535.

Cheng, H., and Leblond, C.P. (1974). Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine III. Entero-endocrine cells. *American Journal of Anatomy* 141, 503–519.

Cheng, H., Merzel, J., and Leblond, C.P. (1969). Renewal of Paneth cells in the small intestine of the mouse. *American Journal of Anatomy* 126, 507–525.

Cheng, H., Leblond, C.P. (with the collaboration of G. Triglydes, A. Grignon and W.W. Chang (1974). Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine I. Columnar cell. *American Journal of Anatomy* 141, 461–479.

Concepcion, J.P., Reh, C.S., Daniels, M., Liu, X., Paz, V.P., Ye, H., Highland, H.M., Hanis, C.L., and Greeley, S.A.W. (2014). Neonatal diabetes, gallbladder agenesis, duodenal atresia, and intestinal malrotation caused by a novel homozygous mutation in RFX6: Neonatal diabetes syndrome from an RFX6 mutation. *Pediatric Diabetes* 15, 67–72.

Cortina, G., Smart, C.N., Farmer, D.G., Bhuta, S., Treem, W.R., Hill, I.D., and Martín, M.G. (2007). Enteroendocrine cell dysgenesis and malabsorption, a histopathologic and immunohistochemical characterization. *Human Pathology* 38, 570–580.

Crosnier, C., Stamatakis, D., and Lewis, J. (2006). Organizing cell renewal in the intestine: stem cells, signals and combinatorial control. *Nature Reviews Genetics* 7, 349–359.

Cui, S., and Chang, P.-Y. (2016). Current understanding concerning intestinal stem cells. *World Journal of Gastroenterology* 22, 7099.

Cummings, D.E., Purnell, J.Q., Frayo, R.S., Schmidova, K., Wisse, B.E., and Weigle, D.S. (2001). A Preprandial Rise in Plasma Ghrelin Levels Suggests a Role in Meal Initiation in Humans. *Diabetes* 50, 1714–1719.

---

## D

---

D'Amour, K.A., Agulnick, A.D., Eliazer, S., Kelly, O.G., Kroon, E., and Baetge, E.E. (2005). Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nature Biotechnology* 23, 1534–1541.

Darmoul, D., Voisin, T., Couvineau, A., Rouyerfessard, C., Salomon, R., Wang, Y.X., Swallow, D.M., and Laburthe, M. (1994). Regional Expression of Epithelial Dipeptidyl Peptidase IV in the Human Intestines. *Biochemical and Biophysical Research Communications* 203, 1224–1229.

Dash, S., Xiao, C., Morgantini, C., Connelly, P.W., Patterson, B.W., and Lewis, G.F. (2014). Glucagon-Like Peptide-2 Regulates Release of Chylomicrons From the Intestine. *Gastroenterology* 147, 1275-1284.e4.

Desai, S., Loomis, Z., Pugh-Bernard, A., Schrunk, J., Doyle, M.J., Minic, A., McCoy, E., and Sussel, L. (2008). Nkx2.2 regulates cell fate choice in the enteroendocrine cell lineages of the intestine. *Developmental Biology* 313, 58–66.

Dessimoz, J., Opoka, R., Kordich, J.J., Grapin-Botton, A., and Wells, J.M. (2006). FGF signaling is necessary for establishing gut tube domains along the anterior–posterior axis in vivo. *Mechanisms of Development* 123, 42–55.

Dianov, G.L., and Hübscher, U. (2013). Mammalian Base Excision Repair: the Forgotten Archangel. *Nucleic Acids Research* 41, 3483–3490.

Dickson, I. (2018). Microbiota modulate ENS maturation. *Nature Reviews Gastroenterology & Hepatology* 15, 454–455.

Dlugosz, A., Muschiol, S., Zakikhany, K., Assadi, G., D'Amato, M., and Lindberg, G. (2014). Human enteroendocrine cell responses to infection with *Chlamydia trachomatis*: a microarray study. *Gut Pathogens* 6, 24.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21.

Dogan, A., Wang, Z.D., and Spencer, J. (1995). E-cadherin expression in intestinal epithelium. *Journal of Clinical Pathology* 48, 143–146.

Driehuis, E., and Clevers, H. (2017). CRISPR/Cas 9 genome editing and its applications in organoids. *American Journal of Physiology-Gastrointestinal and Liver Physiology* 312, G257–G265.

Drucker, D.J. (2005). Biologic actions and therapeutic potential of the proglucagon-derived peptides. *Nat Clin Pract Endocrinol Metab* 1, 22–31.

Du, A., McCracken, K.W., Walp, E.R., Terry, N.A., Klein, T.J., Han, A., Wells, J.M., and May, C.L. (2012). Arx is required for normal enteroendocrine cell development in mice and humans. *Developmental Biology* 365, 175–188.

---

## E

---

Egerod, K.L., Engelstoft, M.S., Grunddal, K.V., Nøhr, M.K., Secher, A., Sakata, I., Pedersen, J., Windeløv, J.A., Füchtbauer, E.-M., Olsen, J., et al. (2012). A Major Lineage of Enteroendocrine Cells Coexpress CCK, Secretin, GIP, GLP-1, PYY, and Neurotensin but Not Somatostatin. *Endocrinology* 153, 5782–5795.

Engelstoft, M.S., Lund, M.L., Grunddal, K.V., Egerod, K.L., Osborne-Lawrence, S., Poulsen, S.S., Zigman, J.M., and Schwartz, T.W. (2015). Research Resource: A Chromogranin A Reporter for Serotonin and Histamine Secreting Enteroendocrine Cells. *Molecular Endocrinology* 29, 1658–1671.

Erlandsen, S.L. Paneth Cell Function: Phagocytosis and Intracellular Digestion of Intestinal Microorganisms. 23.

van Es, J.H., van Gijn, M.E., Riccio, O., van den Born, M., Vooijs, M., Begthel, H., Cozijnsen, M., Robine, S., Winton, D.J., Radtke, F., et al. (2005). Notch/  $\gamma$ -secretase inhibition turns proliferative cells in intestinal crypts and adenomas into goblet cells. *Nature* 435, 959–963.

van Es, J.H., Sato, T., van de Wetering, M., Lyubimova, A., Yee Nee, A.N., Gregorieff, A., Sasaki, N., Zeinstra, L., van den Born, M., Korving, J., et al. (2012). Dll1+ secretory progenitor cells revert to stem cells upon crypt damage. *Nature Cell Biology* 14, 1099–1104.

---

## F

---

Faith, J.J., Guruge, J.L., Charbonneau, M., Subramanian, S., Seedorf, H., Goodman, A.L., Clemente, J.C., Knight, R., Heath, A.C., Leibel, R.L., et al. (2013). The Long-Term Stability of the Human Gut Microbiota. *Science* 341, 1237439–1237439.

Fichtner, F., Urrea Castellanos, R., and Ülker, B. (2014). Precision genetic modifications: a new era in molecular biology and crop improvement. *Planta* 239, 921–939.

Finkbeiner, S.R., Freeman, J.J., Wieck, M.M., El-Nachef, W., Altheim, C.H., Tsai, Y.-H., Huang, S., Dyal, R., White, E.S., Grikscheit, T.C., et al. (2015). Generation of tissue-engineered small intestine using embryonic stem cell-derived human intestinal organoids. *Biology Open* 4, 1462–1472.

Flasse, L.C., Pirson, J.L., Stern, D.G., Von Berg, V., Manfroid, I., Peers, B., and Voz, M.L. (2013). *Ascl1b* and *Neurod1*, instead of *Neurog3*, control pancreatic endocrine cell fate in zebrafish. *BMC Biology* 11, 78.

Fothergill, L.J., and Furness, J.B. (2018). Diversity of enteroendocrine cells investigated at cellular and subcellular levels: the need for a new classification scheme. *Histochem. Cell Biol.* 150, 693–702.

Foulke-Abel, J., In, J., Kovbasnjuk, O., Zachos, N.C., Ettayebi, K., Blutt, S.E., Hyser, J.M., Zeng, X.-L., Crawford, S.E., Broughman, J.R., et al. (2014). Human enteroids as an ex-vivo model of host-pathogen interactions in the gastrointestinal tract. *Experimental Biology and Medicine* 239, 1124–1134.

Fre, S., Huyghe, M., Mourikis, P., Robine, S., Louvard, D., and Artavanis-Tsakonas, S. (2005). Notch signals control the fate of immature progenitor cells in the intestine. *Nature* 435, 964–968.

Fu, L., and Shi, Y.-B. (2017). The Sox transcriptional factors: Functions during intestinal development in vertebrates. *Seminars in Cell & Developmental Biology* 63, 58–67.

Fujii, M., Matano, M., Nanki, K., and Sato, T. (2015). Efficient genetic engineering of human intestinal organoids using electroporation. *Nat Protoc* 10, 1474–1485.

Funakoshi, S., Ezaki, T., Kong, J., Guo, R.J., and Lynch, J.P. (2008). Repression of the desmocollin 2 gene expression in human colon cancer cells is relieved by the homeodomain transcription factors Cdx1 and Cdx2. *Mol. Cancer Res.* 6, 1478–1490.

Furness, J.B., Rivera, L.R., Cho, H.-J., Bravo, D.M., and Callaghan, B. (2013). The gut as a sensory organ. *Nature Reviews Gastroenterology & Hepatology* 10, 729–740.

---

## G

---

Gabdoulline, R., Kaisers, W., Gaspar, A., Meganathan, K., Doss, M.X., Jagtap, S., Hescheler, J., Sachinidis, A., and Schwender, H. (2015). Differences in the Early Development of Human and Mouse Embryonic Stem Cells. *PLOS ONE* 10, e0140803.

Gage, B.K., Baker, R.K., and Kieffer, T.J. (2014). Overexpression of PAX4 reduces glucagon expression in differentiating hESCs. *Islets* 6, e29236.

Gerbe, F., Legraverend, C., and Jay, P. (2012). The intestinal epithelium tuft cells: specification and function. *Cellular and Molecular Life Sciences* 69, 2907–2917.

Gerbe, F., Sidot, E., Smyth, D.J., Ohmoto, M., Matsumoto, I., Dardalhon, V., Cesses, P., Garnier, L., Pouzolles, M., Brulin, B., et al. (2016). Intestinal epithelial tuft cells initiate type 2 mucosal immunity to helminth parasites. *Nature* 529, 226–230.

Germán-Díaz, M., Rodríguez-Gil, Y., Cruz-Rojo, J., Charbit-Henrion, F., Cerf-Bensussan, N., Manzanares-López Manzanares, J., and Moreno-Villares, J.M. (2017). A New Case of Congenital Malabsorptive Diarrhea and Diabetes Secondary to Mutant Neurogenin-3. *Pediatrics* 140, e20162210.

Germann, M., Xu, H., Malaterre, J., Sampurno, S., Huyghe, M., Cheasley, D., Fre, S., and Ramsay, R.G. (2014). Tripartite interactions between Wnt signaling, Notch and Myb for stem/progenitor cell functions during intestinal tumorigenesis. *Stem Cell Research* 13, 355–366.

Gierl, M.S. (2006). The Zinc-finger factor *Insm1* (IA-1) is essential for the development of pancreatic beta cells and intestinal endocrine cells. *Genes & Development* 20, 2465–2478.

Ginis, I., Luo, Y., Miura, T., Thies, S., Brandenberger, R., Gerecht-Nir, S., Amit, M., Hoke, A., Carpenter, M.K., Itskovitz-Eldor, J., et al. (2004). Differences between human and mouse embryonic stem cells. *Developmental Biology* 269, 360–380.

Gracz, A.D., Samsa, L.A., Fordham, M.J., Trotier, D.C., Zwarycz, B., Lo, Y.-H., Bao, K., Starmer, J., Raab, J.R., Shroyer, N.F., et al. (2018). Sox4 Promotes Atoh1-Independent Intestinal Secretory Differentiation Toward Tuft and Enteroendocrine Fates. *Gastroenterology* 155, 1508-1523.e10.

Gradwohl, G., Dierich, A., LeMeur, M., and Guillemot, F. (2000). neurogenin3 is required for the development of the four endocrine cell lineages of the pancreas. *Proceedings of the National Academy of Sciences* 97, 1607–1611.

Gregorieff, A., Pinto, D., Begthel, H., Destrée, O., Kielman, M., and Clevers, H. (2005). Expression Pattern of Wnt Signaling Components in the Adult Intestine. *Gastroenterology* 129, 626–638.

Gribble, F.M., and Reimann, F. (2016). Enteroendocrine Cells: Chemosensors in the Intestinal Epithelium. *Annual Review of Physiology* 78, 277–299.

Gribble, F.M., Meek, C.L., and Reimann, F. (2018). Targeted intestinal delivery of incretin secretagogues—towards new diabetes and obesity therapies. *Peptides* 100, 68–74.

Grissa, I., Vergnaud, G., and Pourcel, C. (2007). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* 8, 172.

Gross, S., Garofalo, D.C., Balderes, D.A., Mastracci, T.L., Dias, J.M., Perlmann, T., Ericson, J., and Sussel, L. (2016). The novel enterochromaffin marker *Lmx1a* regulates serotonin biosynthesis in enteroendocrine cell lineages downstream of *Nkx2.2*. *Development* 143, 2616–2628.

Grün, D., Lyubimova, A., Kester, L., Wiebrands, K., Basak, O., Sasaki, N., Clevers, H., and van Oudenaarden, A. (2015). Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* 525, 251–255.

Guerrini, R., Moro, F., Kato, M., Barkovich, A.J., Shiihara, T., McShane, M.A., Hurst, J., Loi, M., Tohyama, J., Norci, V., et al. (2007). Expansion of the first PolyA tract of *ARX* causes infantile spasms and status dystonicus. *Neurology* 69, 427–433.

Guilinger, J.P., Thompson, D.B., and Liu, D.R. (2014). Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nature Biotechnology* 32, 577–582.

Guz, Y., Montminy, M.R., Stein, R., Leonard, J., Gamer, L.W., Wright, C.V.E., and Teitelman, G. Expression of murine STF-1, a putative insulin gene transcription factor, in  $\beta$  cells of pancreas, duodenal epithelium and pancreatic exocrine and endocrine progenitors during ontogeny. 8.

---

## H

---

Haber, A.L., Biton, M., Rogel, N., Herbst, R.H., Shekhar, K., Smillie, C., Burgin, G., Delorey, T.M., Howitt, M.R., Katz, Y., et al. (2017). A single-cell survey of the small intestinal epithelium. *Nature* 551, 333–339.

Habib, A.M., Richards, P., Cairns, L.S., Rogers, G.J., Bannon, C.A.M., Parker, H.E., Morley, T.C.E., Yeo, G.S.H., Reimann, F., and Gribble, F.M. (2012). Overlap of Endocrine Hormone Expression in the Mouse Intestine Revealed by Transcriptional Profiling and Flow Cytometry. *Endocrinology* 153, 3054–3065.

Hancili, S., Bonnefond, A., Philippe, J., Vaillant, E., De Graeve, F., Sand, O., Busiah, K., Robert, J.-J., Polak, M., Froguel, P., et al. (2018). A novel *NEUROG3* mutation in neonatal diabetes associated with a neuro-intestinal syndrome: HANCILI et al. *Pediatric Diabetes* 19, 381–387.

Hanson, A.C., and Hagerman, R.J. (2014). Serotonin dysregulation in Fragile X Syndrome: implications for treatment. *Intractable & Rare Diseases Research* 3, 110–117.

Haramis, A.-P.G. (2004). De Novo Crypt Formation and Juvenile Polyposis on BMP Inhibition in Mouse Intestine. *Science* 303, 1684–1686.

He, X.C., Zhang, J., Tong, W.-G., Tawfik, O., Ross, J., Scoville, D.H., Tian, Q., Zeng, X., He, X., Wiedemann, L.M., et al. (2004). BMP signaling inhibits intestinal stem cell self-renewal through suppression of Wnt- $\beta$ -catenin signaling. *Nature Genetics* 36, 1117–1121.

Heiskala, K., and Andersson, L.C. (2013). Reg IV is differently expressed in enteroendocrine cells of human small intestine and colon. *Regulatory Peptides* 183, 27–34.

Hill, C.S. (2018). Spatial and temporal control of NODAL signaling. *Current Opinion in Cell Biology* 51, 50–57.

Hill, M.E., Asa, S.L., and Drucker, D.J. (1999b). Essential Requirement for Pax6 in Control of Enteroendocrine Proglucagon Gene Transcription. 13, 13.

Hooton, D., Lentle, R., Monro, J., Wickham, M., and Simpson, R. (2015). The Secretion and Action of Brush Border Enzymes in the Mammalian Small Intestine. In *Reviews of Physiology, Biochemistry and Pharmacology*, B. Nilius, T. Gudermann, R. Jahn, R. Lill, O.H. Petersen, and P.P. de Tombe, eds. (Cham: Springer International Publishing), pp. 59–118.

Hopfer, U. (1977). Kinetics of Na<sup>+</sup>-dependent D-glucose transport. *Journal of Supramolecular Structure* 7, 1–13.

Hou, B., Zhang, D., Zhao, S., Wei, M., Yang, Z., Wang, S., Wang, J., Zhang, X., Liu, B., Fan, L., et al. (2015). Scalable and Dil-compatible optical clearance of the mammalian brain. *Frontiers in Neuroanatomy* 9.

Huang, H.-P., Liu, M., El-Hodiri, H.M., Chu, K., Jamrich, M., and Tsai, M.-J. (2000). Regulation of the Pancreatic Islet-Specific Gene BETA2 (neuroD) by Neurogenin 3. *Molecular and Cellular Biology* 20, 3292–3307.

Hughes, C.S., Postovit, L.M., and Lajoie, G.A. (2010). Matrigel: A complex protein mixture required for optimal growth of cell culture. *PROTEOMICS* 10, 1886–1890.

Hussain, M.M. (2014). Intestinal lipid absorption and lipoprotein formation: Current Opinion in Lipidology 25, 200–206.

---

## I

---

Ikeguchi, E., Harada, N., Kanemaru, Y., Sankoda, A., Yamane, S., Iwasaki, K., Imajo, M., Murata, Y., Suzuki, K., Joo, E., et al. (2018). Transcriptional factor Pdx1 is involved in age-related GIP hypersecretion in mice. *American Journal of Physiology-Gastrointestinal and Liver Physiology* 315, G272–G282.

Imuta, Y., Nishioka, N., Kiyonari, H., and Sasaki, H. (2009). Short limbs, cleft palate, and delayed formation of flat proliferative chondrocytes in mice with targeted disruption of a putative protein kinase gene, Pkdcc ( AW548124 ). *Developmental Dynamics* 238, 210–222.

---

## J

---

Jacobsen, C.M., Narita, N., Bielinska, M., Syder, A.J., Gordon, J.I., and Wilson, D.B. (2002). Genetic Mosaic Analysis Reveals That GATA-4 Is Required for Proper Differentiation of Mouse Gastric Epithelium. *Developmental Biology* 241, 34–46.

Jadhav, U., Saxena, M., O'Neill, N.K., Saadatpour, A., Yuan, G.-C., Herbert, Z., Murata, K., and Shivdasani, R.A. (2017). Dynamic Reorganization of Chromatin Accessibility Signatures during Dedifferentiation of Secretory Precursors into Lgr5<sup>+</sup> Intestinal Stem Cells. *Cell Stem Cell* 21, 65-77.e5.

Jenkins, G.W., and Tortora, G.J. (2013). *Anatomy and physiology: from science to life* (Hoboken, NJ: Wiley).

Jenny, M. (2002). Neurogenin3 is differentially required for endocrine cell fate specification in the intestinal and gastric epithelium. *The EMBO Journal* 21, 6338–6347.

Jensen, J., Pedersen, E.E., Galante, P., Hald, J., Heller, R.S., Ishibashi, M., Kageyama, R., Guillemot, F., Serup, P., and Madsen, O.D. (2000). Control of endodermal endocrine development by Hes-1. *Nature Genetics* 24, 36–44.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* 337, 816–821.

Jung, L., Tropel, P., Moal, Y., Teletin, M., Jeandidier, E., Gayon, R., Himmelspach, C., Bello, F., Andre, C., Tosch, A., et al. (2014). ONSL and OSKM cocktails act synergistically in reprogramming human somatic cells into induced pluripotent stem cells. *Molecular Human Reproduction* 20, 538–549.

---

## K

---

Karlsson, L. PDGF-A in gut development. 10.

Kato, M., Das, S., Petras, K., Kitamura, K., Morohashi, K., Abuelo, D.N., Barr, M., Bonneau, D., Brady, A.F., Carpenter, N.J., et al. (2004). Mutations of ARX are associated with striking pleiotropy and consistent genotype-phenotype correlation. *Human Mutation* 23, 147–159.

Kawazoe, Y., Sekimoto, T., Araki, M., Takagi, K., Araki, K., and Yamamura, K. (2002). Region-specific gastrointestinal Hox code during murine embryonal gut development. *Development, Growth and Differentiation* 44, 77–84.

Ke, M.-T., Fujimoto, S., and Imai, T. (2013). SeeDB: a simple and morphology-preserving optical clearing agent for neuronal circuit reconstruction. *Nature Neuroscience* 16, 1154–1161.

Keszthelyi, d., troost, f. j., and masclee, a. a. m. (2009). Understanding the role of tryptophan and serotonin metabolism in gastrointestinal function. *Neurogastroenterology & Motility* 21, 1239–1249.

Khan, N., Dandan, W., Hassani, N.A., and Had, S. (2016). A Newly-Discovered Mutation in the RFX6 Gene of the Rare Mitchell-Riley Syndrome. *Journal of Clinical Research in Pediatric Endocrinology* 8, 246–249.

Kim, B.-M., Buchner, G., Miletich, I., Sharpe, P.T., and Shivdasani, R.A. (2005). The Stomach Mesenchymal Transcription Factor Barx1 Specifies Gastric Epithelial Identity through Inhibition of Transient Wnt Signaling. *Developmental Cell* 8, 611–622.

Kim, C.-K., Yang, V.W., and Bialkowska, A.B. (2017). The Role of Intestinal Stem Cells in Epithelial Regeneration Following Radiation-Induced Gut Injury. *Current Stem Cell Reports* 3, 320–332.

Kimelman, D., and Griffin, K.J. (2000). Vertebrate mesendoderm induction and patterning. *Current Opinion in Genetics & Development* 10, 350–356.

Kitamura, K., Yanazawa, M., Sugiyama, N., Miura, H., Iizuka-Kogo, A., Kusaka, M., Omichi, K., Suzuki, R., Kato-Fukui, Y., Kamiirisa, K., et al. (2002). Mutation of ARX causes abnormal development of forebrain and testes in mice and X-linked lissencephaly with abnormal genitalia in humans. *Nature Genetics* 32, 359–369.

Konno-Shimizu, M., Yamamichi, N., Inada, K., Kageyama-Yahara, N., Shiogama, K., Takahashi, Y., Asada-Hirayama, I., Yamamichi-Nishina, M., Nakayama, C., Ono, S., et al. (2013). Cathepsin E Is a Marker of Gastric Differentiation and Signet-Ring Cell Carcinoma of Stomach: A Novel Suggestion on Gastric Tumorigenesis. *PLoS ONE* 8, e56766.

Korinek, V., Barker, N., Moerer, P., van Donselaar, E., Huls, G., Peters, P.J., and Clevers, H. (1998). Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. *Nature Genetics* 19, 379–383.

Kosicki, M., Tomberg, K., and Bradley, A. (2018). Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. *Nature Biotechnology*.

Kubo, A. (2004). Development of definitive endoderm from embryonic stem cells in culture. *Development* 131, 1651–1662.

Kursan, S., McMillen, T.S., Beesetty, P., Dias-Junior, E., Almutairi, M.M., Sajib, A.A., Kozak, J.A., Aguilar-Bryan, L., and Di Fulvio, M. (2017). The neuronal K<sup>+</sup>Cl<sup>-</sup> co-transporter 2 (Slc12a5) modulates insulin secretion. *Scientific Reports* 7.

---

## L

---

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9, 357–359.

Larsson, L.-I., Madsen, O.D., Serup, P., Jonsson, J., and Edlund, H. (1996). Pancreatic-duodenal homeobox 1 -role in gastric endocrine patterning. *Mechanisms of Development* 60, 175–184.

Larsson, L.-I., St-Onge, L., Hougaard, D.M., Sosa-Pineda, B., and Gruss, P. (1998). Pax 4 and 6 regulate gastrointestinal endocrine cell development. *Mechanisms of Development* 79, 153–159.

Le Guen, L., Marchal, S., Faure, S., and de Santa Barbara, P. (2015). Mesenchymal–epithelial interactions during digestive tract development and epithelial stem cell regeneration. *Cellular and Molecular Life Sciences* 72, 3883–3896.

Lee, C.S. (2002). Neurogenin 3 is essential for the proper specification of gastric enteroendocrine cells and the maintenance of gastric epithelial cell identity. *Genes & Development* 16, 1488–1497.

Lee, C.Y. (2016). Glucagon-Like Peptide-1 Formulation - the Present and Future Development in Diabetes Treatment. *Basic & Clinical Pharmacology & Toxicology* 118, 173–180.

Lee, J., Hollenberg, S., Snider, L., Turner, D., Lipnick, N., and Weintraub, H. (1995). Conversion of *Xenopus* ectoderm into neurons by NeuroD, a basic helix-loop-helix protein. *Science* 268, 836–844.

Li, J., Song, J., Zaytseva, Y.Y., Liu, Y., Rychahou, P., Jiang, K., Starr, M.E., Kim, J.T., Harris, J.W., Yiannikouris, F.B., et al. (2016a). An obligatory role for neurotensin in high-fat-diet-induced obesity. *Nature* 533, 411–415.

Li, N., Nakauka-Ddamba, A., Tobias, J., Jensen, S.T., and Lengner, C.J. (2016b). Mouse Label-Retaining Cells Are Molecularly and Functionally Distinct From Reserve Intestinal Stem Cells. *Gastroenterology* 151, 298-310.e7.

Liang, T., Qin, T., Xie, L., Dolai, S., Zhu, D., Prentice, K.J., Wheeler, M., Kang, Y., Osborne, L., and Gaisano, H.Y. (2017). New Roles of Syntaxin-1A in Insulin Granule Exocytosis and Replenishment. *Journal of Biological Chemistry* 292, 2203–2216.

Liao, M., Shen, J., Zhang, Y., Li, S.-H., Li, X.-J., and Li, H. (2005). Immunohistochemical Localization of Huntingtin-associated Protein 1 in Endocrine System of the Rat. *Journal of Histochemistry & Cytochemistry* 53, 1517–1524.

Lickert, H., and Kemler, R. (2002). Functional analysis of cis-regulatory elements controlling initiation and maintenance of early *Cdx1* gene expression in the mouse. *Developmental Dynamics* 225, 216–220.

Lieverse, R.J., Jansen, J.B.M.J., Masclee, A.A.M., and Lamers, C.B.H.W. (1994). Satiety effects of cholecystokinin in humans. *Gastroenterology* 106, 1451–1454.

Liu, P., Wakamiya, M., Shea, M.J., Albrecht, U., Behringer, R.R., and Bradley, A. (1999). Requirement for Wnt3 in vertebrate axis formation. *Nature Genetics* 22, 361–365.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15.

Lund, M.L., Egerod, K.L., Engelstoft, M.S., Dmytriyeva, O., Theodorsson, E., Patel, B.A., and Schwartz, T.W. (2018). Enterochromaffin 5-HT cells – A major target for GLP-1 and gut microbial metabolites. *Molecular Metabolism* 11, 70–83.

Luo, T., Zhao, J., Lu, Z., Bi, J., Pang, T., Cui, H., Yang, B., Li, W., Wang, Y., Wu, S., et al. (2018). Characterization of long non-coding RNAs and MEF2C-AS1 identified as a novel biomarker in diffuse gastric cancer. *Translational Oncology* 11, 1080–1089.

Luo, W., Friedman, M.S., Shedden, K., Hankenson, K.D., and Woolf, P.J. (2009). GAGE: generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics* 10, 161.

---

## M

---

Mabbott, N.A., Donaldson, D.S., Ohno, H., Williams, I.R., and Mahajan, A. (2013). Microfold (M) cells: important immunosurveillance posts in the intestinal epithelium. *Mucosal Immunology* 6, 666–677.

Mace, O.J., and Marshall, F. (2013). DIGESTIVE PHYSIOLOGY OF THE PIG SYMPOSIUM: Gut chemosensing and the regulation of nutrient absorption and energy supply1. *Journal of Animal Science* 91, 1932–1945.

Mace, O.J., Tehan, B., and Marshall, F. (2015). Pharmacology and physiology of gastrointestinal enteroendocrine cells. *Pharmacology Research & Perspectives* 3, e00155.

Mace, O.J., Schindler, M., and Patel, S. The regulation of K- and L-cell activity by GLUT2 and CasR in rat small intestine. 45.

Madara, J.L. (1982). Cup cells: structure and distribution of a unique class of epithelial cells in guinea pig, rabbit, and monkey small intestine. *Gastroenterology* 83, 981–994.

Madara, J.L., Nash, S., Moore, R., and Atisook, K. (1990). Structure and function of the intestinal epithelial barrier in health and disease. *Monogr Pathol* 306–324.

Madison, B.B. (2005). Epithelial hedgehog signals pattern the intestinal crypt-villus axis. *Development* 132, 279–289.

Madison, B.B., McKenna, L.B., Dolson, D., Epstein, D.J., and Kaestner, K.H. (2009). FoxF1 and FoxL1 Link Hedgehog Signaling and the Control of Epithelial Proliferation in the Developing Stomach and Intestine. *Journal of Biological Chemistry* 284, 5936–5944.

Magenheim, J., Klein, A.M., Stanger, B.Z., Ashery-Padan, R., Sosa-Pineda, B., Gu, G., and Dor, Y. (2011). Ngn3+ endocrine progenitor cells control the fate and morphogenesis of pancreatic ductal epithelium. *Developmental Biology* 359, 26–36.

Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., Barrangou, R., Brouns, S.J.J., Charpentier, E., Haft, D.H., et al. (2015). An updated evolutionary classification of CRISPR–Cas systems. *Nature Reviews Microbiology* 13, 722–736.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-Guided Human Genome Engineering via Cas9. *Science* 339, 823–826.

Mani, B.K., and Zigman, J.M. (2017). Ghrelin as a Survival Hormone. *Trends in Endocrinology & Metabolism* 28, 843–854.

Mani, S.A., Guo, W., Liao, M.-J., Eaton, E.N., Ayyanan, A., Zhou, A.Y., Brooks, M., Reinhard, F., Zhang, C.C., Shipitsin, M., et al. (2008). The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 133, 704–715.

Marchand, A.R., Luck, D., and DiScala, G. (2003). Evaluation of an improved automated analysis of freezing behaviour in rats and its use in trace fear conditioning. *Journal of Neuroscience Methods* 126, 145–153.

Marinou, K., Christodoulides, C., Antoniadis, C., and Koutsilieris, M. (2012). Wnt signaling in cardiovascular physiology. *Trends in Endocrinology & Metabolism* 23, 628–636.

Marraffini, L.A., and Sontheimer, E.J. (2010). Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* 463, 568–571.

May, C.L., and Kaestner, K.H. (2010). Gut endocrine cell development. *Molecular and Cellular Endocrinology* 323, 70–75.

McCracken, K.W., Howell, J.C., Wells, J.M., and Spence, J.R. (2011). Generating human intestinal tissue from pluripotent stem cells in vitro. *Nature Protocols* 6, 1920–1928.

McCracken, K.W., Catá, E.M., Crawford, C.M., Sinagoga, K.L., Schumacher, M., Rockich, B.E., Tsai, Y.-H., Mayhew, C.N., Spence, J.R., Zavros, Y., et al. (2014). Modelling human development and disease in pluripotent stem-cell-derived gastric organoids. *Nature* 516, 400–404.

McCubrey, J.A., Steelman, L.S., Bertrand, F.E., Davis, N.M., Abrams, S.L., Montalto, G., D'Assoro, A.B., Libra, M., Nicoletti, F., Maestro, R., et al. (2014). Multifaceted roles of GSK-3 and Wnt/  $\beta$ -catenin in hematopoiesis and leukemogenesis: opportunities for therapeutic intervention. *Leukemia* 28, 15–33.

Mellitzer, G., and Gradwohl, G. (2011). Enteroendocrine cells and lipid absorption: Current Opinion in Lipidology 22, 171–175.

Mellitzer, G., Beucher, A., Lobstein, V., Michel, P., Robine, S., Keding, M., and Gradwohl, G. (2010). Loss of enteroendocrine cells in mice alters lipid absorption and glucose homeostasis and impairs postnatal survival. *Journal of Clinical Investigation* 120, 1708–1721.

Meran, L., Baulies, A., and Li, V.S.W. (2017). Intestinal Stem Cell Niche: The Extracellular Matrix and Cellular Components. *Stem Cells International* 2017, 1–11.

Mitchell, J., Punthakee, Z., Lo, B., Bernard, C., Chong, K., Newman, C., Cartier, L., Desilets, V., Cutz, E., Hansen, I.L., et al. (2004). Neonatal diabetes, with hypoplastic pancreas, intestinal atresia and gall bladder hypoplasia: search for the aetiology of a new autosomal recessive syndrome. *Diabetologia* 47, 2160–2167.

Mithieux, G., Andreelli, F., and Magnan, C. (2009). Intestinal gluconeogenesis: key signal of central control of energy and glucose homeostasis: Current Opinion in Clinical Nutrition and Metabolic Care 12, 419–423.

Miyawaki, K., Yamada, Y., Yano, H., Niwa, H., Ban, N., Ihara, Y., Kubota, A., Fujimoto, S., Kajikawa, M., Kuroe, A., et al. (1999). Glucose intolerance caused by a defect in the entero-insular axis: A study in gastric inhibitory polypeptide receptor knockout mice. *Proceedings of the National Academy of Sciences* 96, 14843–14847.

Miyoshi, H., Ajima, R., Luo, C.T., Yamaguchi, T.P., and Stappenbeck, T.S. (2012). Wnt5a potentiates TGF- $\beta$  signaling to promote colonic crypt regeneration after tissue injury. *Science* 338, 108–113.

Montgomery, R.K., Carlone, D.L., Richmond, C.A., Farilla, L., Kranendonk, M.E.G., Henderson, D.E., Baffour-Awuah, N.Y., Ambruzs, D.M., Fogli, L.K., Algra, S., et al. (2011). Mouse telomerase reverse transcriptase (mTert) expression marks slowly cycling intestinal stem cells. *Proceedings of the National Academy of Sciences* 108, 179–184.

Mooseker, M.S. (1984). Brush border cytoskeleton and integration of cellular functions. *The Journal of Cell Biology* 99, 104s – 112.

Mortensen, K., Christensen, L.L., Holst, J.J., and Orskov, C. (2003). GLP-1 and GIP are colocalized in a subset of endocrine cells in the small intestine. *Regulatory Peptides* 114, 189–196.

Moss, C.E., Marsh, W.J., Parker, H.E., Ogunnowo-Bada, E., Riches, C.H., Habib, A.M., Evans, M.L., Gribble, F.M., and Reimann, F. (2012). Somatostatin receptor 5 and cannabinoid receptor 1 activation inhibit secretion of glucose-dependent insulinotropic polypeptide from intestinal K cells in rodents. *Diabetologia* 55, 3094–3103.

Müller, T.S., Ebensperger, C., Neubüser, A., Koseki, H., Balling, R., Christ, B., and Wilting, J. (1996). Expression of Avian Pax1 and Pax9s Intrinsically Regulated in the Pharyngeal Endoderm, but Depends on Environmental Influences in the Paraxial Mesoderm. *Developmental Biology* 178, 403–417.

Múnera, J.O., Sundaram, N., Rankin, S.A., Hill, D., Watson, C., Mahe, M., Vallance, J.E., Shroyer, N.F., Sinagoga, K.L., Zarzoso-Lacoste, A., et al. (2017). Differentiation of Human Pluripotent Stem Cells into Colonic Organoids via Transient Activation of BMP Signaling. *Cell Stem Cell* 21, 51–64.e6.

Muñoz, J., Stange, D.E., Schepers, A.G., van de Wetering, M., Koo, B.-K., Itzkovitz, S., Volckmann, R., Kung, K.S., Koster, J., Radulescu, S., et al. (2012). The Lgr5 intestinal stem cell signature: robust expression of proposed quiescent '+4' cell markers: Transcriptomic and proteomic signature of Lgr5 + stem cells. *The EMBO Journal* 31, 3079–3091.

Mutoh, H., Naya, F.J., Tsai, M.-J., and Leiter, A.B. (1998). The basic helix-loop-helix protein BETA2 interacts with p300 to coordinate differentiation of secretin-expressing enteroendocrine cells. *Genes & Development* 12, 820–830.

---

## N

---

Nagatake, T., Fujita, H., Minato, N., and Hamazaki, Y. (2014). Enteroendocrine Cells Are Specifically Marked by Cell Surface Expression of Claudin-4 in Mouse Small Intestine. *PLoS ONE* 9, e90638.

Nagy, N., and Goldstein, A.M. (2017). Enteric nervous system development: A crest cell's journey from neural tube to colon. *Seminars in Cell & Developmental Biology* 66, 94–106.

Nakanishi, Y., Seno, H., Fukuoka, A., Ueo, T., Yamaga, Y., Maruno, T., Nakanishi, N., Kanda, K., Komekado, H., Kawada, M., et al. (2013). Dcl1 distinguishes between tumor and normal stem cells in the intestine. *Nature Genetics* 45, 98–103.

Naya, F.J., Huang, H.-P., Qiu, Y., Mutoh, H., DeMayo, F.J., Leiter, A.B., and Tsai, M.-J. (1997). Diabetes, defective pancreatic morphogenesis, and abnormal enteroendocrine differentiation in BETA2/NeuroD-deficient mice. *Genes & Development* 11, 2323–2334.

Naya, F.J., Stellrecht, C.M.M., and Tsai, M.-J. Tissue-specific regulation of the insulin gene by a novel basic helix-loop-helix transcription factor. *Development* 124, 1117–1126.

Nelson, S.B., Janiesch, C., and Sander, M. (2005). Expression of Nkx6 Genes in the Hindbrain and Gut of the Developing Mouse. *Journal of Histochemistry & Cytochemistry* 53, 787–790.

Netter, F.H. (2011). *Atlas d'anatomie humaine* (Philadelphia: Elsevier Masson).

Nguyen, T.L.A., Vieira-Silva, S., Liston, A., and Raes, J. (2015). How informative is the mouse for human gut microbiota research? *Disease Models & Mechanisms* 8, 1–16.

Noguchi, T.K., Ninomiya, N., Sekine, M., Komazaki, S., Wang, P.-C., Asashima, M., and Kurisaki, A. (2015). Generation of stomach tissue from mouse embryonic stem cells. *Nature Cell Biology* 17, 984–993.

Nokes, R.L., Fields, I.C., Collins, R.N., and Fölsch, H. (2008). Rab13 regulates membrane trafficking between TGN and recycling endosomes in polarized epithelial cells. *The Journal of Cell Biology* 182, 845–853.

---

O

Offield, M.F., Jetton, T.L., Labosky, P.A., Ray, M., Stein, R.W., Magnuson, M.A., Hogan, B.L.M., and Wright, C.V.E. PDX-1 is required for pancreatic outgrowth and differentiation of the rostral duodenum. 13.

Oidovsambuu, O., Nyamsuren, G., Liu, S., Göring, W., Engel, W., and Adham, I.M. (2011). Adhesion Protein VSIG1 Is Required for the Proper Differentiation of Glandular Gastric Epithelia. *PLoS ONE* 6, e25908.

Ootani, A., Li, X., Sangiorgi, E., Ho, Q.T., Ueno, H., Toda, S., Sugihara, H., Fujimoto, K., Weissman, I.L., Capecchi, M.R., et al. (2009). Sustained in vitro intestinal epithelial culture within a Wnt-dependent stem cell niche. *Nature Medicine* 15, 701–706.

---

P

Pabst, O., and Zweigerdt, R. (1999) Nkx2-3 function in development of mouse gut and spleen. 11.

Pancrazi, L., Di Benedetto, G., Colombaioni, L., Della Sala, G., Testa, G., Olimpico, F., Reyes, A., Zeviani, M., Pozzan, T., and Costa, M. (2015). Foxg1 localizes to mitochondria and coordinates cell differentiation and bioenergetics. *Proceedings of the National Academy of Sciences* 112, 13910–13915.

Paulus, U., Loeffler, M., Zeidler, J., Owen, G., and Potten, C.S. The differentiation and lineage development of goblet cells in the murine small intestinal crypt: experimental and modelling studies. 11.

Pelegri, F., Danilchik, M., and Sutherland, A. (2017). Vertebrate Development Maternal to Zygotic Control.

Penhoat, A., Fayard, L., Stefanutti, A., Mithieux, G., and Rajas, F. (2014). Intestinal gluconeogenesis is crucial to maintain a physiological fasting glycemia in the absence of hepatic glucose production in mice. *Metabolism* 63, 104–111.

Petersen, N., Reimann, F., Bartfeld, S., Farin, H.F., Ringnalda, F.C., Vries, R.G.J., van den Brink, S., Clevers, H., Gribble, F.M., and de Koning, E.J.P. (2014). Generation of L Cells in Mouse and Human Small Intestine Organoids. *Diabetes* 63, 410–420.

Petersen, N., Reimann, F., van Es, J.H., van den Berg, B.M., Kroone, C., Pais, R., Jansen, E., Clevers, H., Gribble, F.M., and de Koning, E.J.P. (2015). Targeting development of incretin-producing cells increases insulin secretion. *Journal of Clinical Investigation* 125, 379–385.

Petersen, N., Frimurer, T.M., Terndrup Pedersen, M., Egerod, K.L., Wewer Albrechtsen, N.J., Holst, J.J., Grapin-Botton, A., Jensen, K.B., and Schwartz, T.W. (2018). Inhibiting RHOA Signaling in Mice Increases Glucose Tolerance and Numbers of Enteroendocrine and Other Secretory Cells in the Intestine. *Gastroenterology* 155, 1164-1176.e2.

Piccand, J., Strasser, P., Hodson, D.J., Meunier, A., Ye, T., Keime, C., Birling, M.-C., Rutter, G.A., and Gradwohl, G. (2014). Rfx6 Maintains the Functional Identity of Adult Pancreatic  $\beta$  Cells. *Cell Reports* 9, 2219–2232.

Pinney, S.E., Oliver-Krasinski, J., Ernst, L., Hughes, N., Patel, P., Stoffers, D.A., Russo, P., and De León, D.D. (2011). Neonatal Diabetes and Congenital Malabsorptive Diarrhea Attributable to a Novel Mutation in the Human Neurogenin-3 Gene Coding Sequence. *The Journal of Clinical Endocrinology & Metabolism* 96, 1960–1965.

Pinto, D. (2003). Canonical Wnt signals are essential for homeostasis of the intestinal epithelium. *Genes & Development* 17, 1709–1713.

Ponder, B.A., Schmidt, G.H., Wilkinson, M.M., Wood, M.J., Monk, M., and Reid, A. (1985). Derivation of mouse intestinal crypts from single progenitor cells. *Nature* 313, 689–691.

Posovszky, C., and Wabitsch, M. (2014). Regulation of Appetite, Satiation, and Body Weight by Enteroendocrine Cells. Part 1: Characteristics of Enteroendocrine Cells and Their Capability of Weight Regulation. *Hormone Research in Paediatrics* 83, 1–10.

Potten, C.S. (1998). Stem cells in gastrointestinal epithelium: numbers, characteristics and death. *Philosophical Transactions of the Royal Society B: Biological Sciences* 353, 821–830.

Potten, C.S., Owen, G., and Booth, D. (2002). Intestinal stem cells protect their genome by selective segregation of template DNA strands. *J. Cell. Sci.* 115, 2381–2388.

---

## R

---

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013a). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols* 8, 2281–2308.

Ran, F.A., Hsu, P.D., Lin, C.-Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., et al. (2013b). Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. *Cell* 154, 1380–1389.

Ratineau, C., Petry, M.W., Mutoh, H., and Leiter, A.B. (2002). Cyclin D1 Represses the Basic Helix-Loop-Helix Transcription Factor, BETA2/NeuroD. *Journal of Biological Chemistry* 277, 8847–8853.

Richards, P., Pais, R., Habib, A.M., Brighton, C.A., Yeo, G.S.H., Reimann, F., and Gribble, F.M. (2016). High fat diet impairs the function of glucagon-like peptide-1 producing L-cells. *Peptides* 77, 21–27.

Rindi, G. (1999) Secretin cell ablation. 8.

Ritze, Y., Hengelhaupt, C., Bárdos, G., Ernst, B., Thurnheer, M., D’Haese, J.G., Bischoff, S.C., and Schultes, B. (2015). Altered intestinal neuroendocrine gene expression in humans with obesity: Neuroendocrine Genes Are Altered in Obesity. *Obesity* 23, 2278–2285.

Roberts, G.P., Larraufie, P., Richards, P., Kay, R.G., Galvin, S.G., Miedzybrodzka, E.L., Leiter, A., Li, J.H., Glass, L.L., Ma, M.K., et al. (2018). Comparison of human and murine enteroendocrine cells by transcriptomic and peptidomic profiling.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. *Nature Biotechnology* 29, 24–26.

Rothenberg, M.E., Nusse, Y., Kalisky, T., Lee, J.J., Dalerba, P., Scheeren, F., Lobo, N., Kulkarni, S., Sim, S., Qian, D., et al. (2012). Identification of a cKit+ Colonic Crypt Base Secretory Cell That Supports Lgr5+ Stem Cells in Mice. *Gastroenterology* 142, 1195–1205.e6.

Rubin, D.C. (2007). Intestinal morphogenesis: Current Opinion in Gastroenterology 23, 111–114.

Rubio-Cabezas, O., Jensen, J.N., Hodgson, M.I., Codner, E., Ellard, S., Serup, P., and Hattersley, A.T. (2011). Permanent Neonatal Diabetes and Enteric Anendocrinosis Associated With Biallelic Mutations in NEUROG3. *Diabetes* 60, 1349–1353.

Russell, W.C., Graham, F.L., Smiley, J., and Nairn, R. (1977). Characteristics of a Human Cell Line Transformed by DNA from Human Adenovirus Type 5. *Journal of General Virology* 36, 59–72.

## S

- Salzman, N.H. (2011). Microbiota-immune system interaction: an uneasy alliance. *Current Opinion in Microbiology* 14, 99–105.
- Sancho, E., Batlle, E., and Clevers, H. (2004). SIGNALING PATHWAYS IN INTESTINAL DEVELOPMENT AND CANCER. *Annual Review of Cell and Developmental Biology* 20, 695–723.
- Sansbury, F.H., Kirel, B., Caswell, R., Lango Allen, H., Flanagan, S.E., Hattersley, A.T., Ellard, S., and Shaw-Smith, C.J. (2015). Biallelic RFX6 mutations can cause childhood as well as neonatal onset diabetes mellitus. *European Journal of Human Genetics* 23, 1744–1748.
- de Santa Barbara, P., van den Brink, G.R., and Roberts, D.J. (2003). Development and differentiation of the intestinal epithelium. *Cellular and Molecular Life Sciences (CMLS)* 60, 1322–1332.
- Sato, T., and Clevers, H. (2013). Growing Self-Organizing Mini-Guts from a Single Intestinal Stem Cell: Mechanism and Applications. *Science* 340, 1190–1194.
- Sato, T., Vries, R.G., Snippert, H.J., van de Wetering, M., Barker, N., Stange, D.E., van Es, J.H., Abo, A., Kujala, P., Peters, P.J., et al. (2009). Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* 459, 262–265.
- Sato, T., Stange, D.E., Ferrante, M., Vries, R.G.J., van Es, J.H., van den Brink, S., van Houdt, W.J., Pronk, A., van Gorp, J., Siersema, P.D., et al. (2011a). Long-term Expansion of Epithelial Organoids From Human Colon, Adenoma, Adenocarcinoma, and Barrett's Epithelium. *Gastroenterology* 141, 1762–1772.
- Sato, T., van Es, J.H., Snippert, H.J., Stange, D.E., Vries, R.G., van den Born, M., Barker, N., Shroyer, N.F., van de Wetering, M., and Clevers, H. (2011b). Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature* 469, 415–418.
- Sayar, E., Islek, A., Yilmaz, A., Akcam, M., Flanagan, S.E., and Artan, R. (2013). Extremely rare cause of congenital diarrhea: Enteric anendocrinosis: Rare cause of congenital diarrhea. *Pediatrics International* 55, 661–663.
- Schild, H., and Rammensee, H.-G. (2000). gp96—The immune system's Swiss army knife. *Nature Immunology* 1, 100–101.
- Schnerch, A., Cerdan, C., and Bhatia, M. (2010). Distinguishing Between Mouse and Human Pluripotent Stem Cell Regulation: The Best Laid Plans of Mice and Men. *STEM CELLS N/A-N/A*.
- Schonhoff, S.E., Giel-Moloney, M., and Leiter, A.B. (2004). Neurogenin 3-expressing progenitor cells in the gastrointestinal tract differentiate into both endocrine and non-endocrine cell types. *Developmental Biology* 270, 443–454.
- Schulze, S.K., Kanwar, R., Gölzenleuchter, M., Therneau, T.M., and Beutler, A.S. (2012). SERE: Single-parameter quality control and sample comparison for RNA-Seq. *BMC Genomics* 13, 524.
- Scrocchi, L.A., Brown, T.J., McClusky, N., Brubaker, P.L., Auerbach, A.B., Joyner, A.L., and Drucker, D.J. (1996). Glucose intolerance but normal satiety in mice with a null mutation in the glucagon-like peptide 1 receptor gene. *Nat. Med.* 2, 1254–1258.
- Shaw, E.J., Smith, E.E., Whittingham-Dowd, J., Hodges, M.D., Else, K.J., and Rigby, R.J. (2017). Intestinal epithelial suppressor of cytokine signaling 3 (SOCS3) impacts on mucosal homeostasis in a model of chronic inflammation: SOCS3 limits intestinal epithelial turnover andIDO expression. *Immunity, Inflammation and Disease* 5, 336–345.

Shen, Z., Zhang, X., Chai, Y., Zhu, Z., Yi, P., Feng, G., Li, W., and Ou, G. (2014). Conditional Knockouts Generated by Engineered CRISPR-Cas9 Endonuclease Reveal the Roles of Coronin in *C. elegans* Neural Development. *Developmental Cell* 30, 625–636.

Shroyer, N.F., Helmrath, M.A., Wang, V.Y. –C., Antalffy, B., Henning, S.J., and Zoghbi, H.Y. (2007). Intestine-Specific Ablation of Mouse atonal homolog 1 (*Math1*) Reveals a Role in Cellular Homeostasis. *Gastroenterology* 132, 2478–2488.

Sikandar, S.S., Pate, K.T., Anderson, S., Dizon, D., Edwards, R.A., Waterman, M.L., and Lipkin, S.M. (2010). NOTCH Signaling Is Required for Formation and Self-Renewal of Tumor-Initiating Cells and for Repression of Secretory Cell Differentiation in Colon Cancer. *Cancer Research* 70, 1469–1478.

Simon-Assmann, P., and Kedinger, M. (1993). Heterotypic cellular cooperation in gut morphogenesis and differentiation. *Semin. Cell Biol.* 4, 221–230.

Sinagoga, K.L., McCauley, H.A., Múnera, J.O., Reynolds, N.A., Enriquez, J.R., Watson, C., Yang, H.-C., Helmrath, M.A., and Wells, J.M. (2018). Deriving functional human enteroendocrine cells from pluripotent stem cells. *Development* 145, dev165795.

Skopkova, M., Ciljakova, M., Havlicekova, Z., Vojtkova, J., Valentinova, L., Danis, D., Murgas, D., Szepeova, R., Stanik, J., Banovcin, P., et al. (2016). Two novel RFX6 variants in siblings with Mitchell-Riley syndrome with later diabetes onset and heterotopic gastric mucosa. *European Journal of Medical Genetics* 59, 429–435.

Smith, S.B., Gasa, R., Watada, H., Wang, J., Griffen, S.C., and German, M.S. (2003). Neurogenin3 and Hepatic Nuclear Factor 1 Cooperate in Activating Pancreatic Expression of Pax4. *Journal of Biological Chemistry* 278, 38254–38259.

Smith, S.B., Watada, H., and German, M.S. (2004). Neurogenin3 Activates the Islet Differentiation Program while Repressing Its Own Expression. *Molecular Endocrinology* 18, 142–149.

Smith, S.B., Qu, H.-Q., Taleb, N., Kishimoto, N.Y., Scheel, D.W., Lu, Y., Patch, A.-M., Grabs, R., Wang, J., Lynn, F.C., et al. (2010). Rfx6 directs islet formation and insulin production in mice and humans. *Nature* 463, 775–780.

Sommer, L., Ma, Q., and Anderson, D.J. (1996). neurogenins, a Novel Family of atonal-Related bHLH Transcription Factors, Are Putative Mammalian Neuronal Determination Genes That Reveal Progenitor Cell Heterogeneity in the Developing CNS and PNS. *Molecular and Cellular Neuroscience* 8, 221–241.

Sosa-Pineda, B., Chowdhury, K., Torres, M., Oliver, G., and Gruss, P. (1997). The Pax4 gene is essential for differentiation of insulin-producing beta cells in the mammalian pancreas. *Nature* 386, 399–402.

Soyer, J., Flasse, L., Raffelsberger, W., Beucher, A., Orvain, C., Peers, B., Ravassard, P., Vermot, J., Voz, M.L., Mellitzer, G., et al. (2010). Rfx6 is an Ngn3-dependent winged helix transcription factor required for pancreatic islet cell development. *Development* 137, 203–212.

Spence, J.R., Mayhew, C.N., Rankin, S.A., Kuhar, M.F., Vallance, J.E., Tolle, K., Hoskins, E.E., Kalinichenko, V.V., Wells, S.I., Zorn, A.M., et al. (2011). Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro. *Nature* 470, 105–109.

Spiegel, R., Dobbie, A., Hartman, C., de Vries, L., Ellard, S., and Shalev, S.A. (2011). Clinical characterization of a newly described neonatal diabetes syndrome caused by RFX6 mutations. *American Journal of Medical Genetics Part A* 155, 2821–2825.

Stainier, D.Y.R. (2002). A glimpse into the molecular entrails of endoderm formation. *Genes & Development* 16, 893–907.

Strege, P.R., Knutson, K., Eggers, S.J., Li, J.H., Wang, F., Linden, D., Szurszewski, J.H., Milesco, L., Leiter, A.B., Farrugia, G., et al. (2017). Sodium channel NaV1.3 is important for enterochromaffin cell excitability and serotonin release. *Scientific Reports* 7.

Stringer, E.J., Pritchard, C.A., and Beck, F. (2008). Cdx2 initiates histodifferentiation of the midgut endoderm. *FEBS Letters* 582, 2555–2560.

Strømme, P., Mangelsdorf, M.E., Shaw, M.A., Lower, K.M., Lewis, S.M.E., Bruyere, H., Lütcherath, V., Gedeon, A.K., Wallace, R.H., Scheffer, I.E., et al. (2002). Mutations in the human ortholog of *Aristaless* cause X-linked mental retardation and epilepsy. *Nature Genetics* 30, 441–445.

Sui, L., Bouwens, L., and Mfopou, J.K. (2013). Signaling pathways during maintenance and definitive endoderm differentiation of embryonic stem cells. *The International Journal of Developmental Biology* 57, 1–12.

Suzuki, K., Harada, N., Yamane, S., Nakamura, Y., Sasaki, K., Nasteska, D., Joo, E., Shibue, K., Harada, T., Hamasaki, A., et al. (2013). Transcriptional Regulatory Factor X6 (Rfx6) Increases Gastric Inhibitory Polypeptide (GIP) Expression in Enteroendocrine K-cells and Is Involved in GIP Hypersecretion in High Fat Diet-induced Obesity. *Journal of Biological Chemistry* 288, 1929–1938.

---

## T

---

Takahashi, K., Okita, K., Nakagawa, M., and Yamanaka, S. (2007a). Induction of pluripotent stem cells from fibroblast cultures. *Nature Protocols* 2, 3081–3089.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007b). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* 131, 861–872.

Tanaka, T., Kishi, K., Igawa, M., Takase, S., and Goda, T. (1998). Dietary carbohydrates enhance lactase/phlorizin hydrolase gene expression at a transcription level in rat jejunum. *Biochemical Journal* 331, 225–230.

Terry, N.A., Walp, E.R., Lee, R.A., Kaestner, K.H., and May, C.L. (2014). Impaired enteroendocrine development in intestinal-specific *Islet1* mouse mutants causes impaired glucose homeostasis. *American Journal of Physiology-Gastrointestinal and Liver Physiology* 307, G979–G991.

Terry, N.A., Lee, R.A., Walp, E.R., Kaestner, K.H., and Lee May, C. (2015). Dysgenesis of Enteroendocrine Cells in *Aristaless*-Related Homeobox Polyalanine Expansion Mutations: *Journal of Pediatric Gastroenterology and Nutrition* 60, 192–199.

Terry, N.A., Ngaba, L.V., Wilkins, B.J., Pi, D., Gheewala, N., and Kaestner, K.H. (2018). Lipid malabsorption from altered hormonal signaling changes early gut microbial responses. *American Journal of Physiology-Gastrointestinal and Liver Physiology* 315, G580–G591.

Tetteh, P.W., Basak, O., Farin, H.F., Wiebrands, K., Kretschmar, K., Begthel, H., van den Born, M., Korving, J., de Sauvage, F., van Es, J.H., et al. (2016). Replacement of Lost Lgr5-Positive Stem Cells through Plasticity of Their Enterocyte-Lineage Daughters. *Cell Stem Cell* 18, 203–213.

Thomson, J.A. (1998). Embryonic Stem Cell Lines Derived from Human Blastocysts. *Science* 282, 1145–1147.

Tian, H., Biehs, B., Warming, S., Leong, K.G., Rangell, L., Klein, O.D., and de Sauvage, F.J. (2011). A reserve stem cell population in small intestine renders Lgr5-positive cells dispensable. *Nature* 478, 255–259.

Tomic, G., Morrissey, E., Kozar, S., Ben-Moshe, S., Hoyle, A., Azzarelli, R., Kemp, R., Chilamakuri, C.S.R., Itzkovitz, S., Philpott, A., et al. (2018). Phospho-regulation of ATOH1 Is Required for Plasticity of Secretory Progenitors and Tissue Regeneration. *Cell Stem Cell* 23, 436-443.e7.

Trabelsi, M.-S., Daoudi, M., Prawitt, J., Ducastel, S., Touche, V., Sayin, S.I., Perino, A., Brighton, C.A., Sebti, Y., Kluza, J., et al. (2015). Farnesoid X receptor inhibits glucagon-like peptide-1 production by enteroendocrine L cells. *Nature Communications* 6.

Tsai, S.Q., and Joung, J.K. (2016). Defining and improving the genome-wide specificities of CRISPR–Cas9 nucleases. *Nature Reviews Genetics* 17, 300–312.

Tsai, Y.-H., Nattiv, R., Dedhia, P.H., Nagy, M.S., Chin, A.M., Thomson, M., Klein, O.D., and Spence, J.R. (2017). In vitro patterning of pluripotent stem cell-derived intestine recapitulates in vivo human development. *Development* 144, 1045–1055.

Tuazon, F.B., and Mullins, M.C. (2015). Temporally coordinated signals progressively pattern the anteroposterior and dorsoventral body axes. *Seminars in Cell & Developmental Biology* 42, 118–133.

---

## U

---

Ueo, T., Imayoshi, I., Kobayashi, T., Ohtsuka, T., Seno, H., Nakase, H., Chiba, T., and Kageyama, R. (2012). The role of Hes genes in intestinal development, homeostasis and tumor formation. *Development* 139, 1071–1082.

Umar, S. (2010). Intestinal Stem Cells. *Current Gastroenterology Reports* 12, 340–348.

Ussar, S., Haering, M.-F., Fujisaka, S., Lutter, D., Lee, K.Y., Li, N., Gerber, G.K., Bry, L., and Kahn, C.R. (2017). Regulation of Glucose Uptake and Enteroendocrine Function by the Intestinal Epithelial Insulin Receptor. *Diabetes* 66, 886–896.

---

## V

---

Verdich, C., Toubro, S., Buemann, B., Madsen, J.L., Holst, J.J., and Astrup, A. The role of postprandial releases of insulin and incretin hormones in meal-induced satiety & effect of obesity and weight reduction. *International Journal of Obesity* 9.

Volovitz, I., Shapira, N., Ezer, H., Gafni, A., Lustgarten, M., Alter, T., Ben-Horin, I., Barzilai, O., Shahar, T., Kanner, A., et al. (2016). A non-aggressive, highly efficient, enzymatic method for dissociation of human brain-tumors and brain-tissues to viable single-cells. *BMC Neuroscience* 17.

Vouillot, L., Thélie, A., and Pollet, N. (2015). Comparison of T7E1 and Surveyor Mismatch Cleavage Assays to Detect Mutations Triggered by Engineered Nucleases. *Genes & Genomes Genetics* 5, 407–415.

---

## W

---

Waring, R. (2006). Mutant Neurogenin-3 in Congenital Malabsorptive Diarrhea. *N Engl J Med* 11.

Watada, H., Scheel, D.W., Leung, J., and German, M.S. (2003). Distinct Gene Expression Programs Function in Progenitor and Mature Islet Cells. *Journal of Biological Chemistry* 278, 17130–17140.

Watson, C.L., Mahe, M.M., Múnera, J., Howell, J.C., Sundaram, N., Poling, H.M., Schweitzer, J.I., Vallance, J.E., Mayhew, C.N., Sun, Y., et al. (2014). An in vivo model of human small intestine using pluripotent stem cells. *Nature Medicine* 20, 1310–1314.

Wells, J.M., and Melton, D.A. (1999). Vertebrate Endoderm Development. *Annual Review of Cell and Developmental Biology* 15, 393–410.

van de Wetering, M., Sancho, E., Verweij, C., de Lau, W., Oving, I., Hurlstone, A., Batlle, E., Coudreuse, D., Haramis, A.-P., Tjon-Pon-Fong, M., et al. The  $\beta$ -Catenin/TCF-4 Complex Imposes a Crypt Progenitor Phenotype on Colorectal Cancer Cells. 10.

Whitman, M. Smads and early developmental signaling by the TGF  $\beta$  superfamily. 19.

Willet, S.G., and Mills, J.C. (2016). Stomach Organ and Cell Lineage Differentiation: From Embryogenesis to Adult Homeostasis. *Cellular and Molecular Gastroenterology and Hepatology* 2, 546–559.

Winslow, J.W., Moran, P., Valverde, J., Shih, A., Yuan, J.O., Wong, S.C., Tsai, S.P., Goddard, A., Henzel, W.J., Hefti, F., et al. (1995). Cloning of AL-1, a ligand for an Eph-related tyrosine kinase receptor involved in axon bundle formation. *Neuron* 14, 973–981.

Wong, W.M., and Wright, N.A. (1999). Cell proliferation in gastrointestinal mucosa. *Journal of Clinical Pathology* 52, 321–333.

Workman, M.J., Mahe, M.M., Trisno, S., Poling, H.M., Watson, C.L., Sundaram, N., Chang, C.-F., Schiesser, J., Aubert, P., Stanley, E.G., et al. (2017). Engineered human pluripotent-stem-cell-derived intestinal tissues with a functional enteric nervous system. *Nature Medicine* 23, 49–59.

Worthington, J.J., Reimann, F., and Gribble, F.M. (2018). Enteroendocrine cells-sensory sentinels of the intestinal environment and orchestrators of mucosal immunity. *Mucosal Immunol* 11, 3–20.

---

## X

---

Xi, J., Yue, J., and Yang, Z. (2015). Expression profiles of retinoic acid synthetases ALDH1As and metabolic enzymes CYP26s in adult and embryonic zebrafish (*Danio rerio*). *Genetics and Molecular Research* 14, 3948–3956.

Xu, Q., Mellitzer, G., Robinson, V., and Wilkinson, D.G. (1999). In vivo cell sorting in complementary segmental domains mediated by Eph receptors and ephrins. *Nature* 399, 267–271.

---

## Y

---

Yan, K.S., Chia, L.A., Li, X., Ootani, A., Su, J., Lee, J.Y., Su, N., Luo, Y., Heilshorn, S.C., Amieva, M.R., et al. (2012). The intestinal stem cell markers *Bmi1* and *Lgr5* identify two functionally distinct populations. *Proceedings of the National Academy of Sciences* 109, 466–471.

Yan, K.S., Gevaert, O., Zheng, G.X.Y., Anchang, B., Probert, C.S., Larkin, K.A., Davies, P.S., Cheng, Z., Kaddis, J.S., Han, A., et al. (2017). Intestinal Enteroendocrine Lineage Cells Possess Homeostatic and Injury-Inducible Stem Cell Activity. *Cell Stem Cell* 21, 78-90.e6.

Yan, Y., Frisen, J., Lee, M.H., Massague, J., and Barbacid, M. (1997). Ablation of the CDK inhibitor *p57Kip2* results in increased apoptosis and delayed differentiation during mouse development. *Genes & Development* 11, 973–983.

Yang, Q. (2001). Requirement of *Math1* for Secretory Cell Lineage Commitment in the Mouse Intestine. *Science* 294, 2155–2158.

Ye, D.Z., and Kaestner, K.H. (2009). Foxa1 and Foxa2 Control the Differentiation of Goblet and Enteroendocrine L- and D-Cells in Mice. *Gastroenterology* 137, 2052–2062.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. *Science* 318, 1917–1920.

Yuasa, Y. (2003). Control of gut differentiation and intestinal-type gastric carcinogenesis. *Nature Reviews Cancer* 3, 592–600.

---

## Z

---

Zietek, T., Rath, E., Haller, D., and Daniel, H. (2015). Intestinal organoids for assessing nutrient transport, sensing and incretin secretion. *Scientific Reports* 5.

Zorn, A.M., and Wells, J.M. (2007). Molecular Basis of Vertebrate Endoderm Development. In *International Review of Cytology*, (Elsevier), pp. 49–111.

Zorn, A.M., and Wells, J.M. (2009). Vertebrate Endoderm Development and Organ Formation. *Annual Review of Cell and Developmental Biology* 25, 221–251.

# ANNEXES

**Annexe 1: Tableau récapitulatif des phénotypes observés dans les modèles déficients pour nos gènes d'intérêt (*Ngn3*, *Arx*, *Pax4*), décrits dans la littérature**

		Cellules entéroendocrines et facteurs de transcription associés						Autres types cellulaires								
Gène	Génotype	Espèce	Modèle	Tissu	Non différent	Augmentation	Diminution	Absence	Cellules à mucus	Cellules de Paneth	Entérocytes	Anomalies morphologiques	Anomalies physiologiques	Références	Survie	
NG3	Ngn3 <sup>-/-</sup>	Souris	<i>in vivo</i>	intestin				toutes les cellules EE	↗	=	=	taille estomac réduite estomac: métaplasie intestinale	diabète	Jenny et al., 2002 Lee et al., 2002 Gradwohl et al., 2000	Mort postnatale	
				estomac	ghréline, HDC*		5-HT	SST, gastrine, glucagon								
NG3	Ngn3 <sup>Δint</sup>	Souris	<i>in vivo</i>	intestin				toutes les cellules EE	=	=	microvillosités plus courtes	mutants de plus petite taille villosités atrophiées et dilatées détachement épithélial de la membrane basale ↗ taille cryptes ↘ bordure en brosse	↗ renouvellement cellulaire ↗ prolifération stéatorrhée perturbation de l'homéostasie du glucose	Mellitzer et al., 2010	Mort post-natale à 50% des souris	
				HICs				absence des cellules EE? (CHGA)			ND	ND	ND			Spence et al., 2011
ARX	Arx <sup>-/-</sup>	Souris	<i>in vivo</i>	intestin	Tph1, 5-HT, Pax4, Pax6, Insml, Pdx1 (ig)	SST, ghrléline (duo), Pdx1 (colon), Foxa1/2 (ig)	Sécrétine, gastrine (ig)	GLP1, GIP, CCK, PYY, NTS	=	ND	ND		retard de croissance	déshydratation, hypoglycémie	Beucher et al., 2012 Colombat et al., 2003	Mort postnatale P2
				intestin	ghrléline, Tph1, gastrine	SST	Préproglucagon, GIP, CCK, sécrétine, NTS, PYY			ND	↗ Lysozyme		présence vécule	stéatorrhée		Terry et al., 2018
PAX4	Pax4 <sup>-/-</sup>	Souris	<i>in vivo</i>	HICs	SST, ghrléline		CCK, Sécrétine, Préproglucagon (tendance)		=	=	=				Du et al., 2012	
				intestin	CCK, sécrétine, NTS (colon), Insml, Pdx1, Foxa1/2	GLP1, PYY (colon), Ghrléline, Arx	NTS (ig), gastrine (ig), sécrétine (ig), Pax6 (ig)	GIP, Tph1, SST		ND	ND	ND		diabète		Beucher et al., 2012 Napolitano et al., 2015

\* Histidine décarboxylase ⇒ histamine

**Annexe 2: Protocole de différenciation  
intestinale *in vitro***

# ***In vitro* Generation of Intestinal Organoids from Human Pluripotent Stem Cells**

(Adapted from Mùnera and Wells, 2017)

## **I. hiPSC culture**

1. Appropriate Matrigel-coated dishes can be ordered from cell culture facility and kept at 4 °C for up to 1 week (BD Matrigel hESC-qualified Matrix #354277).

*Note: Order Matrigel-coated dishes at least one day prior to the planned date of passaging, as the Matrigel-coated dish can be used the next day only after coating.*

2. Grow hiPS cells in 2 mL mTESR1 medium on hESC-qualified Matrigel coated P35 dishes.

3. Check cells daily for differentiation and remove differentiated colonies if necessary.

4. Replace medium daily and passage cells after 5–6 days.

5. An hour before passaging the cells, take out Matrigel-coated dishes, DMEM-F12 and an aliquot of dispase and keep them at room temperature (inside flow hood).

6. Passage cells by aspirating medium and washing once with DMEM-F12.

7. Aspirate DMEM-F12 and add 1 mL of dispase solution.

8. Place plates in a 5% CO<sub>2</sub> incubator at 37 °C for 4 min. After 4 min observe for folding of the edges of colonies.

*Note: While your cells are being incubated, replace Matrigel with 2ml of mTESR1 medium in the dishes in which cells will be seeded.*

9. Aspirate dispase and wash 3 times with DMEM-F12.

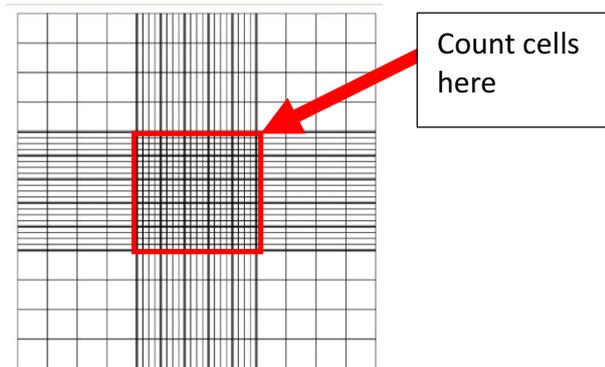
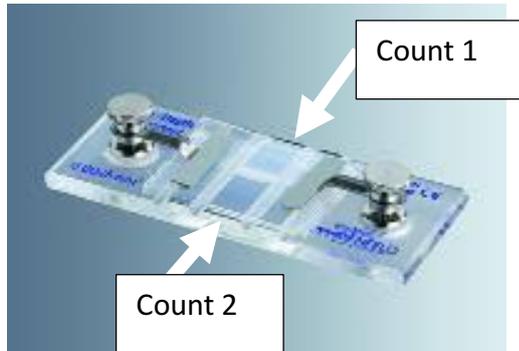
- 10.** Aspirate DMEM-F12 and add 1 mL of mTESR1.
- 11.** Use a cell scraper to lift colonies off the plate.
- 12.** Add 2 mL mTESR1 and break colonies into small chunks by pipetting up and down 3 times while holding pipette against the side of the plate (use a 5 mL-pipette). Observe under microscope to ensure that no large clumps remain.
- 13.** Plate cells into new Matrigel-coated P35 by dispensing 250  $\mu$ L of cells into each P35 containing 2 mL of mTESR1.
- 14.** To ensure even dispersal of cells in the plates, shake carefully the plates. Check cells the next day and repeat **steps 1–13** as needed.

## **II. Single-cell plating of hiPSCs**

- 1.** Grow hiPS cells to 70–80% confluence and remove differentiated cells before starting. Wash cells once with 2 mL DMEM-F12.
- 2.** Add 1 mL of Accutase and place in incubator for 6–7 min (until cells detach).
- 3.** Add 2 mL of DMEM-F12 and pipette up and down using a 1000  $\mu$ L pipette to detach adherent cells.
- 4.** Pool cells into a 15-mL Falcon tube and mix by pipetting up and down to break cells into single cells.
- 5.** Spin down cells at 200  $\times g$  for 4 min.
- 6.** Remove supernatant without disturbing pellet and resuspend in 1 mL mTESR1 supplemented with 10- $\mu$ M Y-27632 ROCK inhibitor. Mix cells well and take out small aliquot (50 $\mu$ l) in one well of 96 well plate for cell counting.

## 7. Count cells in Neubauer improved counting chambers:

- Mix 10 $\mu$ l of cell suspension and 10 $\mu$ l of trypan blue in a well of 96-well plate. Load 10 $\mu$ l of this in top grid of the counting chamber. Again, Mix 10 $\mu$ l of cell suspension and 10 $\mu$ l of trypan blue in another well of 96-well plate. Load 10 $\mu$ l of this in bottom grid of the counting chamber.



- Count cells in both top (count 1) and bottom (count 2) chambers inside the red square. Calculate mean of two counts.

Cell concentration = mean cell count X dilution factor X 10<sup>4</sup> cells/ml

**8.** Plate 52 600 cells/cm<sup>2</sup>: 100,000 cells per well or 420 000 cells per P35 dish. Add 2 mL of mTESR1 supplemented with 10- $\mu$ M Y-27632 ROCK inhibitor if using P35 or 500  $\mu$ L if using 4-well plates.

**9.** To ensure even dispersal of cells in the plates, shake carefully the plates. Allow cells to reach a 90-95% confluency (2 to 3 days) and start differentiation.

## **III. Definitive endoderm induction**

**1.** Replace medium with Day 1 medium : RPMI 1640; L-glutamine 2 mM; Activin A 100ng/ml.

**2.** After 24 h in Day 1 medium, cell death should be evident. Aspirate Day 1 medium and replace with Day 2 medium : RPMI 1640; L-glutamine 2 mM; Activin A 100ng/ml; dFBS 0.2%.

**3.** After 48 h cell death should still be evident. Aspirate Day 2 medium and replace with Day 3 medium : RPMI 1640; L-glutamine 2 mM; Activin A 100ng/ml; dFBS 2%.

## **IV. Mid-hindgut formation**

- 1.** Aspirate Day 3 medium and replace with mid-hindgut (MHG) differentiation medium : RPMI 1640; L-glutamine 2mM; CHIR99021 3  $\mu$ M; FGF4 500ng/ml; dFBS 2%. Change medium every 24 h for a total 4 days.
  - 2.** After 48 h floating three-dimensional spheres will already be present (very few and occasionally may not be present as well). Collect spheroids (see **steps 4-8**) and replace MHG medium.
- Optional step:* if you have small spheroids, filter them through a 70  $\mu$ m cell strainer to collect only bigger spheroids.
- 3.** After 72 h more floating spheres should be evident. Replace MHG medium. After 96 h hundreds of floating spheres should be evident. If spheroids remain attached to the cell layer: pick them carefully with a needle + syringe.
  - 4.** Pre-warm 4-well Nunc plates or P35 dishes by placing them in the incubator. Thaw on ice as many Matrigel aliquots you need (Corning® Matrigel® Growth Factor Reduced Basement Membrane Matrix, Phenol Red-Free, \*LDEV-Free #356231)
  - 5.** Collect spheroids in maximum 25  $\mu$ L and put them into an aliquot of 50  $\mu$ L Matrigel. Mix by slowly pipetting up and down 3–5 times. Be careful not to introduce bubbles.

*Optional step:* if you have small spheroids, filter them through a 70  $\mu$ m cell strainer to collect only bigger spheroids.

*Notes: to take as less medium as possible during spheroid collection, first collect the medium with floating spheroids in a new P35 dish and let spheroids settle at the bottom of the dish. Swirl it in order to move spheroids to the center of plate and collect them.*

- 6.** Plate spheroids by putting the 75  $\mu$ l at the bottom of a new 4 well Nunc plate in the middle of the well to form a dome of matrigel + spheroids. Adapt volumes if you are using P35 dishes.

7. Place plates for 10 min back in the incubator to allow Matrigel to solidify (being careful not to disturb the Matrigel droplets).

8. Add 500  $\mu$ l per well of intestine differentiation medium (IDM) : Advanced DMEM F12; B27 (1/50); R-Spondin1 500ng/ml; Noggin 100ng/ml; EGF 50ng/ml; HEPES 15mM; P/S (1/100).

## **V. Organoids splitting**

1. After approximately 14 days, organoids will have degraded the Matrigel droplet. To allow further expansion, replate at a lower density (5–10 organoids per well) into fresh Matrigel.

2. Cut the tip of a 200  $\mu$ L pipette to increase the bore size of the pipette. Pipette the Matrigel bead containing organoids using the cut 200  $\mu$ L tip and put them into a 10 cm petri dish with 10 mL of warm Advanced DMEM-F12.

3. Using a sterile disposable scalpel and a forceps, dislodge any large pieces of Matrigel which are still attached to the organoids.

4. Collect 5–10 organoids in as little medium as possible and replate them as it was done in **steps 5–8 (part IV)**.

5. Place plates containing organoids back into the incubator and grow tissue for an additional 14 days.

6. After 14 days, organoids need to be splitted in two. Proceed as in **steps 2–5** and cut organoids in two halves at step 3.

7. Replace medium every 4 days and split organoids every 14 days.

**Annexe 3: Liste des potentiels sites de fixation  
non-spécifique des sgRNAs**

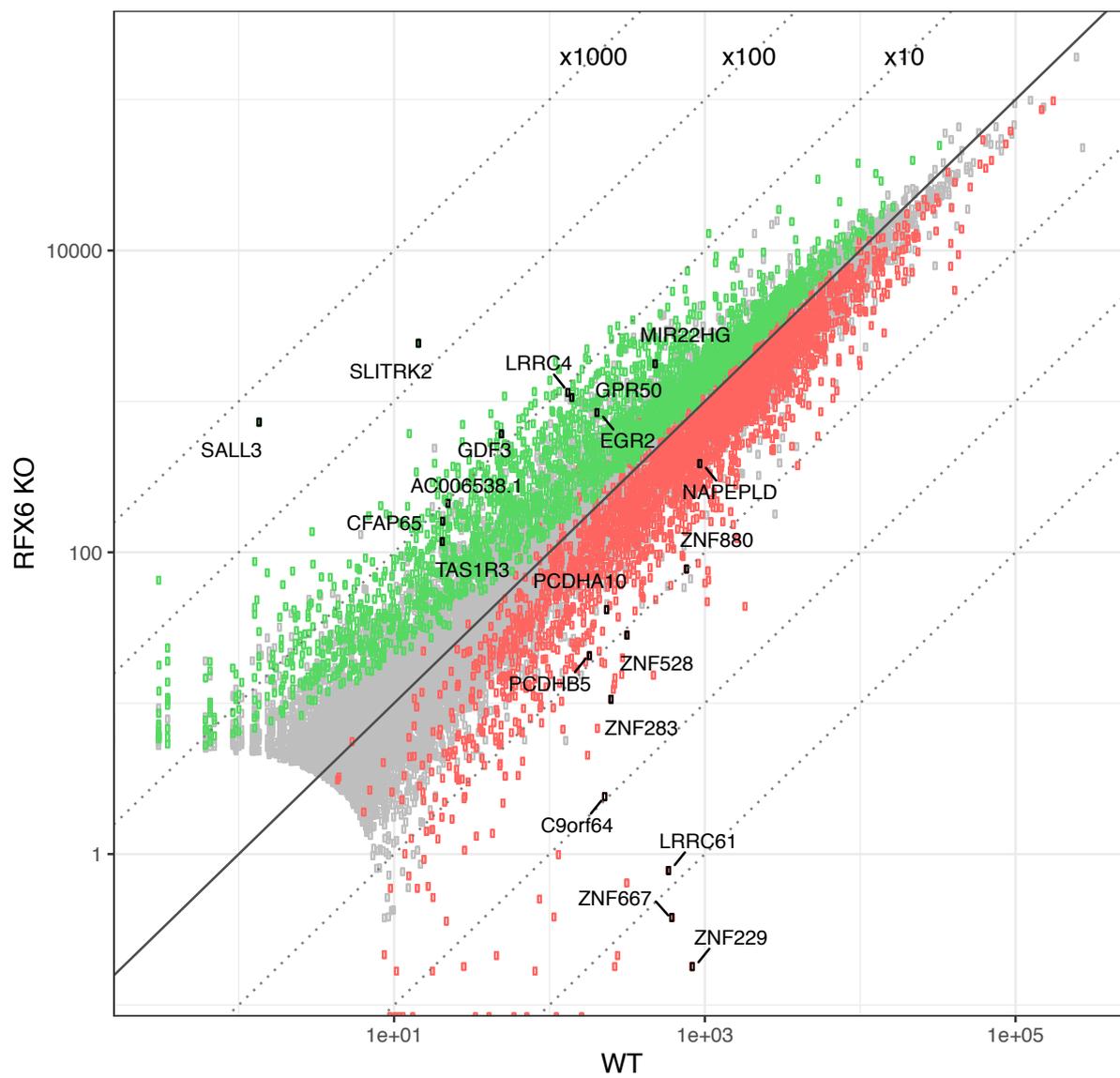
sgRNA#17

chr	strand	position	sequence	n_mismatches	score	ontarget	gene
chr5	1	165008607	TGTCCAAGTGACCCGTGAGATAG	0	100.0	True	None
chr18	-1	40150892	TAACCAAGTGACCCGTGAGGAGG	3	1.47120112994	False	None
chrX	1	88148873	TGAACAACCTACCCGTGAGAGGG	4	1.25880452335	False	None
chr2	1	152095136	AATCCAAGTGACCCCTGAGATGG	3	0.844444444444	False	None
chr8	-1	70952361	GGGACAAGTGCCCCGTGAGATAG	4	0.795717091837	False	None
chr22	-1	27046038	CATCCCTGTGACCCGTGAGAGAG	4	0.564014727011	False	None
chr8	1	109717202	TGGCCCTTTGACCCGTGAGATGG	4	0.547724279009	False	None
chr12	-1	102497469	GGGCCTAGGGACCCGTGAGAGGG	4	0.513225057065	False	None
chr9	-1	5831777	TGGCTTAGGGACCCGTGAGATAG	4	0.497494020833	False	None
chr7	-1	50413964	TGTTCAAGGGACCCGAGAGAAGG	3	0.312480125196	False	None
chr2	1	26985173	TGCCCAAGTGACACTGAGAGAG	3	0.304071290141	False	None
chr21	-1	39446141	TGTCCAGAGACCCATGAGATAG	3	0.271613746032	False	None
chrX	-1	141160847	TGAACATGTGACCCATGAGAAAG	4	0.27129321962	False	None
chr1	-1	81652697	TGACCAAGGGACCCAGTGAGATAG	3	0.259592789346	False	None
chr14	1	28360257	AGTTCAAGAGACCCCTGAGATAG	4	0.254738973799	False	None
chr14	1	40557725	TCTCAGAGTGACCCATGAGAGAG	4	0.247907188841	False	None
chr3	-1	190155920	TGTCCAAGTGTACCGTGAACAAG	3	0.245561551948	False	None
chr11	1	33866282	TGCCTAAGGGACCCATGAGATGG	4	0.242694227215	False	None
chr10	1	44089095	TGTCCTAGTCACCAAGTGAAGAAG	3	0.221863200422	False	None
chr22	1	30460143	TGGCCCAGTTACCCATGAGAGGG	4	0.221326410598	False	None
chr12	-1	72754425	TGGCAAAGTGTCCCTGAGAAAAG	4	0.220450566456	False	None
chr5	1	108551197	TGTCCAAGGCACCAAGTGAGACAG	3	0.208247249804	False	None
chr8	1	135352740	TGGCCTTGTGACCCGTGACACAG	4	0.20688363887	False	None
chr13	1	40275095	TATCAAAGTCACCTGTGAGACAG	4	0.206277768987	False	None
chr5	-1	100621234	AGTCGAAGTGAGCCATGAGAGGG	4	0.205125327511	False	None
chr15	-1	28117586	TCTGCAAGTGAACCCTGAGATGG	4	0.201603862661	False	None
chr11	1	90410318	TGTCAAAGTGAGCAGTGAGATAG	3	0.200988744589	False	None
chr6	1	143649208	TGACCAACTGACTCGTGACATGG	4	0.193758620192	False	None
chr7	-1	33722355	TCTCCAAGTTATCCCTGAGAGAG	4	0.185677157511	False	None
chr2	-1	114700881	TGAGCAAGTGACCTTGAGGTAG	4	0.180901806221	False	None
chr6	1	93489630	TGTGCAAGTCAGCCTTGAGATAG	4	0.179513600415	False	None
chr10	-1	85598737	GGTCCAAGTGCCCCGTGGAAGG	4	0.178584677419	False	None
chr3	-1	192269924	TCTCCATGAGACCCTTGAGAGGG	4	0.170999822639	False	None
chr17	1	80698201	AGTACAAGGGACCCGAGAGAAGG	4	0.166395666667	False	None
chr3	-1	176783898	TGGCCAAGTGACAAGTGAGAGGG	3	0.164422928767	False	None
chr12	-1	84274867	ATTCCAAGTGACCCGTGGAAGG	4	0.158312903226	False	None
chr2	1	59231969	TGTTCAATGTGTCATGAGAAAG	4	0.150171082469	False	None
chr11	-1	78182855	TCTCCAAGTGCCCGTGAGCAAG	4	0.149801193222	False	None
chr5	-1	76276236	TGTCCATGAGCCCCGTGAGCTAG	4	0.147669193853	False	None
chr17	1	5623458	TGTCCAGCTGTCCATGAGATGG	4	0.143048343379	False	None
chr2	1	96151255	TCTCCGAGTGGCCCATGAGACAG	4	0.137588489807	False	None
chr4	1	171024333	TCACCCAGTGACCTGTGAGACAG	4	0.133605730222	False	None
chr11	1	57346558	TGGCCAGGTGATCCATGAGAGGG	4	0.133476264053	False	None
chr10	-1	69672729	TGGACAAGAGACCAAGTGAAGAAG	4	0.132691231276	False	None
chr19	-1	38614234	TCTGCAAGTGACCCGAGAGCGGG	4	0.119960915493	False	None
chr1	-1	113026614	TGTTCAAGGGAACCTTGAGAGAG	4	0.119090998755	False	None
chr12	1	62618471	TGTCAGAGTTACCTGTGAGAGAG	4	0.11878368637	False	None
chr2	1	99413592	TGCCAGGTGCCCCGTGGGACAG	4	0.11598959393	False	None
chr3	-1	186919732	TGCCAGGTGCCCCGTGGGATGG	4	0.11598959393	False	None
chr7	-1	7101178	TGTCGATGGGACCCGAGAGAAG	4	0.106103128942	False	None
chr17	1	79556774	GATCCAAGTGACCTGTGAGGAGG	4	0.105908522727	False	None
chr17	-1	14426322	TGACCCAGTGACCCCTAAGAAGG	4	0.0957518566539	False	None
chr2	1	13694554	TGACCAAGTGAGCCCTGAGTAAG	4	0.0890036886608	False	None
chr8	1	93974695	TGTTCAAATGACCAAGTAAGATAG	4	0.087709361588	False	None
chr2	-1	144946470	TGTTCAAAGTGCCACCTGAGAAGG	4	0.0850896177386	False	None
chr14	-1	40910614	TGAGCAAGTGACTGGTGAAGAAG	4	0.0840450188278	False	None
chr12	1	122755017	TGTCCTAGAGGCCATGAGACAG	4	0.0786646994955	False	None
chr8	-1	142048980	CTTCCAAGTGACCTGTGAAAGAG	4	0.0785004401408	False	None
chr17	1	21466699	TGTCCAATGTGCCGTGTAAGG	4	0.0778221	False	None
chr7	-1	87112734	TGTACAAGTGAACCATAAGAAAG	4	0.0776174871245	False	None

## sgRNA#22

chr	strand	position	sequence	n_mismatches	score	ontarget	gene
chr5	-1	165008778	GTTGAGGCGTCATCCTACGGTGG	0	100.0	True	None
chrX	1	125101712	CTCCAGGTGTCATCCTACGGAAG	4	1.36678015564	False	None
chr2	1	122003201	GGACAGGCCTCATCCTACGGTGG	4	0.835102675097	False	None
chr9	-1	1466211	GTTGAGGAGGCATCCTAGGGGGG	3	0.508119111111	False	None
chr10	1	97933776	GTTAATCCCTCATCCTACGGTAG	4	0.339411292571	False	None
chr2	1	52750480	TTTGAGGAGTAATCCTACTGTAG	4	0.29240096831	False	None
chr6	-1	15712541	GGTGGGGCTTCATCCCACGGCAG	4	0.16348919214	False	None
chr4	1	29638035	GTAGAGGGGTCTTCCTATGGAAG	4	0.150546424	False	None
chr21	-1	21916338	GTCGAGGCTTCATCCTGCTGAAG	4	0.117774700312	False	None
chr2	-1	74701545	GTTGTGGCCTCATCCTCCAGGGG	4	0.115274138237	False	None
chr6	-1	152715325	GTTGGGGCCTCATCCTCCAGAGG	4	0.115274138237	False	None
chr17	1	8277361	CTTCAGGCGTCAGCCCACGGCAG	4	0.105393	False	None
chr6	1	43203012	GTTGAGGGGGCATCCTGGGGTGG	4	0.101056725	False	None
chr2	1	134658087	GTGGAGACGTCATCCTATGCCAG	4	0.090361802221	False	None
chr7	-1	128130922	G TTCAGGCTTCATCCTATGAAGG	4	0.0805000102941	False	None
chr16	-1	16760797	GTTGAGGAGTCATCCTTCAATGG	4	0.0760175494462	False	None
chrX	-1	122944220	GTTGGGGCTTCATCCTCGGGGAG	4	0.0704947805794	False	None
chr2	-1	20880826	GTTGAGGGGTATTCCAACGGTAG	4	0.0661334051383	False	None
chr21	-1	42105934	TTTGAGCCGTCATCCAACAGCAG	4	0.0618920651408	False	None
chr4	-1	148813254	CCTGAGGCGTCATCCATGGAGG	4	0.0553451612903	False	None
chr21	1	22503487	GTTGAGGCTACAGCCCACGGAAG	4	0.0519232660065	False	None
chr8	1	2169213	GTTGAGGGTTCATCATAAGGCAG	4	0.04666818	False	None
chr4	1	36332632	GTAGAGGCGTCAGCCTGGGGCGG	4	0.04559078139	False	None
chr2	-1	101614369	GTTGAGGCTTCAACCAACAGGAG	4	0.0186285891536	False	None
chr2	1	23646006	GTTGAGGCCTCACCCCATGGGGG	4	0.0114049193747	False	None

**Annexe 4: Gènes les plus différentiellement exprimés  
dans les HIOs RFX6<sup>-/-</sup>**



**Annexe 5: Gènes différentiellement exprimés  
communs aux 3 modèles: Rfx6<sup>-/-</sup> E18, Rfx6<sup>ΔAdInt</sup> et  
HIOs RFX6<sup>-/-</sup>**

Nom Gène Homme	Nom gène Souris	Rfx6 <sup>ΔAdInt</sup>	Rfx6 <sup>ΔAdInt</sup>	Rfx6 <sup>-/-</sup> E18	Rfx6 <sup>-/-</sup> E18	HIOs RFX6 <sup>-/-</sup>	HIOs RFX6 <sup>-/-</sup>
		vs WT (LFC)	vs WT (adj. P-value)	vs WT (LFC)	vs WT (adj. P-value)	vs CTRL (LFC)	vs CTRL (adj. P-value)
SCT	Sct	-0,90	1,78E-05	-1,46	2,51E-03	-1,79	4,16E-02
SNX10	Snx10	-0,39	2,15E-03	0,37	1,76E-02	-1,44	2,75E-08
PRPF6	Prpf6	-0,15	4,58E-02	-0,61	1,70E-03	0,46	6,54E-04
MAPRE1	Mapre1	-0,12	1,56E-02	-0,14	2,16E-02	0,72	5,89E-04
TSPAN12	Tspan12	-0,51	1,31E-02	0,29	2,31E-02	-0,99	4,49E-07
CHKA	Chka	0,45	3,71E-05	-0,38	2,49E-02	0,82	1,60E-02
UHRF1BP1L	Uhrf1bp1l	0,22	1,13E-02	0,28	2,19E-02	-0,91	3,74E-10
AIM1	Aim1	-0,35	3,49E-03	1,00	4,96E-02	-1,68	3,38E-04
ADGRG6	Gpr126	-0,79	7,34E-05	0,94	1,92E-02	-1,33	1,26E-08
BCHE	Bche	0,71	2,89E-03	-0,60	2,41E-02	-0,82	3,02E-02
RGS2	Rgs2	0,40	7,48E-03	0,33	8,43E-03	0,69	1,25E-04
AGMAT	Agmat	0,39	2,81E-02	-0,43	3,98E-02	-1,11	2,34E-02
LPGAT1	Lpgat1	0,44	1,61E-08	-0,66	9,87E-03	-0,68	1,29E-02
SLC12A5	Slc12a5	1,38	1,31E-06	1,26	3,07E-02	1,29	4,12E-03
ID1	Id1	-0,49	1,56E-03	0,67	3,47E-02	-1,03	6,99E-03
TSPAN8	Tspan8	-0,22	1,55E-04	0,49	3,60E-02	-0,99	1,88E-02
UBE4B	Ube4b	0,33	7,12E-04	-0,64	3,17E-03	-0,99	4,24E-08
REG4	Reg4	1,48	6,12E-06	0,36	2,46E-02	-1,69	2,69E-02
LPIN1	Lpin1	0,33	3,41E-02	-0,53	3,93E-02	0,45	1,14E-02
NT5E	Nt5e	1,80	1,71E-11	1,11	1,76E-02	-0,79	1,95E-02
FAM13A	Fam13a	0,47	3,65E-02	-0,87	1,86E-03	-2,53	8,16E-07
MYOM3	Myom3	-0,68	3,90E-03	0,42	2,20E-02	-2,10	1,38E-03
SLC17A4	Slc17a4	0,95	2,08E-04	0,48	1,70E-02	-3,07	3,40E-06
PRSS23	Prss23	0,59	3,43E-02	0,35	1,57E-02	0,60	1,68E-02
ME3	Me3	0,75	1,32E-03	0,76	2,91E-02	-0,63	4,39E-02
PTPRR	Ptprr	-0,49	4,56E-02	0,79	2,68E-02	-1,30	1,96E-03
SCN3A	Scn3a	0,99	2,97E-18	0,36	4,90E-02	0,89	4,92E-02
DBI	Dbi	0,36	2,98E-04	-0,56	2,51E-03	-0,46	2,96E-02
GHRL	Ghrl	-3,33	6,68E-23	-6,24	2,62E-04	-2,37	4,16E-03
SLC25A34	Slc25a34	0,91	1,37E-04	-0,42	2,56E-02	1,38	1,08E-02
LMX1A	Lmx1a	1,60	4,10E-11	0,80	1,50E-02	2,96	2,73E-04
SERINC5	Serinc5	0,31	1,56E-02	0,52	1,83E-02	-1,79	6,42E-07
PACSIN3	Pacsin3	0,48	1,02E-02	0,74	3,37E-02	-0,48	2,02E-02
TMEM135	Tmem135	0,23	7,00E-03	-0,34	3,72E-02	-0,90	8,80E-07
ACAA2	Acaa2	0,67	6,84E-05	-0,55	1,76E-02	-0,54	8,00E-03
PXDC1	1300014I06Ri	0,85	4,47E-07	-0,66	1,37E-02	1,21	4,80E-06
SLC38A11	Slc38a11	1,73	7,59E-08	0,72	3,58E-02	-1,27	5,95E-03
CYCS	Cycs	-0,25	2,56E-03	-0,51	2,66E-02	0,94	4,30E-03
MBOAT1	Mboat1	-0,79	1,74E-04	0,95	2,29E-02	-1,58	3,40E-06
TMEM139	Tmem139	0,65	9,19E-04	-0,52	4,22E-02	-1,56	5,87E-03
DNAH2	Dnahc2	0,79	7,84E-04	-0,67	3,55E-02	2,15	9,00E-04
GCNT1	Gcnt1	-1,30	1,20E-04	0,50	2,51E-02	-1,43	8,96E-08
RYR1	Ryr1	-0,89	1,70E-03	-0,91	2,87E-02	1,34	1,41E-02
SEPP1	Sepp1	0,71	1,89E-02	0,41	1,39E-02	-1,05	9,60E-03
TXNIP	Txnip	0,41	4,97E-02	0,62	4,48E-02	-2,28	2,15E-06



## Etude de la régulation transcriptionnelle de la différenciation des cellules entéroendocrines dans un modèle d'organoïde intestinal humain

### Résumé

Les cellules entéroendocrines sécrétrices d'hormones représentent 1% de l'épithélium intestinal mais sont des régulateurs essentiels du métabolisme énergétique et l'altération de leur différenciation provoque de graves pathologies métaboliques. Leur différenciation est régie par une cascade de régulations transcriptionnelles qui est encore peu décrite, particulièrement chez l'homme. L'objectif de ce projet de thèse était d'évaluer l'implication de plusieurs facteurs de transcription préalablement identifiés chez la souris (NGN3, RFX6, ARX, PAX4) dans la différenciation entéroendocrine humaine. Pour ce faire, ces gènes ont été inactivés par la technique CRISPR/Cas9 dans des cellules souches pluripotentes induites humaines (hiPSCs), qui ont ensuite été différenciées *in vitro* en organoïdes intestinaux (HIOs). Les analyses des HIOs déficients pour *NGN3* ne permettent pas de conclure quant à un rôle dans la différenciation entéroendocrine mais indiquent une altération de la régionalisation du tissu formé chez les mutants. RFX6 semble important pour la différenciation et/ou la fonction des cellules entéroendocrines, bien que sa fonction précise n'ait pas pu être déterminée.

**Mots-clés** : Différenciation entéroendocrine, hormones, hiPSCs, organoïdes intestinaux, facteurs de transcription

### Summary

Hormone-producing enteroendocrine cells represent 1% of the intestinal epithelium but are key regulators of the energetic metabolism and alteration of their differentiation is associated with severe metabolic disorders. Enteroendocrine differentiation is governed by a transcriptional regulatory cascade that is poorly described, especially in humans. This thesis project aimed to evaluate the implication of several transcription factors, previously identified in mice (NGN3, RFX6, ARX, PAX4), in human enteroendocrine differentiation. To do so, these genes were disrupted with the CRISPR/Cas9 system in human inducible pluripotent stem cells (hiPSCs), which were then differentiated in intestinal organoids (HIOs). Preliminary analysis of NGN3-deficient HIOs did not allow a firm conclusion regarding NGN3 implication in enteroendocrine differentiation but showed a tissue regionalization alteration. RFX6 seems important for the differentiation/function of enteroendocrine cells, although its precise function is still to be determined.

**Keywords** : Enteroendocrine differentiation, hormones, hiPSCs, intestinal organoids, transcription factors