

THESE

Présentée par
Josselin Soyer

Pour l'obtention du grade de
Docteur en Sciences de l'Université de Strasbourg
Discipline : Science de la vie
Spécialité : Biologie du développement

Identification des gènes cibles du facteur de transcription Neurogenin 3 et étude de leur fonction au cours du développement pancréatique

Soutenue publiquement le 1^{er} octobre 2009

Directeur de thèse :	Georg Mellitzer, CR INSERM, Strasbourg
Rapporteur externe :	Bénédicte Durand, Professeur, Lyon Harry Heimberg, Principal Investigator, Bruxelles
Rapporteur interne :	Stéphane Viville, Professeur, Strasbourg
Membre invité :	Gérard Gradwohl, DR INSERM, Strasbourg

A celles et ceux qui m'ont emmené ici aujourd'hui

REMERCIEMENTS

Je tiens tout d'abord à remercier Georg Mellitzer pour m'avoir encadré pendant ces 5 dernières années, pour le temps que tu y as accordé, les discussions et les conseils que tu m'as donnés.

Je remercie également Gérard Gradwohl, pour m'avoir accueilli dans son équipe.

Je remercie Bénédicte Durand, Michèle Kedingler, Harry Heimberg et Stéphane Viville d'avoir accepté de juger ce travail.

Je remercie également mes financeurs, le ministère de l'Enseignement Supérieur et de la Recherche, et l'Association de Recherche Contre le cancer pour leur soutien au cours de ces quatre années.

Vivi, la première alsacienne avec qui j'ai travaillé, merci pour avoir partagé cette folie commune du stock et notre petite pièce.

Merce, catalane avant tout, merci pour les conseils que tu m'as apportés, pour l'ambiance que tu as développée.

Isabelle, transfuge partiel qui nous a quittés, merci d'avoir partagé notre bureau pendant toutes ces années, et pour tes conseils.

Mélanie, Nathalie, duo inséparable, merci pour votre dynamisme, vos conseils sur les transfections.

Olivier, toi le petit, mon maître en qPCR, merci pour cette initiation.

Marjorie, ma prédécesseur chez GG, merci pour m'avoir fait goûter à l'Addal.

Christophe, l'homme de la Bio. Mol., merci pour tous tes conseils en WB, et la synthèse des anticorps.

Aline, ce manuscrit aurait été un champ de mines sans tes conseils, merci pour tes précieuses relectures.

Anthony, collègue de doctorat, merci pour tes infos sorties de nulle part.

Yaelle, merci d'avoir égayé notre labo pendant les mois que tu as passés parmi nous.

Julie, notre nouvelle doctorante, toi qui prendras ma succession, profite de ton doctorat, c'est 3 ans de bonheur.

Martine, notre dernière arrivée, bienvenue parmi nous et merci d'avoir tenté de décoder la séquence cachée.

Carole, merci pour tes conseils sur les ShRNA.

L'U682, toi qui m'as accueilli en DEA, bien que nous t'ayons quitté il y a presque deux ans maintenant, tu resteras un endroit merveilleux. Pour cela, je remercie notre directrice Michèle Kedinger, qui a su conserver cet esprit que nous partagions et qui rendait l'U682 comme un second chez soi.

Je tiens également à remercier tous les membres qui la compose, et notamment Oliver le grand et Annick pour leurs conseils sur les Northern Blot, Elisabeth pour ses conseils en culture cellulaire, Christiane pour ses conseils en histologie, Isabelle et Erwan pour leurs pauses cafés inoubliables.

Je ne peux écrire ces remerciements sans citer Philippe Ravassard et Corinne Pardanaud pour leur initiation à la dissection des tractus pancréatiques, et l'ensemble des services communs de l'IGBMC, toujours disponibles, et qui nous aident au quotidien, et notamment le service de culture cellulaire et la plateforme d'imagerie.

Il me reste finalement plusieurs personnes pour qui des remerciements ne suffiront jamais à vous dire ce que je pense. Je n'ai alors d'autre solution que de vous dédier ma thèse. Cette thèse est ainsi pour mes parents, mon adoratrice des tournesols, mon ex Montignienne, mes bretons préférés...

SOMMAIRE

TABLE DES FIGURES.....	9
TABLE DES TABLES.....	11
INTRODUCTION	12
A) L'organe, le pancréas	13
1. Anatomie du pancréas	13
2. Histologie du pancréas	13
3. Rôles physiologiques du pancréas	15
3.A Rôle du pancréas exocrine	15
3.B Rôle du pancréas endocrine.....	16
4. Maladies du pancréas.....	18
4.A Les diabètes.....	18
4.B Les pancréatites	20
4.C Les tumeurs pancréatiques	20
B) Organogénèse du pancréas	21
1. Régionalisation de l'endoderme	21
2. L'engagement dans la différenciation pancréatique	22
2.A La notochorde	23
2.B La signalisation du TGF β	24
2.C La signalisation du FGF	25
2.D La signalisation Hedgehog	26
2.E La signalisation de l'acide rétinoïque	27
2.F La signalisation Wnt.....	29
2.G Les facteurs de transcription impliqués dans la différenciation précoce du pancréas	31
2.G.a HLXB9	31
2.G.b ISL1.....	31
2.G.c HEX	32
2.G.d PDX1.....	32
2.G.e PTF1a.....	33
2.G.f SOX9.....	34
C) Contrôle génétique de la différenciation du pancréas endocrine.....	35
1. Induction de la différenciation des différents lignages cellulaires pancréatiques	35
2. Neurog3 est nécessaire à l'engagement de la différenciation du pancréas endocrine	38
3. les facteurs de transcription impliqués dans la différenciation des progéniteurs pancréatiques endocrines.....	41
3.A Les facteurs de transcription pancréatiques pan-endocriniens	41
3.A.a NEUROD1	41
3.A.b PAX 6	42
3.A.c INSM1	43
3.A.d ISL1	43
3.B Les facteurs de transcription impliqués dans le choix des sous types endocriniens	44
3.B.a PAX4	44
3.B.b ARX	44

3.B.b Arx et Pax4	45
3.C Les facteurs de transcription impliqués dans la différenciation terminale des cellules pancréatiques endocrines	45
3.C.a MAFA et MAFB	46
3.C.b HLXB9	46
3.C.c PDX1	47
3.C.d NKX6.1 et NKX6.2	47
3.C.e NKX2.2	48
3.C.f BRN4.....	48
D) La famille des gènes à domaine SERTA.....	49
E) La famille des gènes Rfx	51
1. <i>Rfx1</i>	52
2. <i>Rfx2</i>	53
3. <i>Rfx3</i>	53
4. <i>Rfx4</i>	54
5. <i>Rfx5</i>	55
6. <i>Rfx6</i> et <i>Rfx7</i>	55
OBJECTIFS	57
RESULTATS.....	60
A) Caractérisation de la souris transgénique Pdx1-Neurog3ER TM	61
1. La lignée de souris transgénique Pdx1-Neurog3ER TM	61
2. Expression ectopique de <i>Neurog3</i> dans les cellules épithéliales exprimant PDX1 à E12,5.....	62
3. Le transgène Pdx1-Neurog3ER TM restaure les lignages α et β dans le mutant <i>Neurog3</i>	62
4. L'activation de Neurog3ER TM conduit à l'induction des gènes impliqués dans la différenciation des cellules endocrines.....	63
B) Identification des potentiels gènes cibles de Neurogenin3	64
C) Etude du gène RIKEN cDNA 1700086L19.....	67
1. Caractérisation de la séquence de l'ARNm du gène <i>Riken L19</i>	67
2. Expression du gène <i>Riken L19</i>	68
D) Etude du gène Regulatory Factor X 6	70
1. Caractérisation de la séquence de l'ARNm du gène <i>Rfx6</i>	70
2. Expression du gène <i>rfx6</i>	72
2.A Cinétique d'expression de <i>rfx6</i> au cours du développement pancréatique	72
2.B <i>Rfx6</i> est spécifiquement exprimé dans le lignage endocrine du pancréas.....	72
2.C Expression précoce de RFX6 dans l'endoderme intestinal primitif et l'épithélium pancréatique	73
2.D RFX6 est exprimé dans l'ensemble des sous types cellulaires insulaires	74
2.E Les cellules <i>Rfx6</i> ne prolifèrent pas.....	75
2.F Les cellules <i>Rfx6</i> dans les îlots de Langerhans sont ciliées.....	75
2.G <i>Rfx6</i> est exprimé dans l'estomac et l'intestin.....	76
3. Caractérisation de la fonction de <i>rfx6</i> au cours du développement pancréatique murin	78
3.A Caractérisation des ShRNA <i>Rfx6</i>	78
3.B Etude de la perte de fonction de <i>Rfx6</i> sur la différenciation endocrine	79

3.C Le ShRfx6 n'induit pas de dégradation de la forme endogène de l'ARNm de <i>Rfx6</i>	80
4. Caractérisation de la fonction de <i>Rfx6</i> chez le poisson zèbre.....	81
DISCUSSION	83
A) Caractérisation du transcriptome pancréatique induit par NEUROG3 au cours de la différenciation du pancréas endocrine	84
B) Le gène RIKEN cDNA 1700086L19.....	86
C) Le gène <i>Rfx6</i>	87
1. La fonction précoce de <i>Rfx6</i> dans l'endoderme intestinal primitif.....	88
2. <i>Rfx6</i> dans la cascade génétique de différenciation du pancréas endocrine.....	88
3. La famille des gènes <i>Rfx</i> dans la différenciation endocrine.....	89
4. La fonction de <i>Rfx6</i> dans le pancréas adulte.....	90
5. La fonction de <i>Rfx6</i> dans la différenciation de l'intestin.....	91
6. La fonction de <i>Rfx6</i> dans la différenciation pancréatique du poisson zèbre.....	91
7. Implication de RFX6 dans la thérapie cellulaire pancréatique.....	92
D) Conclusion générale et perspectives	94
MATERIEL ET METHODES	100
A) Matériel biologique.....	101
1. Les lignées murines	101
2. Les lignées cellulaires	101
2.A. culture de la lignée HEK293t.....	101
2.B. culture de la lignée β TC3.....	101
B) Culture organotypique	102
1. Dissection et culture des épithélia pancréatiques.....	102
2. Infection des épithélia	102
C) Techniques d'histologie	103
D) Techniques d'immuno-histochimie	104
1. Démasquage.....	104
2. Blocage des sites aspécifiques et incubation de l'anticorps primaire	104
3. Anticorps secondaires et révélation.....	104
3.A. Immuno-histochimie à la peroxydase.....	104
3.B. Immunofluorescence	105
4. Acquisition des images	106
E) Analyse quantitative.....	106
1. Dans le pancréas embryonnaire.....	106
2. Dans les épithélia pancréatiques embryonnaires après la culture.....	106
F) Technique d'hybridation <i>in situ</i>	107
1. Synthèse de la sonde ARN.....	107
2. Hybridation <i>in situ</i>	107
G) Techniques de biologie moléculaire	108
1. Génotypage	108
1.A Préparation de l'ADN génomique	108
1.B Réaction de polymérisation en chaîne (PCR).....	109
2. Extraction d'ARN	110
3. Reverse-transcription (RT)/qPCR	110
3.A Transcription Reverse	110
3.B qPCR	111

4. Puce à ADN	112
5. RACE	112
6. Clonage.....	112
7. Northern Blot	113
8. Extraction de protéines	114
9. Western Blot.....	115
10. Clonage et test des ShRNA	115
10.A Clonage des ShRNA et de leur séquence cible.....	115
10.B Test de l'efficacité des ShRNA.....	118
10.C Production des ShRNA	118
H) Test Statistique	119
I) Analyse bioinformatique.....	119
REFERENCES	120
ABREVIATIONS.....	141
ANNEXE	145

TABLE DES FIGURES

Les figures 9 à 32 sont localisées au verso des pages indiquées.

Figure 1 : Organisation du pancréas adulte.....	Page 14
Figure 2 : Régulation de la glycémie.....	Page 17
Figure 3 : Développement de l'endoderme.....	Page 22
Figure 4 : Impact de la signalisation FGF, Hedgehog et TGF β sur la différenciation de l'endoderme en pancréas.....	Page 24
Figure 5 : L'induction de la voie Notch par <i>Neurog3</i> induit la différenciation des différents lignages pancréatiques.....	Page 36
Figure 6 : Modèle schématique de la différenciation du pancréas endocrine.....	Page 40
Figure 7 : Représentation des protéines à domaine SERTA humaines et de drosophile...Page	50
Figure 8 : Représentation des protéines humaines de la famille RFX.....	Page 52
Figure 9 : Caractérisation de l'expression de <i>Neurog3</i> contrôlée par les séquences régulatrices du gène <i>Pdx1</i> dans la lignée de souris transgénique <i>Pdx1-Neurog3ERTM</i>	Page 60
Figure 10 : Le transgène <i>Pdx1-Neurog3ERTM</i> restaure les lignages α et β dans le mutant <i>Neurog3</i>	Page 61
Figure 11 : L'activation de <i>NEUROG3ERTM</i> conduit à l'induction des gènes impliqués dans la différenciation des cellules endocrines.....	Page 62
Figure 12 : Filtres et critères de sélections des potentiels gènes cibles de Neurogenin3.....	Page 64
Figure 13 : Séquence putative du gène <i>Riken cDNA 1700086L19</i>	Page 66
Figure 14 : <i>Riken L19</i> est exprimé au cours du développement pancréatique embryonnaire et dans les îlots de Langerhans à l'exception des cellules β	Page 67
Figure 15 : <i>Riken L19</i> est spécifiquement exprimé dans les cellules du pancréas endocrine.....	Page 68
Figure 16 : Séquence nucléique et peptidique de <i>Rfx6</i>	Page 69
Figure 17 : Le transcrit de <i>Rfx6</i> dans le pancréas mesure environs 3487 nucléotides.....	Page 69
Figure 18 : Comparaison des protéines de la famille RFX.....	Page 70

Figure 19 : <i>Rfx6</i> est exprimé tout au long du développement pancréatique et dans les îlots de Langerhans.....	Page 71
Figure 20 : <i>Rfx6</i> est exprimé dans la majorité des cellules exprimant NEUROG3.....	Page 71
Figure 21 : <i>Rfx6</i> est spécifiquement exprimé dans les cellules du pancréas endocrine....	Page 72
Figure 22 : Expression précoce de RFX6.....	Page 73
Figure 23 : RFX6 est exprimé dans les différents sous types cellulaires pancréatiques endocriniens.....	Page 73
Figure 24 : Les cellules RFX6 ne prolifèrent pas pendant la différenciation des cellules du pancréas endocrine.....	Page 74
Figure 25 : Les cellules RFX6 dans les îlots de Langerhans sont ciliées.....	Page 74
Figure 26 : Expression de <i>Rfx6</i> dans l'intestin embryonnaire.....	Page 75
Figure 27 : RFX6 est exprimé dans l'épithélium intestinal.....	Page 76
Figure 28 : Le ShRNA <i>Rfx6</i> #2 induit une dégradation du transcrit <i>Rfx6-ires-GFP</i>	Page 77
Figure 29 : Le Sh <i>rfx6</i> induit une diminution de la différenciation endocrine.....	Page 78
Figure 30 : Aucune diminution du nombre de cellule à insuline ou à glucagon n'est induite par le Sh <i>Rfx6</i>	Page 79
Figure 31 : Le Sh <i>Rfx6</i> n'induit pas de dégradation de la forme endogène de l'ARNm de <i>Rfx6</i>	Page 80
Figure 32 : La perte de fonction de <i>zRfx6</i> induit une réduction de la différenciation du pancréas endocrine et une accumulation des progéniteurs pancréatiques endocrines...	Page 81
Figure 33 : Schéma des procédures de différenciation des cellules souches en cellules pancréatiques endocrines et expression protéique au cours de la différenciation pancréatique.....	Page 93
Figure 34 : Alignement des séquences nucléiques correspondant au domaine SERTA de <i>Riken L19</i> chez la souris, le gorille, le macaque et le cheval	Page 95
Figure 35 : Vecteurs utilisés pour le clonage et le test d'efficacité in vitro des ShRNA....	Page 117
Figure 36 : Vecteur utilisé pour la production des ShRNA, pTrip-CMV-eGFP-DeltaU3....	Page 119

TABLE DES TABLES

La table 1 est localisée au verso de la page indiquée.

Table 1 : Liste des potentiels gènes cibles de NEUROG3	Page 65
Table 2 : Temps d'incubations pour l'inclusion des tissus	Page 103
Table 3 : Anticorps primaires et secondaires	Page 105
Table 4 : Oligonucléotides utilisés et conditions de RT-PCR pour le géotypage	Page 109
Table 5 : Oligonucléotides utilisés pour les qPCR	Page 111
Table 6 : Oligonucléotides utilisés pour les clonages	Page 113
Table 7 : Oligonucléotides utilisés pour la synthèse des ShRNA	Page 116
Table 8 : Oligonucléotides utilisés pour le clonage des séquences cibles des ShRNA.....	Page 116

INTRODUCTION

A) L'organe, le pancréas

1. Anatomie du pancréas

Le pancréas présente une forme de cône presque horizontal, divisé en trois parties : la « tête », le « corps » et la « queue » du pancréas. Il est situé à l'extrémité supérieure de la région épigastrique, contre la paroi de l'abdomen, à l'arrière de l'estomac, entre la rate et le duodénum chez l'homme. Le pancréas est une glande mixte, abouchée au duodénum par le canal de Wirsung (Fig. 1)

2. Histologie du pancréas

Le pancréas est constitué de deux parties distinctes : le pancréas exocrine et le pancréas endocrine. Cette dernière constitue environ 2 % du pancréas mature (Githens, 1988). Chez les vertébrés supérieurs, le pancréas endocrine est disséminé dans le pancréas exocrine, pour ne former qu'une seule structure résultante : le pancréas. Chez les vertébrés inférieurs, ces deux parties sont soit séparées en 2 organes distincts (chez les Agnathes), soit le pancréas endocrine est regroupé en un seul gros îlot (appelé corps de Brockmann) localisé dans le pancréas exocrine (chez les téléostéens ; Biemar et al., 2001).

Le pancréas exocrine est constitué d'une part, de sacs glandulaires ou acini, et d'autre part de canaux excréteurs ou canaux pancréatiques (Fig. 1 ; Edlund, 2002). De plus, à la jonction des acini et des canaux pancréatiques se situent les cellules centroacinaires (Stanger et al., 2006). Les cellules acinaires pancréatiques, regroupées en acini, sécrètent différents zymogènes (ou proenzymes) et enzymes. Les cellules centroacinaires, comme les cellules canaliculaires sont à l'origine de la sécrétion de l'eau et des différents composés électrolytiques composant les sucs pancréatiques (Freedman et al., 1994 ; Garcia et al., 2008). Le suc pancréatique, sécrété dans les canaux pancréatiques intralobulaires va se déverser dans le canal de Wirsung, ainsi que la bile (le canal cholédoque y étant abouché au niveau de l'ampoule de Vater). Ce fluide composé du suc pancréatique et de la bile se déversera ensuite dans le duodénum au niveau de la papille de Vater (Fig. 1).

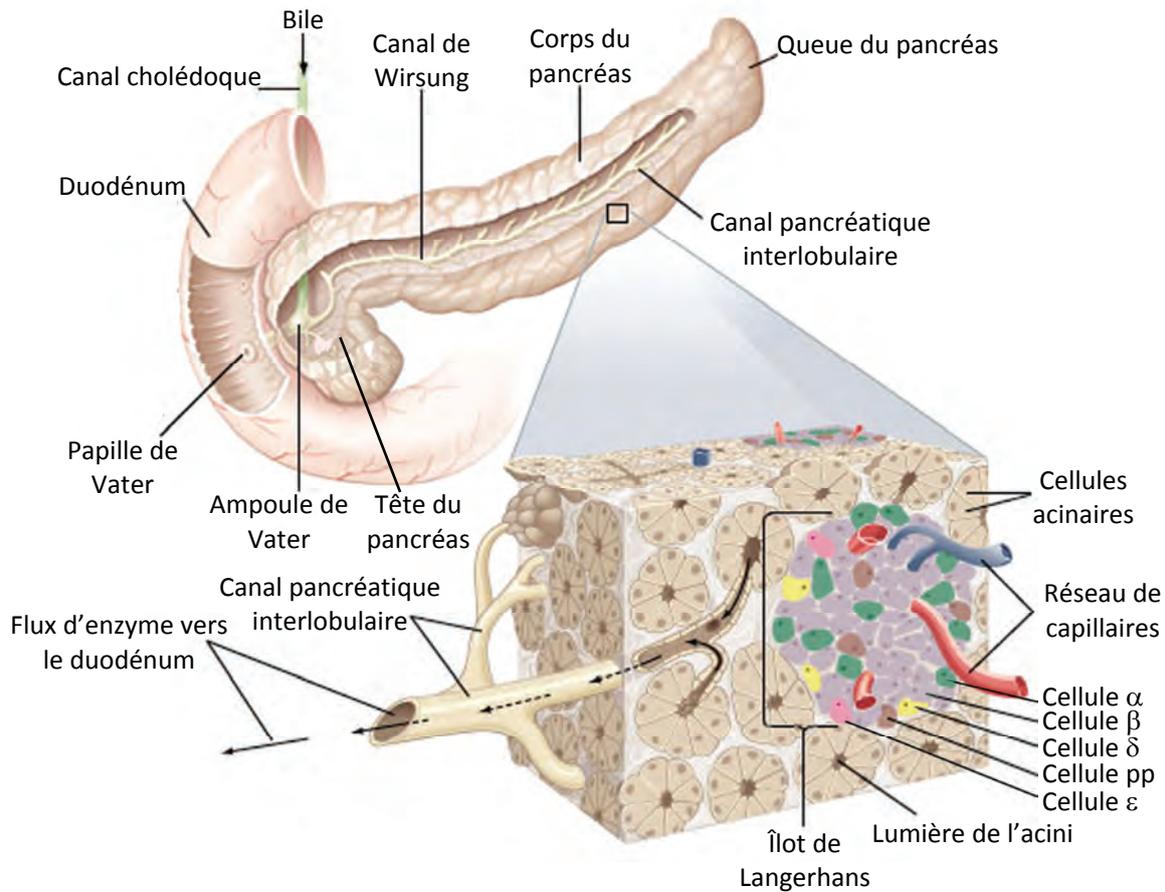


Figure 1 : Organisation du pancréas adulte. Le pancréas adulte, divisé en trois parties (la tête, le corps et la queue), est abouché au duodénum par le papille de Vater. Le pancréas est composé d'acini et canaux pancréatiques formant le pancréas exocrine, et d'îlots de Langerhans composant le pancréas endocrine. Les îlots de Langerhans sont constitués de cinq types cellulaires: les cellules α , β , δ , PP et ϵ sécrétant respectivement l'insuline, le glucagon, la somatostatine, le polypeptide pancréatique et la ghréline (adaptée de l'encyclopédie Britannica).

Le pancréas endocrine est quand à lui constitué par les îlots de Langerhans (Fig. 1). Ces agrégats cellulaires, de forme sphérique, sont principalement constitués de quatre types cellulaires endocriniens : les cellules α , les cellules β , les cellules δ et les cellules PP, sécrétant respectivement les hormones glucagon, insuline, somatostatine et polypeptide pancréatique (Edlund, 2002). Chez les rongeurs, les cellules sécrétant l'insuline sont localisées au centre des îlots, et les autres types cellulaires en périphérie. Chez l'homme, cette structure est conservée bien que la ségrégation des cellules sécrétant l'insuline par rapport aux autres types cellulaires endocriniens est moins marquée (Slack, 1995). Un cinquième sous type cellulaire a également été caractérisé : les cellules epsilon, sécrétant la

ghréline (Wierup et al., 2002). Toutefois, l'expression de la ghréline chez la souris semble restreinte au pancréas en développement (Heller et al., 2005). Chez l'homme, la ghréline est également détectée dans les îlots de Langerhans (Andralojc et al., 2009). Les îlots de Langerhans sont vascularisés afin de permettre le transport des hormones pancréatiques (Fig. 1).

3. Rôles physiologiques du pancréas

3.A Rôle du pancréas exocrine

Les rôles du pancréas exocrine sont dépendants de ses deux principaux constituants : l'eau et les composants électrolytiques d'une part, et les zymogènes et enzymes d'autre part.

Le suc pancréatique va permettre de neutraliser le chyme gastrique, riche en acide chloridrique, grâce à la présence d'ions bicarbonates, principal composant électrolytique du suc pancréatique (Meyer, 1977).

Les enzymes pancréatiques vont poursuivre la métabolisation du chyme gastrique afin de permettre l'absorption des glucides, lipides et composés protéiques. Pour cela, les zymogènes sécrétés par les cellules acinaires vont être activés dans la lumière intestinale. Pour cela, le trypsinogène produit par les cellules acinaires va être métabolisé en trypsine par l'entérokinase, un enzyme intestinale ou, la trypsine activée. La trypsine activera à son tour les autres zymogènes (le chymotrypsinogène en chymotrypsine, la proélastase en élastase, la procarboxypeptidase A en carboxypeptidase A, la procarboxypeptidase B en carboxypeptidase B et la prophospholipase A en phospholipase A). Les cellules acinaires sécrètent également des enzymes sous une forme active.

L'ensemble des zymogènes et enzymes pancréatiques peut être classé en 5 catégories : les enzymes lipolytiques, protéolytiques, glycolytiques, nucléolytique et les inhibiteurs pancréatiques. Parmi les enzymes lipolytiques, la lipase hydrolise les triglycérides en acides gras et glycérol ; et les phospholipases sont des hydrolases capables de couper les liaisons esters des phosphoglycérides. Parmi les enzymes protéolytiques, la trypsine, au-delà de son rôle d'activation des zymogènes, permet l'hydrolyse des protéines au niveau de la fonction carboxyle des acides aminés basiques. L'élastase est une endopeptidase qui coupe la liaison peptidique située après un acide aminé ayant une chaîne latérale aliphatique, non chargée et non aromatique. Parmi les enzymes glycolytiques, l' α -amylase catalyse l'hydrolyse des

liaisons (1→4) des polymères du glucose. Les enzymes nucléolytiques sont des endonucléases permettant la dégradation de l'ARN (la ribonucléase) ou de l'ADN (la désoxyribonucléase). Finalement, la dernière catégorie d'enzymes sécrétées par les cellules acinaires sont les inhibiteurs pancréatiques. Ces enzymes ont un rôle fondamental en empêchant l'autolyse du pancréas en inhibant la trypsine (inhibiteur de Kazal) ou les enzymes sécrétées sous forme active (Clemente et al., 1972 ; Figarella, 1973).

L'ensemble des sécrétions du pancréas exocrine sont notamment régulées par la sécrétine, et la cholecystokinine-pancréozymine, deux hormones produites respectivement par les cellules S (Chey et al., 1976) et I (Polak et al., 1975) du duodénum ; et la gastrine, hormone produite par les cellules G (Soll et al., 1984) de l'estomac. La sécrétine, et la cholecystokinine-pancréozymine induisent respectivement la sécrétion des composés hydroélectrolytiques et, des enzymes et zymogènes ; la gastrine induit un accroissement de l'ensemble des sécrétions du pancréas exocrine (Meyer, 1977). Le polypeptide pancréatique, sécrété par le pancréas endocrine, régule également les sécrétions du pancréas exocrine en réduisant les sécrétions (Beglinger et al., 1984).

3.B Rôle du pancréas endocrine

Le rôle principal du pancréas endocrine est la régulation de la glycémie. Pour cela, les sécrétions d'insuline et de glucagon vont influencer sur la néoglucogenèse et la glycogénolyse dans le foie et le muscle. L'insuline, sécrétée par les cellules β des îlots de Langerhans, est une hormone hypoglycémisante qui induit l'augmentation du stockage du glycogène dans le foie et inhibe la néoglucogenèse et la glycogénolyse hépatique (Michael et al., 2000 ; Bergman et al., 2000). En parallèle, la sécrétion d'insuline induit l'absorption du glucose par les muscles et les cellules adipocytaires (Fig. 2 ; Bergman et al., 2000). A l'inverse, le glucagon est une hormone hyperglycémisante dont le rôle est de maintenir un niveau suffisant de glucose dans le sang (0,70 à 1,10 g/L chez l'homme). Pour cela, la sécrétion du glucagon stimule la synthèse de glucose en induisant la glycogénolyse hépatique et la lipolyse dans les adipocytes, permettant ainsi de libérer le glycérol qui sera ensuite métabolisé en glucose (Fig. 2).

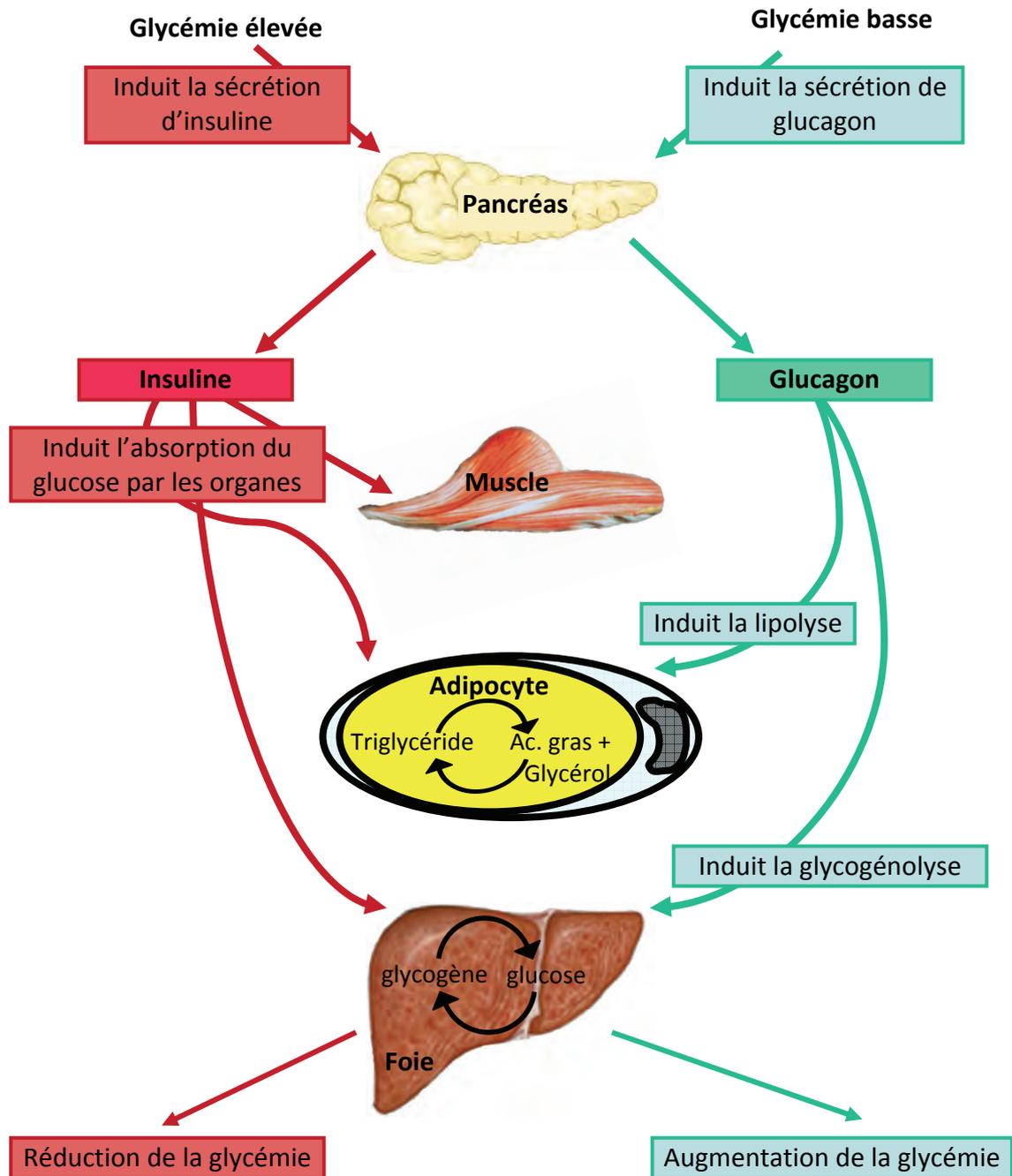


Figure 2 : Régulation de la glycémie. En cas d'excès de glucose dans le sang, l'insuline sécrétée par les cellules β , induit l'augmentation du stockage du glycogène dans le foie. En parallèle, la sécrétion de l'insuline induit l'absorption du glucose par les muscles et les cellules adipocytaires. A l'inverse, le glucagon maintient un niveau suffisant de glucose dans le sang (0,70 à 1,10 g/L chez l'homme). Pour cela, le glucagon stimule la synthèse de glucose en induisant la glycogénolyse hépatique, et la lipolyse dans les adipocytes permettant ainsi de libérer le glycérol qui sera ensuite métabolisé en glucose par le foie.

Les trois autres hormones sécrétées par les îlots de Langerhans : la somatostatine, le polypeptide pancréatique et la ghréline ne sont pas liées directement à la régulation de la glycémie, mais à l'absorption de nutriments. En effet, la somatostatine inhibe les sécrétions

acides de l'estomac directement ou via l'inhibition de la sécrétion de la ghréline et de la gastrine (Schubert, 2003 ; Strowski et al., 2003).

Ces sécrétions acides de l'estomac ont pour but de prédigérer les aliments absorbés en induisant notamment la rupture des ponts disulfures et des liaisons hydrogènes des protéines, afin de permettre la métabolisation des aliments dans l'intestin. Les sécrétions acides de l'estomac ont également une action antibactérienne (Schubert, 2003). Le polypeptide pancréatique inhibe les sécrétions exocrines pancréatiques, ainsi que les sécrétions acides gastriques et la motilité intestinale (Murphy et al., 2004). Finalement à l'inverse de la somatostatine et du polypeptide pancréatique, la ghréline qui est également sécrétée par les cellules X/A de l'estomac, est caractérisée pour stimuler la vidange gastrique et inhiber l'état de satiété (Kojima et al., 1999 ; Murphy et al., 2004).

4. Maladies du pancréas

Les maladies du pancréas sont de trois types : les diabètes, les cancers pancréatiques et les pancréatites.

4.A Les diabètes

Le diabète est une maladie très répandue dans le monde, qui induit une dérégulation du taux de glucose sanguin. Cette absence de régulation de la glycémie, ainsi que les injections d'insuline peuvent entraîner de nombreux effets secondaires, tels que les rétinopathies, néphropathies, neuropathies ou amputations. En l'an 2000, 170 millions de personnes étaient diabétiques, et ce chiffre pourrait atteindre plus de 350 millions en 2030 (Wild et al., 2004). Selon leur étiologie, les diabètes sont classés en trois types : le diabète de type 1, de type 2 et les MODY.

Le diabète de type 1, appelé également diabète insulino-dépendant est causé par une destruction spécifique des cellules β des îlots de Langerhans sécrétant l'insuline par une réaction auto-immune (Eizirik et al., 2009). Cette forme de diabète atteint l'enfant ou le jeune adulte, et représente environ 10 % des cas de diabète. L'absence complète d'insuline déclenche une production massive de glucose. Par conséquent, les diabétiques de type 1 doivent donc s'injecter de l'insuline afin de réguler leur glycémie, tout en suivant un régime

alimentaire afin de réguler au mieux leur apport en glucose. Aucun traitement curatif n'existe pour le diabète de type 1. Pour les personnes ne pouvant contrôler leur diabète par injections d'insuline ou en cas de complications telles que les néphropathies, une greffe d'îlot de Langerhans ou de pancréas à l'occasion de greffe rein-pancréas peut être envisagée (White et al., 2009). Cependant, le manque de greffons disponibles en restreint considérablement l'utilisation. Afin de pallier aux effets secondaires dus à l'injection quotidienne d'insuline chez les diabétiques et au manque de greffons de pancréas, le développement de cellules β à partir de cellules souches est aujourd'hui étudié. Nos connaissances sur la différenciation de l'endoderme et du pancréas endocrine permettent ainsi de développer *in vitro* à partir de cellules souches des cellules sécrétant l'insuline. Cependant ces cellules ne peuvent réguler leur sécrétion en fonction du taux de glucose (D'Amour et al., 2006). D'autres études seront donc encore nécessaires afin que les cellules β dérivées de cellules souches puissent être utilisées comme thérapie dans le cadre du diabète de type 1.

Le diabète de type 2 (diabète non insulino-dépendant) est souvent lié à une obésité ou une surcharge pondérale du patient. De plus, une prédisposition génétique est aussi liée au diabète de type 2. Le diabète de type 2 représente 90 % des cas de diabète, et touche principalement des personnes adultes de plus de 40 ans. Cependant, suite à l'apparition de surcharge pondérale de plus en plus tôt au cours de la vie, l'âge moyen d'apparition du diabète de type 2 est de plus en plus précoce (Wild et al., 2004). Au début de la maladie, la production d'insuline par le pancréas est normale. Cependant, les organes périphériques n'y répondent plus correctement. De plus, la prise de poids par les malades stimule les cellules des muscles afin d'utiliser de préférence les acides gras pour en tirer de l'énergie. Par conséquent, le glucose ne pouvant entrer dans les cellules, les cellules β du pancréas vont produire de plus en plus d'insuline pour forcer la prise de glucose par les cellules. Ainsi, au fur et à mesure de l'avancée de la maladie, les cellules β vont s'épuiser, jusqu'à la disparition de la sécrétion d'insuline (Doria et al., 2008). Les traitements actuels sont la définition d'un régime alimentaire, la pratique d'exercice physique ainsi que des médicaments stimulant la sécrétion de l'insuline ou diminuant la résistance à l'insuline. Malgré tout, après plusieurs années de médication, l'injection d'insuline peut devenir nécessaire afin de pallier au manque de sécrétion d'insuline par les cellules β (Buse et al., 2009).

Les MODY sont un sous type de diabète de type 2. Elles représentent 5 à 10 % des cas de diabète de type 2 (Ledermann, 1995). Les MODY sont dus à la mutation de facteurs de transcription ou de protéines, impliqués dans l'expression de l'insuline ou la détection du glucose par les cellules β . Six types de MODY existent. Les MODY-1, -2, -3, -4, -5 et -6 correspondent respectivement à la mutation de *Hnf4 α* , la glucokinase, *Hnf1 α* , *Pdx1*, *HnfF1 β* et *NeuroD1*.

4.B Les pancréatites

Les pancréatites comprennent deux formes différentes : les pancréatites aiguës et les pancréatites chroniques. La pancréatite aiguë est caractérisée par une inflammation du pancréas. Les deux causes principales des pancréatites aiguës sont un excès d'alcool et une obstruction du canal biliaire. Elles sont bénignes dans 80% des cas (Wang et al., 2009a). La pancréatite chronique est caractérisée par une calcification du pancréas. L'étiologie est chez 80 à 90% des patients un alcoolisme chronique. Il n'existe pas de traitement curatif, seul un régime alimentaire et un suivi médical permettent d'en limiter l'évolution (Dimagno et al., 2009).

4.C Les tumeurs pancréatiques

L'adénocarcinome est le cancer pancréatique le plus fréquent, avec plus de 80% des cas. Il est très rarement guérissable car son diagnostic a souvent lieu alors que d'autres organes adjacents sont atteints. Même chez les personnes pour qui une opération de résection est possible, le taux de survie à 5 ans est inférieur à 20 % (Lau et al., 2006). Les tumeurs neuroendocrines ont contrairement un pronostic beaucoup plus favorable. La majorité d'entre elles sont des insulinomes, dont le risque de malignité est exceptionnel. Les glucagonomes, plus rares, ont eux un risque de malignité plus élevé (Davies et al., 2009).

B) Organogénèse du pancréas

1. Régionalisation de l'endoderme

Le pancréas se développe à partir de l'endoderme. Ce dernier est obtenu après la gastrulation qui permet d'organiser l'embryon en trois feuillets : l'ectoderme d'où se développera la peau et le système nerveux central ; le mésoderme d'où se développera le système circulatoire, le squelette et les muscles ; et l'endoderme d'où se développera notamment le système digestif (Wells et al., 1999).

En fin de gastrulation, l'endoderme n'est encore qu'un feuillet de 500 à 1000 cellules. Cependant, bien que l'endoderme apparaisse morphologiquement homogène, on peut diviser ce feuillet en quatre régions (I à IV, Fig. 3a). Cette division est faite par l'expression de certains gènes tels que *Cerberus* (Bouwmeester et al., 1996), *Otx1* (Rhinn et al., 1998) et *Hex1* (Thomas et al., 1996) qui marquent spécifiquement la partie antérieure de l'endoderme. La partie postérieure est caractérisée par l'expression de gènes tels qu'*I-FABP* (Green et al., 1992) et *Cdx2* (Beck et al., 1995). La région I va ensuite se rabattre ventralement et progresser vers la partie postérieure de l'embryon. A l'inverse, après s'être rabattue ventralement, la région IV va progresser vers la partie antérieure de l'embryon. Ces mouvements morphogénétiques, en combinaison avec le retournement de l'embryon permettront la formation d'un tube intestinal primitif à E9 (Fig. 3a ; Wells et al., 1999).

La région I donnera naissance au sac vitellin, ainsi qu'au poumon, au foie, au pancréas ventral et à l'estomac. La région II contribuera à l'œsophage, à l'estomac, au pancréas ventral et au duodénum. Les régions III et IV formeront respectivement le jéjunum et l'iléon ; et le colon (Fig. 3 ; Wells et al., 1999). Wells a montré que le FGF4 est impliqué dans ce processus de détermination de l'identité antéropostérieure de l'endoderme (Wells et al., 2000). Des travaux récents ont confirmé ces résultats et ont montré que la formation d'un gradient de FGF4 influence la régionalisation de l'expression de gènes tels que *Hex1*, *Pdx1* et *CdxB*, marquant respectivement les territoires de l'endoderme antérieur ; de l'estomac postérieur, du pancréas et du duodénum ; et de l'endoderme postérieur (Dessimoz et al., 2006).

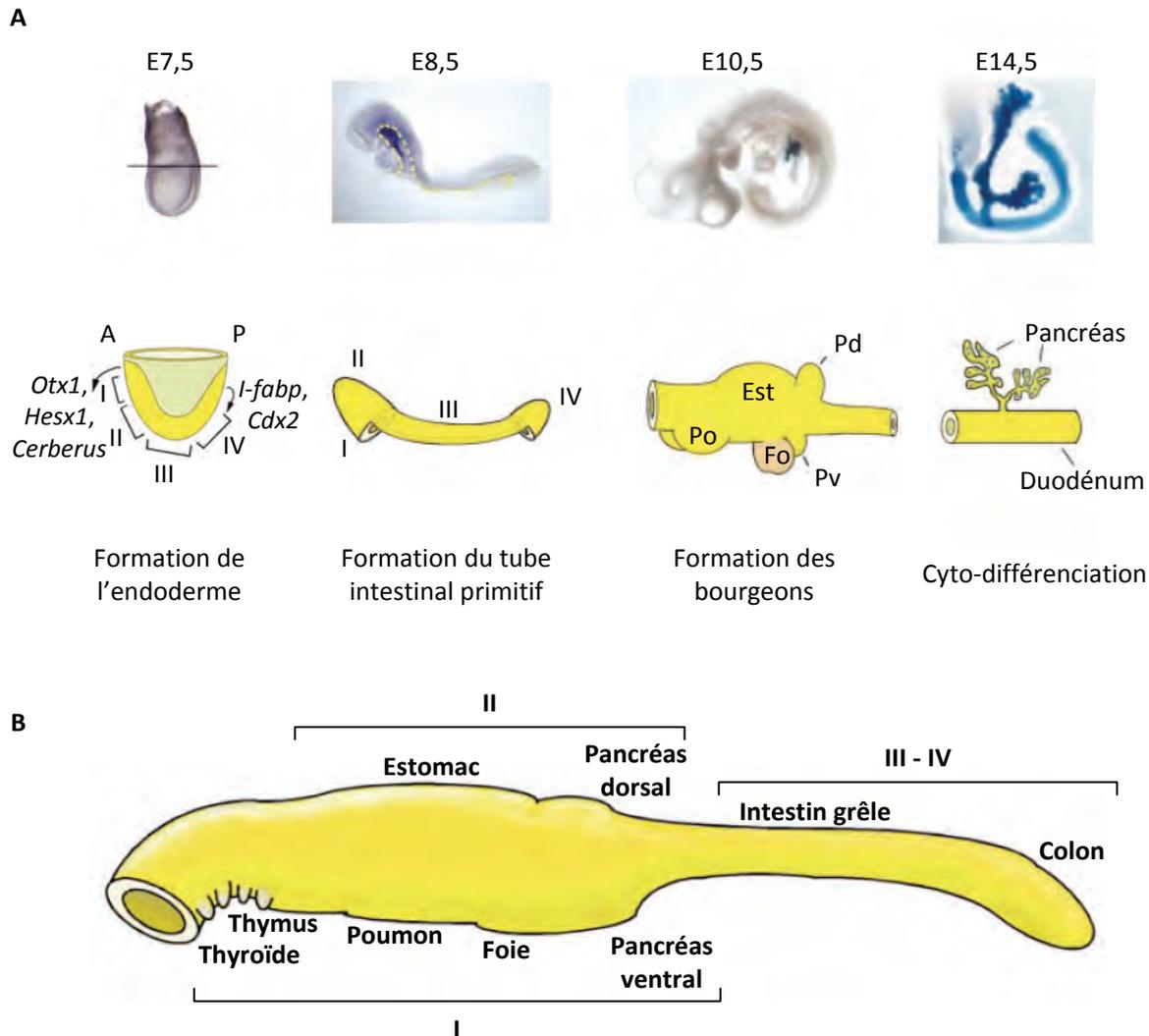


Figure 3 : Développement de l'endoderme. (A) Evolution de l'endoderme entre E7,5 et E14,5. A E7,5, l'endoderme issu de la gastrulation est formé d'un feuillet d'environ 500 cellules. Les gènes tels que *Otx1*, *Hesx1* et *Cerberus* vont délimiter la partie antérieure de l'endoderme ; alors que *I-fabp* et *Cdx2* vont en délimiter la partie postérieure. A ce stade, l'endoderme peut être divisé en quatre régions numérotées de I à IV. Un jour après la gastrulation, l'endoderme commence à former le tube intestinal primitif suite à des mouvements morphogénétiques, qui sera clos à E9,5. La ligne en pointillé jaune montre le tube intestinal primitif en formation, marqué par l'expression de l'albumine (bleu). A E10,5, l'endoderme va ensuite bourgeonner afin de donner naissance aux poumons, au foie et au pancréas qui se développeront ensuite en organes fonctionnels. L'expression de *Pdx1* (bleu) est révélée à E10,5 et E14,5. A: antérieur, P: postérieur, Pd: Pancréas dorsal, Po: Poumon, Pv: Pancréas ventral et Est: estomac (B) schéma représentant les différents territoires présomptifs de l'endoderme à E9,5. (adaptée de Wells et al., 1999).

2. L'engagement dans la différenciation pancréatique

Le premier signe visible de la formation du pancréas est l'expression de *Pdx1* à E8,5 (Ohlsson et al., 1993). Cependant, le pancréas n'est pas encore morphologiquement distinguable de l'endoderme à ce stade. Les premiers signes morphologiques de la formation

du pancréas sont marqués à E8,5-E9,5 par l'apparition d'un bourgeon dorsal et de deux bourgeons ventraux, précisément là où l'endoderme est en contact avec l'endothélium de l'aorte sur la face dorsale et les deux veines vitellines sur la face ventrale (Lammert et al., 2001). Chez les vertébrés supérieurs, seul le bourgeon ventral droit persistera, le second régressera en même temps que la veine vitelline gauche (Lammert et al., 2001). En revanche, chez les vertébrés inférieurs, les deux bourgeons antérieurs (correspondant aux bourgeons ventraux chez les vertébrés supérieurs) fusionneront et seront impliqués dans la formation du pancréas endocrine (Field et al., 2003).

Le développement précoce du pancréas est lié à un ensemble de morphogènes tels que la famille des FGF, des hedgehog et de la signalisation via l'acide rétinoïque.

La formation du pancréas ventral est faite par défaut, contrairement à celle du pancréas dorsal. En effet, l'endoderme ventral antérieur, qui n'est pas en contact avec le mésoderme cardiaque sera dirigé dans la différenciation pancréatique, et sera à l'origine du pancréas ventral (Deutsch et al., 2001). En revanche, l'endoderme ventral en contact avec le mésoderme cardiaque, percevra les molécules de FGF émises par le mésoderme cardiaque et se différenciera en foie.

2.A La notochorde

Le développement dorsal du pancréas est contrôlé par la notochorde. Chez la souris, la notochorde est en contact avec l'endoderme prépancréatique dorsal, jusqu'à la fusion des aortes dorsales qui se placeront entre l'épithélium intestinal primitif et la notochorde (Kim et al., 1997). Kim et al. ont montré que la suppression de la notochorde chez l'embryon de poulet empêche la formation correcte du pancréas dorsal, et inhibe l'expression de gènes spécifiques du pancréas tels que *Pdx1*, *Pax6* et *Isl1* (Kim et al., 1997). De plus, des expériences de greffe de notochorde sur de l'endoderme de poulet ont montré que la notochorde permet la suppression de Sonic hedgehog (shh) dans l'endoderme, cette suppression étant nécessaire à la formation du pancréas (Fig. 4 ; Hebrok et al., 1998). Hebrok et al. ont également montré que les facteurs sécrétés par la notochorde, responsables de la formation du pancréas sont l'activine β B et le FGF2.

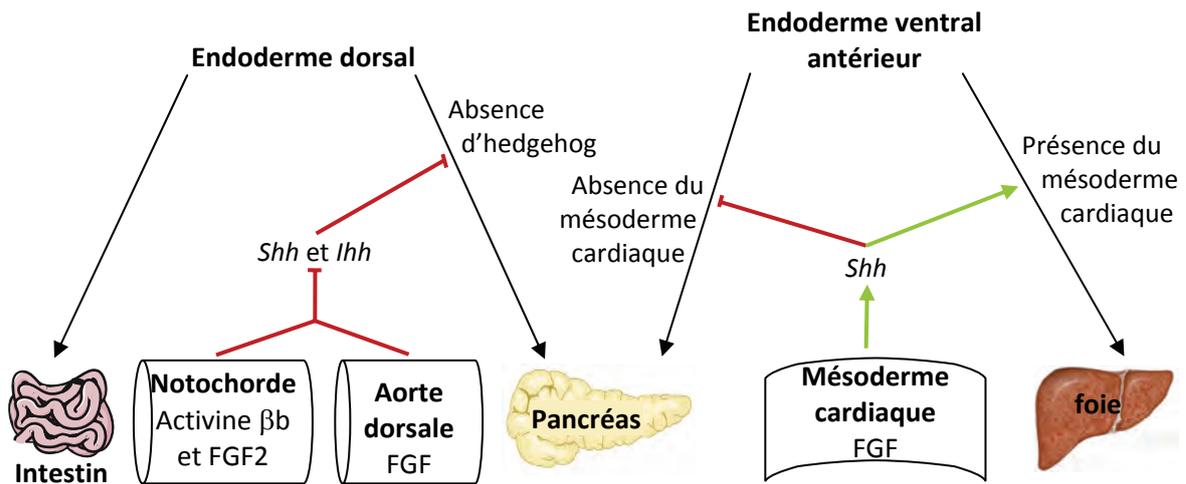


Figure 4 : Impact de la signalisation FGF, Hedgehog et TGF β sur la différenciation de l'endoderme en pancréas. L'endoderme ventral antérieur se différenciera par défaut en pancréas, et formera le bourgeon pancréatique ventral. L'engagement de la différenciation de mésoderme dorsal en pancréas se fera suite à l'inhibition de l'expression de *Shh* et *Ihh* par la notochorde et l'aorte dorsale (adaptée de Kim et al., 2002).

2.B La signalisation du TGF β

La signalisation du TGF- β (Balemans et al., 2002 ; von Bubnoff et al., 2001) a été impliqué dans le développement précoce du pancréas. La première étude montrant l'implication de la voie du TGF β dans la formation du pancréas est celle de Sanvito et al. qui montrent qu'un excès de TGF β induit une augmentation de la différenciation endocrine (Sanvito et al., 1994). L'étude inverse, dans laquelle l'activine est inhibée dans l'épithélium pancréatique montre une induction de la différenciation exocrine au dépend de la différenciation endocrine (Miralles et al., 1998). Ces résultats soulignent un effet tardif de la voie du TGF β sur la différenciation du pancréas en influençant la balance endocrine/exocrine. Hebrok et al. ont également mis en évidence que l'ajout d'activine β B sur des cultures d'endoderme de poulet permet d'induire l'expression de PDX1 et de l'insuline, mimant ainsi le rôle de la notochorde (Hebrok et al., 1998). Kim et al. ont également montré que la mutation du récepteur de l'activine de type IIB, chez des souris hétérozygotes mutantes pour le récepteur de l'activine de type IIA, induit la formation d'un pancréas hypoplasique enroulé autour de l'intestin (Kim et al., 2000). Ce phénotype est similaire à celui observé chez les souris mutantes pour Indian hedgehog quand elles sont croisées avec les souris mutantes pour *Shh*. En effet, on observe chez ces souris une

hypoplasie du pancréas, avec un enroulement de ce dernier autour du duodénum (Hebrok et al., 2000).

L'ensemble de ces résultats, montre que la régulation de l'expression des gènes de la famille Hedgehog par la voie de signalisation du TGF β est nécessaire à la fois à la différenciation précoce du pancréas, ainsi qu'à la régulation plus tardive (vers E12,5) de la balance de la différenciation des cellules endocrines/exocrines.

2.C La signalisation du FGF

Tout comme la voie de signalisation du TGF β , la signalisation du FGF (Bottcher et al., 2005) est impliquée dans différents processus de développement du pancréas.

Comme préalablement cité, le FGF4 est nécessaire très tôt au cours de la régionalisation de l'endoderme (Wells et al., 2000).

Puis au cours de la différenciation pancréatique, la signalisation du FGF a un impact opposé dans le pancréas ventral par rapport au pancréas dorsal. En effet, l'expression de FGF par le mésoderme cardiaque induit la différenciation de l'endoderme ventral antérieur en foie suite à l'induction de l'expression de Shh (Fig. 4 ; Deutsch et al., 2001).

En revanche, la signalisation du FGF est nécessaire à une différenciation correcte du pancréas dorsal. La perte de l'expression de FGF10 induit une forte réduction de la croissance de l'épithélium pancréatique, qui par conséquent induit l'arrêt de la croissance et de la morphogénèse des bourgeons pancréatiques (Bhushan et al., 2001). De plus, la mutation du récepteur FGFR-2 IIIB, qui est le récepteur de FGF10, entraîne un phénotype similaire à celui observé par Bhushan et al. En effet, Pulkkinen et al. ont montré que la perte de fonction du récepteur FGFR-2 IIIB induit une réduction de la croissance et de la ramification des bourgeons pancréatiques (Pulkkinen et al., 2003). Toutefois, à l'inverse du phénotype de la mutation de FGF10, une différenciation partielle du pancréas endocrine et exocrine est conservée chez les souris mutantes pour FGFR-2 IIIB. Sachant que FGFR-2 IIIB est le récepteur de FGF10 (Miralles et al., 1999), ces résultats suggèrent que via FGFR-2 IIIB, FGF10 aurait pour fonction d'induire la prolifération de l'épithélium pancréatique. Hebrok et al. ont également montré que la culture d'endoderme de poulet en présence d'une dose de 1 ng/ml de FGF2 permet d'induire l'expression de *Pdx1* et de l'insuline, tout en supprimant l'expression de Shh et *Ptc*, la cible de Shh (Hebrok et al., 1998).

La voie du FGF, au delà de son impact sur la régionalisation de l'endoderme et sur la formation des bourgeons pancréatiques, semble être impliquée dans la balance de la différenciation endocrine et exocrine. Dans ce sens, Hebrok et al. ont également montré qu'une dose élevée de FGF2 (10-50 ng/ml) induit une réduction de l'expression de *pdx1* et de l'insuline par rapport à une dose plus faible (Hebrok et al., 1998). De plus, l'expression ectopique de FGF4 selon le patron d'expression de *Pdx1* induit une destruction des cellules endocrines, avec une persistance des cellules exocrines (Dichmann et al., 2003). Un phénotype similaire a été obtenu par l'ajout prolongé de FGF7 sur des explants pancréatiques en culture, induisant la prolifération de l'épithélium tout en supprimant la différenciation endocrine. De plus, l'induction de la prolifération de l'épithélium par ajout de FGF7 est suivie de la différenciation de cet épithélium en cellules endocrines quand le FGF7 est enlevé du milieu de culture (Elghazi et al., 2002).

L'ensemble de ces résultats suggère donc que la voie de signalisation du FGF est impliquée au cours du développement du pancréas, dans la prolifération de l'épithélium pancréatique, tout en empêchant la différenciation de cet épithélium en cellules endocrines. Il faut aussi noter que la signalisation du FGF régule la signalisation Hedgehog.

2.D La signalisation Hedgehog

La voie de signalisation hedgehog (Lum et al., 2004) est également impliquée dans la formation du pancréas. Sonic Hedgehog (*Shh*) est exprimé tôt dans l'ensemble de l'épithélium intestinal primitif (Apelqvist et al., 1997). Puis, son expression va être spécifiquement inhibée dans le territoire présomptif du pancréas sous l'influence de la notochorde (Fig. 4 ; Hebrok et al., 1998). Toutefois, la délétion de *Shh* ne permet pas d'induire un agrandissement du territoire pancréatique ; bien que proportionnellement à la taille de l'embryon qui lui diminue chez les mutants *Shh*, la taille du pancréas augmente (Hebrok et al., 2000). Par contre, l'ajout de cyclopamine (qui bloque l'ensemble de la signalisation Hedgehog : Indian, Desert et Sonic hedgehog) induit une augmentation de la taille du pancréas (Thomas et al., 2000). L'étude de Thomas et al. suggère donc que d'autres facteurs que *Shh* seraient impliqués dans la régulation de la voie Hedgehog. Hebrok et al. ont dans ce sens mis en évidence que la mutation *Ihh*, quand elle est combinée avec celle de

Shh, entraîne un hypoplasie du pancréas, avec un enroulement de ce dernier autour du duodénum (Hebrok et al., 2000).

Bien que l'absence de notochorde n'empêche pas la formation du pancréas ventral, Deutsch et Al. ont montré la nécessité de l'inhibition de la signalisation hedgehog pour la formation du pancréas ventral. Tout comme le pancréas dorsal, l'ensemble de l'endoderme ventral antérieur n'exprime pas Shh (Deutsch et al., 2001). Cette totale absence de l'expression de Shh peut expliquer la différenciation par défaut de l'endoderme ventral antérieur en pancréas. Deutsch et Al. ont également montré que l'inhibition de la différenciation pancréatique dans l'endoderme ventral antérieur formant le foie est permise grâce à l'expression de Shh dans le mésoderme cardiaque, qui est en contact de l'épithélium (Fig. 4 ; Deutsch et al., 2001).

L'ensemble de ces résultats montre donc que la signalisation Hedgehog inhibe la différenciation de l'endoderme en pancréas. Par conséquent, son inhibition par les FGF et la voie du TGF β est nécessaire pour la formation du pancréas dorsal (Fig.4).

2.E La signalisation de l'acide rétinoïque

La signalisation de l'acide rétinoïque (AR ; Bastien et al., 2004 ; Ross et al., 2000) a également été impliquée dans le développement du pancréas.

La signalisation via l'AR a été impliquée dans l'initiation du développement du pancréas murin (Martin et al., 2005 ; Molotkov et al., 2005). Les souris mutantes pour *Raldh2* sont caractérisées par une perte du développement du pancréas dorsal, alors que le développement du pancréas ventral reste inchangé (Martin et al., 2005 ; Molotkov et al., 2005). Cette différence sur le développement des deux bourgeons pancréatiques peut s'expliquer par l'absence d'expression de *Raldh2* dans le mésoderme autour du pancréas ventral de l'embryon de souris à E9,5. Martin et Al. ont également montré que la perte de *Raldh2* entraîne une perte des cellules du mésenchyme autour du bourgeon dorsal, similaire au phénotype observé chez le mutant *Is11* (Martin et al., 2005 ; Ahlgren et al., 1997). Ainsi, l'absence de développement du pancréas dorsal, chez les souris *RALDH2* mutantes ne pourrait être qu'un effet secondaire de l'absence de signal provenant du mésenchyme autour du pancréas dorsal.

Des travaux récents ont mis en évidence que l'AR serait également nécessaire à la formation du pancréas ventral. En effet, l'expression d'une forme dominante négative du récepteur à l'AR α sous le contrôle du promoteur de *pdx1* induit une totale agénésie du pancréas (Ostrom et al., 2008). Bien que ces résultats suggèrent un rôle direct de l'AR dans le développement du pancréas, se pose la question du mécanisme par lequel l'AR est synthétisé au pôle ventral depuis que Martin et al. n'ont pu mettre en évidence l'expression d'aucune des isoformes de RALDH dans le mésenchyme entourant le bourgeon ventral (Martin et al., 2005).

Öström et al. ont en outre mis en évidence que l'AR serait impliqué dans la formation de l'épithélium pancréatique exprimant *Neurog3*, puis dans la différenciation de ces cellules en cellules β par l'induction d'un programme normal de différenciation (Ostrom et al., 2008).

En utilisant des poissons zèbres mutés pour *Raldh2* (enzyme nécessaire à la formation de l'AR) ou par injection d'antagoniste de la signalisation de l'AR, Stafford et al. ont montré que l'AR est nécessaire à la formation du pancréas (Stafford et al., 2002) ainsi qu'à la formation des cellules β (Stafford et al., 2006) chez le poisson zèbre. Des analyses similaires chez le xénope ont également mis en évidence la nécessité de la signalisation de l'AR afin d'induire le développement normal du pancréas (Chen et al., 2004). De plus, Stafford et al. ont également montré que l'expression ectopique de l'AR permet d'étendre le territoire d'où se développera le pancréas en postériorisant l'endoderme antérieur (Stafford et al., 2002).

De récents travaux chez l'embryon de poulet ont conforté ces derniers résultats. En effet, Bayha et al. ont mis en évidence que la mise en place d'un gradient d'AR est nécessaire à la spécification de l'endoderme sur l'ensemble de l'axe antéro-postérieur (Bayha et al., 2009).

L'ensemble de ces résultats montre donc que la signalisation de l'AR est nécessaire à la formation du pancréas dorsal et ventral, et serait également impliquée dans la régionalisation de l'endoderme.

2.F La signalisation Wnt

Comme l'ensemble des voies de signalisation décrites précédemment, la voie Wnt (Logan et al., 2004) est également impliquée dans la différenciation pancréatique.

Une analyse détaillée de l'expression des différents acteurs de la voie Wnt a montré qu'à E11, l'expression des Wnt est localisée dans le mésenchyme. Dans le pancréas, le niveau d'expression de *Wnt5a* est le plus élevé. L'expression de *Wnt7b* et *Wnt2b* sont également détectées dans le pancréas en développement. *Wnt5b*, *Wnt11* et *Wnt16* sont également exprimés de façon constante au cours du développement pancréatique mais à des niveaux plus faibles. Différents récepteurs de la voie Wnt (*Frz2, 3, 4, 5, 6, 7, 8* et *9*) sont également détectés dans les cellules épithéliales pancréatiques embryonnaires. En revanche, l'expression des inhibiteurs de la voie Wnt, sFRP, change au cours du développement. sFRP1 est détecté à E11 dans le mésenchyme pancréatique, alors que sFRP2 et sFRP3 sont détectés dans les cellules endocrines à E17. Finalement, LRP5 et LRP6, des corécepteurs de Wnt, sont exprimés pendant le développement pancréatique (Heller et al., 2002).

La fonction de Wnt a également été intensément étudiée. La première de ces études a été réalisée par l'expression ectopique de *Wnt1* ou *Wnt5* selon le patron d'expression de *Pdx1*. Heller et al. ont ainsi mis en évidence que la surexpression de *Wnt1* ou *Wnt5* induisent respectivement une agénésie ou une sévère hypoplasie du pancréas (Heller et al., 2002). Ces résultats suggèrent donc que l'inhibition de la voie Wnt est nécessaire à la différenciation des cellules pancréatiques.

L'étude de la voie Wnt s'est ensuite focalisée sur la voie canonique de Wnt. Pour cela, l'expression de la β -caténine a été spécifiquement supprimée selon le patron d'expression de *Pdx1* (Murtaugh et al., 2005 ; Wells et al., 2007). Ces deux études ont montré que la perte d'expression de la β -caténine n'induit pas de perte des progéniteurs pancréatiques exprimant *Pdx1* à E11,5. Par contre, alors que les cellules endocrines se développent normalement, ces deux études ont mis en évidence une perte de la différenciation exocrine presque totale (Murtaugh et al., 2005 ; Wells et al., 2007).

L'étude inverse a montré que l'expression précoce de la β -caténine selon le patron d'expression de *Pdx1* induisait, comme l'expression de *Wnt-1* et *Wnt-5*, une agénésie du pancréas (Heiser et al., 2006). De plus, cette agénésie du pancréas est corrélée avec une augmentation de l'expression de Shh et une perte de l'expression de FGF10. Ces derniers

résultats suggèrent donc que l'expression anormale de la voie Wnt déclenche à son tour, une expression anormale de la voie hedgehog et FGF. En revanche, l'induction d'une expression plus tardive de la β -caténine induit une augmentation de la taille du pancréas, avec une augmentation du nombre de cellules exocrines (Heiser et al., 2006).

Papadopoulou et al. ont montré que l'inhibition de la voie Wnt, par l'expression d'une forme dominante négative de FRZ8 sous le contrôle du promoteur *Pdx1* induit une réduction de la croissance pancréatique (Papadopoulou et al., 2005), alors que la perte d'expression de la β -caténine n'induit pas de perte des progéniteurs pancréatiques exprimant *pdx1*, mais une perte quasi complète de la différenciation exocrine (Murtaugh et al., 2005 ; Wells et al., 2007). Cependant, malgré cette hypoplasie drastique, ces souris ne présentent pas d'hyperglycémie. Ceci peut être expliqué par le fait que chez l'adulte la proportion des cellules endocrines par rapport à la surface pancréatique est augmentée par rapport à une souris sauvage, bien que le rapport des cellules α/β reste similaire à celui d'une souris sauvage.

Une autre étude, réalisée chez la souris, montre que la perte d'activité de la β -caténine selon le patron d'expression de *Pdx1* n'induit qu'une faible réduction de la différenciation endocrine, et n'induit aucun effet sur la différenciation exocrine (Dessimoz et al., 2006). Ces résultats sont donc en contradiction avec les travaux de Murtaugh et Wells (Murtaugh et al., 2005 ; Wells et al., 2007). Toutefois, des contrôles montrent que l'inactivation de la β -caténine n'est que partielle entre E14,5 et la naissance (Dessimoz et al., 2006). Cette observation suggère donc que le faible phénotype observé par Dessimoz et al. serait plutôt dû à une faible réduction de la signalisation de la voie Wnt, plutôt qu'à une non nécessité de la voie Wnt pour le développement pancréatique.

Finalement, la voie Wnt semble également être impliquée dans la régulation de la prolifération des cellules β dans les îlots de Langerhans. En effet, Rulifson et al. ont montré que l'ajout de WNT3a sur des cultures d'îlots isolés ou que l'expression d'une forme active de la β -caténine dans des cellules exprimant l'insuline induisaient la prolifération des cellules β . Réciproquement, l'inhibition de la voie Wnt, par une surexpression de l'axine dans les îlots de Langerhans, induit une réduction de la prolifération néonatale des cellules β (Rulifson et al., 2007).

L'ensemble de ces résultats suggère donc que les différents composants de la voie Wnt pourraient être impliqués soit dans l'induction de la différenciation pancréatique, soit dans la prolifération cellulaire, soit dans la différenciation exocrine, soit dans plusieurs de ces processus. Toutefois, la variation de phénotype d'une étude à l'autre nécessitera de réaliser d'autres études afin de déterminer le rôle exact de la voie Wnt dans la différenciation pancréatique.

2.G Les facteurs de transcription impliqués dans la différenciation précoce du pancréas

L'action régulée de l'ensemble des facteurs décrits ci dessus va ainsi permettre l'émergence des bourgeons pancréatiques ventral et dorsal, qui auront la capacité de se différencier dans l'ensemble des sous types cellulaires pancréatiques. En parallèle à ces signaux, l'expression de plusieurs facteurs de transcription est également requise aux premières étapes de l'initiation pancréatique. On peut notamment y remarquer les protéines à « homeodomain » HLXB9, ISL1, HEX et PDX1, la protéine à domaine HMG SOX9 et la protéine à domaine « basix Helix Loop Helix » (bHLH) PTF1a.

2.G.a HLXB9

L'expression de *Hlxb9* est induite dès E8 dans l'ensemble de l'endoderme intestinal primitif (Li et al., 1999) puis va rapidement être perdue dans le pancréas en différenciation et sera exprimé dans les îlots de Langerhans spécifiquement dans les cellules β (Harrison et al., 1999). L'étude de la perte de fonction de *Hlxb9* chez la souris a montré que HLXB9 est nécessaire à la formation du pancréas dorsal. En revanche, le développement du bourgeon ventral n'est pas affecté chez les souris mutantes (Harrison et al., 1999 ; Li et al., 1999). De plus, l'absence d'expression de Shh et Ihh dans les bourgeons pancréatiques montrent que HLXB9 n'est pas requis pour la régulation de la voie hedgehog (Li et al., 1999).

2.G.b ISL1

Le facteur de transcription à « homeodomain » LIM, ISL1, a également été impliqué dans le développement précoce du pancréas (Karlsson et al., 1990). Son expression est

biphasique. Dans un premier temps, *Isl1* est exprimé dans le mésenchyme autour du pancréas dorsal. Puis dans un second temps, *Isl1* est exprimé dans l'épithélium dorsal et ventral, et son expression sera conservée dans les cellules pancréatiques endocrines (Ahlgren et al., 1997). La perte de fonction d'*Isl1* est caractérisée par une absence du développement du bourgeon dorsal alors que le bourgeon ventral se forme normalement. Ce processus semble être cellule non autonome puisque des cocultures d'épithélium mutant pour *Isl1* avec du mésenchyme sauvage induit la prolifération de l'épithélium pancréatique et le sauvetage de la différenciation exocrine. Le pancréas des mutants *Isl1* est également caractérisé par une absence totale de cellules endocrines dans les cocultures d'épithélium mutant pour *Isl1* et de mésenchyme sauvage, suggérant qu'ISL1 serait requis pour la différenciation endocrine (Ahlgren et al., 1997).

2.G.c HEX

Le facteur de transcription à « homeodomain » HEX est, à l'opposé d'ISL1 et de HLXB9, impliqué dans la formation du pancréas ventral. En effet, en absence d'expression de *Hex*, le bourgeon pancréatique ventral n'est pas formé ; alors que le foie, qui est issu de l'inhibition par le mésoderme cardiaque de la différenciation du pancréas ventral, se développe normalement (Bort et al., 2004). Bort et al. ont montré que chez le mutant *Hex*, l'hypoplasie du pancréas ventral est due à une réduction de la prolifération de l'endoderme ventral antérieur. Celle-ci ayant pour conséquence de ne pas permettre un développement de l'endoderme antérieur ventral au delà du mésoderme cardiaque, engendrant par conséquent l'inhibition de la différenciation pancréatique du bourgeon ventral (Bort et al., 2004).

2.G.d PDX1

Le facteur de transcription PDX1 a été intensément étudié de par son rôle crucial dans le développement pancréatique ; et de nombreuses études ont utilisé son promoteur pour induire l'expression de gènes au cours du développement pancréatique. L'expression de *Pdx1* dans le pancréas a été caractérisée par Ohlsson et al. (Ohlsson et al., 1993). *Pdx1* est exprimé dès E8 dans l'endoderme intestinal primitif (Guz et al., 1995). A E10,5, son

expression est détectée dans la partie postérieure de l'estomac, le duodénum et les deux bourgeons pancréatiques (Ahlgren et al., 1996). Puis dans le pancréas en développement, son expression est perdue ; elle est ensuite réinduite dans les cellules β en différenciation et dans les îlots de Langerhans (Ohlsson et al., 1993 ; Ahlgren et al., 1998).

Plusieurs études ont mis en évidence le rôle primordial de PDX1 dans la différenciation pancréatique. En effet, la perte de fonction de PDX1 entraîne une absence de pancréas chez les souriceaux nouveaux nés (Jonsson et al., 1994). Seul un début de bourgeonnement dorsal de l'épithélium est observé entre E10,5 et E12,5 (Ahlgren et al., 1996). Offield et al. ont également mis en évidence que l'expression de *Pdx1* est requise pour le développement normal de la jonction pylorique avec le duodénum, et la différenciation de la partie antérieure du duodénum (Offield et al., 1996). D'autres études ont également analysé la capacité de PDX1 à induire une différenciation pancréatique. Il a ainsi été mis en évidence que PDX1 pouvait induire l'expression de HLXB9 et de NKX6.1, ainsi qu'une structure similaire au bourgeon pancréatique (Grapin-Botton et al., 2001 ; Pedersen et al., 2005).

Différents gènes ont été caractérisés pour réguler l'expression de *Pdx1*. L'étude de la souris mutante pour *Tcf2*, qui est caractérisée par une agénésie du pancréas bien que le bourgeonnement dorsal pancréatique s'initie (comme chez la souris mutante pour *Pdx1*), a mis en évidence la perte d'expression de *Pdx1*, suggérant que l'expression de TCF2 est nécessaire à l'expression de *Pdx1*. Il a été également montré que l'expression de FOXa2 est nécessaire pour l'expression de *Pdx1* (Wu et al., 1997). Jacquemin et al. ont également montré que la perte de fonction de *Oc-1* entraîne un retard dans la spécification pancréatique de l'endoderme, induisant ainsi une réduction de l'expression de *Pdx1*, et par conséquent une hypoplasie du pancréas (Jacquemin et al., 2003).

2.G.e PTF1a

PTF1a, un facteur de transcription à bHLH a également été caractérisé comme étant nécessaire au développement précoce du pancréas. Initialement, *Ptf1a* a été caractérisé pour être exprimé dans les cellules exocrines (Krapp et al., 1996). Krapp et al. ont montré que la perte de fonction de *Ptf1a* induit une perte du lignage exocrine. Ils ont également mis en évidence que la perte de fonction de *Ptf1a* influence la formation du pancréas endocrine. En effet, ils ont montré que les cellules pancréatiques endocrines se développent

normalement jusqu'à E16,5. Puis, chez les souriceaux nouveaux nés *Ptf1a* mutants, les cellules pancréatiques endocrines ne sont plus détectées dans le pancréas, mais sont relocalisées dans la rate (Krapp et al., 1998). Ces derniers résultats suggèrent donc que *Ptf1a* serait requis à la structuration du pancréas, et qu'en son absence, les cellules endocrines seraient relocalisées vers la rate, une caractéristique observée chez certains vertébrés inférieurs tels que les reptiles (Krapp et al., 1998). Toutefois, par une expérience de traçage cellulaire, Kawaguchi et al. ont montré que les cellules exprimant les hormones pancréatiques localisées dans la rate n'ont pas exprimé *Ptf1a* (Kawaguchi et al., 2002). D'autres études ont montré que *Ptf1a* est exprimé dans les progéniteurs pancréatiques dès E9 (Kawaguchi et al., 2002 ; Hald et al., 2008). Kawaguchi et al. ont également montré que les progéniteurs pancréatiques exprimant *Ptf1a* donneront naissance à l'ensemble des différents lignages cellulaires constituant le pancréas adulte (Kawaguchi et al., 2002). Ils ont également montré par des expériences de traçage cellulaire que chez les mutants *Ptf1a*, les cellules qui auraient dû exprimer *Ptf1a* sont relocalisées dans le duodénum (Kawaguchi et al., 2002). Ces résultats suggèrent donc que *Ptf1a* serait requis pour la différenciation pancréatique de l'épithélium exprimant *pdx1* et empêcherait une différenciation intestinale de ces progéniteurs. D'autres études ont également montré que l'expression ectopique de PTF1a, combinée à celle de PDX1, ou l'expression d'une forme constitutivement active de PTF1a, permet d'étendre à l'endoderme postérieur la formation de progéniteurs pancréatiques, conduisant au développement d'un pancréas géant chez le xénope ; ou permet la transformation de l'endoderme hépatique en pancréas (Afelik et al., 2006 ; Jarikji et al., 2007).

2.G.f SOX9

La protéine à domaine HMG SOX9 est exprimée dès E9,5 dans les bourgeons pancréatiques (Lioubinski et al., 2003). De plus, il a été montré par une expérience de traçage cellulaire que *Sox9* est exprimé dans l'ensemble de l'épithélium pancréatique suggérant qu'il est exprimé dans les progéniteurs pancréatiques (Akiyama et al., 2005). Seymour et al. ont également mis en évidence que la perte de fonction de *Sox9* induit une sévère hypoplasie pancréatique (Seymour et al., 2007). Cette hypoplasie est due à une réduction du nombre de progéniteurs pancréatiques dont SOX9 contrôle la prolifération par

la régulation de l'expression de *FoxA2*, *Tcf2*, *Oc-1* et *Hes1* (Lynn et al., 2007 ; Seymour et al., 2007). De plus, des travaux récents ont montré que SOX9 est également impliqué dans la formation des progéniteurs pancréatiques endocriniens. En effet, Seymour et al. ont montré que la mutation de *Sox9* à l'état hétérozygote entraîne une réduction spécifique du nombre de cellules endocrines, liée à une baisse du nombre de cellules NEUROG3 (Seymour et al., 2008).

L'ensemble de ces études montre que l'induction du développement du pancréas est à la fois sous le contrôle de signalisation externe à l'épithélium pancréatique, ainsi que sous le contrôle de l'expression de gènes dans l'épithélium, nécessaires au développement du pancréas. Ensemble, ils permettent de générer à partir de l'endoderme intestinal primitif, la formation de deux bourgeons pancréatiques qui permettront de donner naissance au pancréas chez l'individu adulte.

C) Contrôle génétique de la différenciation du pancréas endocrine

1. Induction de la différenciation des différents lignages cellulaires pancréatiques

Une fois les progéniteurs pancréatiques exprimant notamment PDX1 et PTF1a formés, l'induction de leur différenciation dans les trois lignages cellulaires constituant le pancréas va permettre de former les cellules acinaires, les cellules canalaire et les cellules endocrines.

Cette régulation de la différenciation des différents lignages pancréatiques est permise par la voie de signalisation notch (Fig. 5; Schweisguth, 2004 ; Lai, 2004). Différents partenaires de la voie Notch ont été identifiés dans le pancréas en développement. *Notch-1*, *-2* sont exprimés dans l'épithélium exocrine en formation (Jensen et al., 2000a). *Notch-3* est spécifiquement exprimé dans le mésenchyme pancréatique (Jensen et al., 2000b). *Notch-4* est lui exprimé dans les cellules endothéliales présentes dans le pancréas en développement (Lammert et al., 2000). Les ligands de Notch, *Dll1*, *Dll3*, *Jag1*, *Jag2* sont exprimés au cours du développement pancréatique. *Hes1*, une cible de Notch, est également exprimée dans le pancréas embryonnaire (Jensen et al., 2000b).

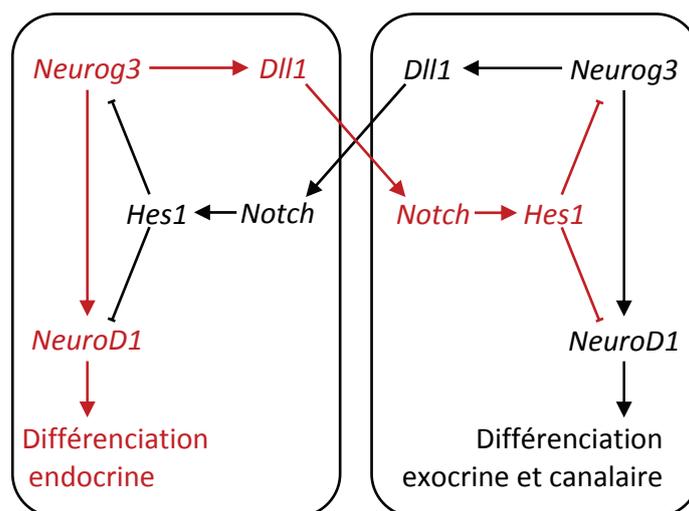


Figure 5 : L'induction de la voie Notch par *Neurog3* induit la différenciation des différents lignages pancréatiques. L'expression de *Neurog3* induit la voie notch qui par un mécanisme d'inhibition latérale, va inhiber l'expression de *Neurog3* dans les cellules adjacentes. Par conséquent, ces cellules s'engageront dans le lignage exocrine ou canalaire. A l'inverse, les cellules exprimant *Neurog3* se différencieront en cellules endocrines (adaptée de Skipper et Lewis, 2000).

La première évidence de l'implication de la voie notch a été montrée par les travaux d'Apelqvist et al.. Ils ont mis en évidence que la perte de fonction pour *Rbp-jκ*, un coactivateur de la voie notch, induit une augmentation à E8,5-E9,5 de l'expression de *neurog3*, qui a été décrit comme un gène proendocrine (Apelqvist et al., 1999). Apelqvist et al. ont également étudié l'impact de la perte de fonction de *Dll1*, le ligand de notch, nécessaire au clivage de la protéine Notch permettant son activité intracellulaire. Ils ont ainsi montré que chez les souris mutantes pour *Dll1*, la taille du bourgeon pancréatique est réduite, et que le nombre de progéniteurs pancréatiques exprimant PDX1 est réduit parallèlement à une augmentation du nombre des cellules exprimant *Neurog3* à E9,5, et *Neurod1* ou le glucagon à E10,5 (Apelqvist et al., 1999). Jensen et al. ont également mis en évidence que la perte de fonction de *hes1* induit une réduction des précurseurs pancréatiques pluripotents suite à une différenciation accélérée en cellules endocrines exprimant le glucagon (Jensen et al., 2000b). Il a également été mis en évidence que la surexpression de *NEUROG3* selon le patron d'expression de *pdx1* induit une réduction de la taille du pancréas en développement, avec une augmentation précoce de la différenciation endocrine (Apelqvist et al., 1999 ; Schwitzgebel et al., 2000). La surexpression de *Notch1* selon le patron de *Pdx1* dès E9,5 induit également une hypoplasie pancréatique, couplée à une différenciation précoce des progéniteurs pancréatiques en cellules endocrines

(Murtaugh et al., 2003). L'ensemble de ces résultats montre donc que l'induction précoce de la voie notch a pour fonction d'inhiber une différenciation précoce des progéniteurs pancréatiques en cellules endocrines, en inhibant l'expression de *Neurog3* via *Hes1*.

Toutefois, Notch, au-delà de son rôle dans l'inhibition de la différenciation précoce des progéniteurs pancréatiques, est également indispensable à la différenciation des différents lignages pancréatiques. Afin d'étudier les fonctions de la voie notch à différents stades au cours du développement pancréatique, Murtaugh et al. ont créé un modèle de souris où le domaine intracellulaire de Notch1 est conditionné à l'expression de la Cre couplée au domaine de liaison du ligand du récepteur des œstrogènes modifié pour reconnaître spécifiquement le 4-hydroxytamoxifène (Murtaugh et al., 2003). Ils ont ainsi montré que l'induction ectopique de la voie notch à E13,5 ou E14,5 induisait la formation d'un épithélium tubulaire dépourvu de cellules endocrines et exocrines. En revanche, l'induction de la voie notch à E15,5 n'a pas d'effet sur la croissance pancréatique, bien que les cellules exprimant le domaine intracellulaire de NOTCH1 n'exprime ni l'insuline, ni l'amylase (Murtaugh et al., 2003). Esni et al. ont également montré que l'activation ectopique de la voie Notch dans des explants pancréatiques murins en culture induit une réduction de la différenciation exocrine. Ils ont également confirmé ce résultat chez le poisson zèbre (Esni et al., 2004). Récemment, il a également été suggéré que Notch serait nécessaire à la formation du lignage canalaire pancréatique. En effet, Lorent et al. ont montré une perte de l'expression des marqueurs canaux pancréatiques en faveur d'une augmentation de la taille des acini chez le poisson zèbre où l'expression des ligands de Notch, Jagged, est perdue (Lorent et al., 2004). Yee et al. ont également mis en évidence chez le poisson zèbre que l'expression ectopique de Notch induit une perte de la formation des canaux pancréatiques (Yee et al., 2005). Récemment, il a aussi été démontré que l'expression du domaine intracellulaire de notch selon le patron d'expression de pax4 chez la souris induit un développement accru des cellules canaux (Greenwood et al., 2007).

L'ensemble de ces résultats montre donc que la signalisation de la voie Notch est nécessaire au maintien des progéniteurs pancréatiques après le bourgeonnement ventral et dorsal du pancréas, ainsi qu'à la régulation de l'induction de la différenciation des trois lignages cellulaires composant le pancréas.

2. Neurog3 est nécessaire à l'engagement de la différenciation du pancréas endocrine

Comme suggéré par les travaux d'Apelqvist et al., le facteur de transcription à bHLH NEUROG3 semble être impliqué dans la différenciation des progéniteurs pancréatiques en cellules endocrines (Apelqvist et al., 1999). L'étude de la perte de fonction de *Neurog3* a permis de définir le rôle de NEUROG3 dans le pancréas en développement. Il a ainsi été montré qu'en l'absence de l'expression de *Neurog3*, le pancréas des souris nouveau nées est caractérisé par une agénésie du pancréas endocrine. Ces souriceaux, présentant une hyperglycémie, meurent rapidement après la naissance (Gradwohl et al., 2000). Ces résultats montrent donc que NEUROG3 est nécessaire à la formation des cellules endocrines, en engageant les progéniteurs pancréatiques dans la différenciation endocrine.

L'expression de NEUROG3 dans le pancréas commence à être détectable dans le bourgeon dorsal à E8,5-E8,75 et dans le bourgeon ventral à E9,5 (Jorgensen et al., 2007). Son expression va ensuite croître dans l'épithélium pancréatique pour atteindre son maximum à E15,5 (Apelqvist et al., 1999 ; Gradwohl et al., 2000). Son expression va ensuite fortement décroître pour être détectée dans quelques rares cellules à E17,5 (Apelqvist et al., 1999 ; Gradwohl et al., 2000). De récents travaux réalisés par le groupe d'Harry Heimberg suggèrent également que des cellules NEUROG3 positives seraient présentes dans le pancréas adulte sous certaines conditions. En effet, Xu et al. ont montré une forte augmentation du nombre de cellules NEUROG3 dans le pancréas adulte après la ligature au niveau de la queue du pancréas induisant la destruction du pancréas par les sucs pancréatiques (Xu et al., 2008). De plus, la greffe des cellules exprimant *Neurog3* permet de restaurer la différenciation endocrine chez des mutants *Neurog3*, chez qui la différenciation endocrine est totalement perdue (Xu et al., 2008). Une étude récente suggère également que NEUROG3 serait exprimée dans la majorité des cellules des îlots de Langerhans chez l'adulte, bien que le niveau de son expression soit beaucoup plus faible que durant l'embryogénèse (Wang et al., 2009b).

De nombreuses études ont étudié l'impact de l'expression de NEUROG3 au cours du développement embryonnaire. Il a ainsi été mis en évidence que l'expression de *Neurog3* dans l'endoderme de poulet est nécessaire et suffisante à l'engagement de l'endoderme dans la différenciation du pancréas endocrine (Grapin-Botton et al., 2001). Ces résultats ont été confirmés par l'expression ectopique de *Neurog3* dans des pancréas de souris en culture

à E13,5 (Dominguez-Bendala et al., 2005). Gu et al. ont également montré par des expériences de traçage cellulaire que NEUROG3 est spécifiquement requis pour la formation du lignage endocrinien, et que l'ensemble des cellules endocrines du pancréas ont exprimé Neurog3 (Gu et al., 2002). Il a également été montré que l'expression ectopique précoce de neurog3 à E10,5 selon le patron d'expression de *Pdx1* induit la formation d'un pancréas hypoplasique où la majorité des cellules endocrines exprime le glucagon (Apelqvist et al., 1999 ; Schwitzgebel et al., 2000 ; Johansson et al., 2007).

L'ensemble de ces résultats met ainsi en évidence le rôle fondamental du facteur de transcription NEUROG3 dans l'induction de cascades génétiques nécessaires à la différenciation du pancréas endocrine (Fig. 6). Rôle qui n'a été que confirmé depuis, puisque l'expression endocrine des facteurs de transcription impliqués dans la différenciation des sous types pancréatiques endocriniens est perdue chez les souris mutées pour *Neurog3* (Gradwohl et al., 2000 ; Mellitzer et al., 2006 ; Petri et al., 2006).

De plus, des travaux récents ont montré que la compétence des progéniteurs pancréatiques endocrines exprimant NEUROG3 à se différencier dans les différents sous types cellulaires endocriniens varie au cours de l'embryogénèse. Johansson et al. ont montré que les progéniteurs endocrines ont la compétence de se différencier en cellules α sécrétant le glucagon dès E8,5, alors qu'ils ne peuvent se différencier en cellules β sécrétant l'insuline ou δ sécrétant la somatostatine qu'à partir de E10,5, et en cellules PP sécrétant le PP qu'à partir de E14,5. De plus, la compétence à se différencier en cellules α est drastiquement réduite entre E8,5 et E14,5 en faveur des autres types cellulaires pancréatiques endocriniens (Johansson et al., 2007).

L'ensemble de ces résultats montre donc qu'au cours du développement pancréatique NEUROG3 est nécessaire et suffisante à l'engagement des progéniteurs pancréatiques dans la différenciation endocrine et à la différenciation des différents sous types cellulaires pancréatiques endocriniens.

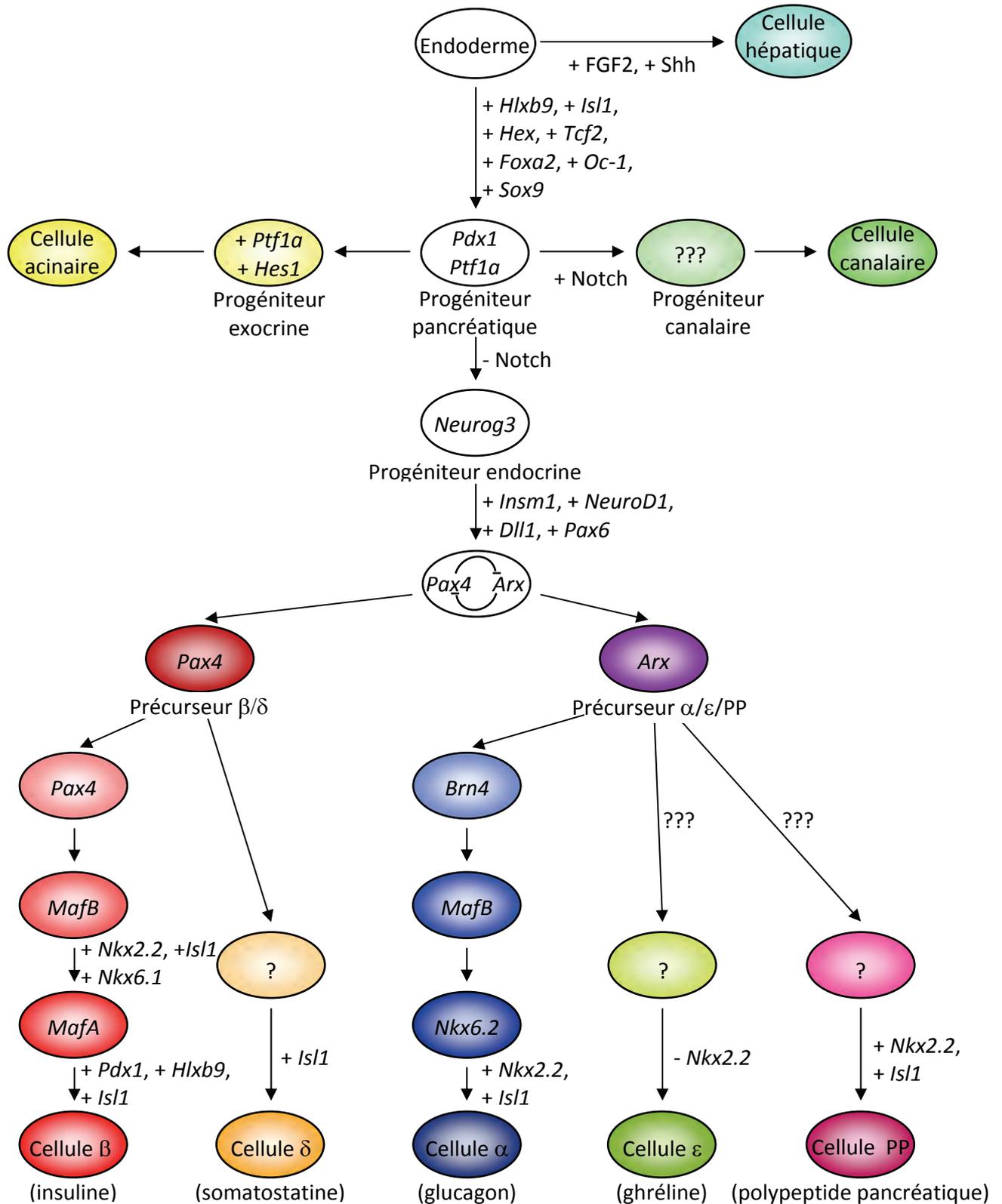


Figure 6 : Modèle schématique de la différenciation du pancréas endocrine. Ce schéma est basé sur le profil d'expression des gènes représentés et sur les phénotypes observés chez les gains et pertes de fonction de ces gènes (adaptée de Collombat et al., 2006).

3. les facteurs de transcription impliqués dans la différenciation des progéniteurs pancréatiques endocrines

En parallèle à l'étude de l'induction de la différenciation pancréatique, et de l'étude de la régulation de la différenciation exocrine/endocrine, de nombreuses études ont été réalisées sur la différenciation des sous types endocriniens. De nombreux facteurs de transcription ont ainsi été caractérisés et se répartissent en trois catégories : les facteurs de transcription pancréatiques pan-endocriniens, ceux impliqués dans le choix des sous types endocriniens et ceux impliqués dans la différenciation terminale des cellules pancréatiques endocrines (Fig. 6).

3.A Les facteurs de transcription pancréatiques pan-endocriniens

Parmi les facteurs de transcription impliqués dans la différenciation endocrine, les facteurs de transcription pan-endocriniens sont impliqués dans la différenciation de l'ensemble des sous types cellulaires pancréatiques endocriniens. Ces facteurs sont notamment NeuroD1, PAX6, INSM1 et ISL1 (Fig. 6).

3.A.a NEUROD1

NEUROD1, facteur de transcription à bHLH, a été caractérisé comme gène pan-endocrinien grâce aux travaux de Naya et al. (Fig. 6 ; Naya et al., 1997). En effet, Naya et al. ont montré que la perte de fonction de *NeuroD1* entraîne une perte partielle de la différenciation des différents sous types cellulaires endocriniens, accompagnée d'une induction de l'apoptose des cellules endocrines à la naissance des souriceaux. Ces deux phénomènes ont pour conséquence l'absence de formation d'îlots, induisant une hyperglycémie fatale pour les souriceaux (Naya et al., 1997). De plus, les cellules exocrines dont la différenciation est conservée chez les mutants *NeuroD1* montrent une accumulation de grains de sécrétion dans les cellules acinaires. Cette accumulation de grains pourrait être due à la perte de l'expression de la cholecystokinine perdue chez le mutant *NeuroD1*, normalement nécessaire pour la sécrétion des cellules acinaires (Naya et al., 1997).

De plus, *NeuroD1* a été caractérisé comme une cible directe de NEUROG3 (Huang et al., 2000). L'expression ectopique de *NeuroD1* selon le patron d'expression de *Pdx1* chez des

souris a mis en évidence que le phénotype induit par cette expression précoce de *NeuroD1* est similaire à celui de *Neurog3* dans les mêmes conditions (Schwitzgebel et al., 2000). De plus, l'expression dans des lignées cellulaires pancréatiques de *NeuroD1* ou de *Neurog3* a montré que comme *NEUROG3*, *NEUROD1* est capable d'induire un programme de différenciation endocrine normale (Heremans et al., 2002).

L'ensemble de ces résultats montre l'implication de *NeuroD1* dans les différents programmes de différenciation conduisant à la formation des sous types cellulaires pancréatiques endocriniens.

3.A.b PAX 6

Comme *NEUROD1*, l'analyse de la souris mutante pour *Pax6* a montré la nécessité de *PAX6* pour la différenciation des différents sous types pancréatiques endocriniens (Sander et al., 1997 ; St-Onge et al., 1997). Il a en effet été montré que la perte de fonction de *Pax6*, un facteur de transcription à domaine « homeodomain », induit une perte de la différenciation de l'ensemble des sous types pancréatiques endocriniens, bien que des différences dans les phénotypes soient observées dans les deux études. St-Onge et al. ont mis en évidence que la perte de fonction de *pax6* induit une perte quasi-totale de la différenciation des cellules α et une réduction de la différenciation des autres sous types endocriniens (St-Onge et al., 1997). Sander et al. ont eux montré que le mutant *pax6* (*Sey^{Neu}*) est caractérisé par une réduction de l'ensemble des sous types cellulaires endocriniens sans observer une perte des cellules α (Sander et al., 1997). Toutefois, l'observation d'un phénotype légèrement plus drastique dans l'étude de St-Onge et al. peut être dû aux différences de construction induisant la perte de fonction de *Pax6*. En effet, St-Onge et al. ont utilisé une souris dont la séquence codant la protéine *Pax6* a été remplacé par le gène de la β -galactosidase, alors que la souris mutante *Sey^{Neu}* est issue d'une mutation ponctuelle de *Pax6* induisant la formation d'une protéine *PAX6* tronquée.

De plus, le fait que l'expression de *Pax6* soit induite par *NEUROD1* peut expliquer le rôle pancréatique pan-endocrinien de *PAX6* (Marsich et al., 2003).

3.A.c INSM1

INSM1, facteur de transcription à « Zinc Finger », a récemment été caractérisé comme nécessaire à la différenciation endocrine. Mellitzer et al. puis Gierl et al. ont montré que la perte de fonction de *Insm1* induisait une réduction de la différenciation du pancréas endocrine (Mellitzer et al., 2006 ; Gierl et al., 2006). De plus, cette réduction de la différenciation endocrine est liée à une accumulation des progéniteurs endocrines chez les souris mutantes pour *Insm1* (Gierl et al., 2006). Cette accumulation pourrait être due au fait qu'INSM1 soit requis pour la progression du cycle cellulaire (Zhang et al., 2009). Le gène *Insm1* a également été caractérisé comme une cible directe de NEUROG3 (Mellitzer et al., 2006).

3.A.d ISL1

Le facteur de transcription à « homeodomain » LIM, ISL1, avait été précédemment caractérisé comme impliqué dans la différenciation précoce du pancréas (cf § 2.G.b ; Ahlgren et al., 1997). Les travaux d'Ahlgren et al. suggéraient également qu'ISL1 pourrait être impliqué dans la différenciation des différents sous types cellulaires pancréatiques endocriniens. En effet, ils ont montré qu'ISL1 est exprimé dans l'ensemble de ces sous types, et que les cultures d'épithélium mutant pour *isl1* en présence du mésenchyme sauvage ne permet pas d'observer de différenciation du pancréas endocrine (Ahlgren et al., 1997). Toutefois, due à une létalité précoce du mutant *isl1*, sa fonction au cours de la différenciation des progéniteurs endocrines n'avait pu être déterminée précisément. De récents travaux ont permis de répondre à cette question en générant un mutant conditionnel d'*isl1* (Du et al., 2009). Ils ont ainsi montré que la perte d'expression d'*isl1* à partir de E13,5 entraîne une réduction de la différenciation des cellules α , β , δ et PP mais n'affecte pas la différenciation des cellules ϵ . Conjugée à une réduction de la prolifération des cellules endocrines lors de la formation des îlots (à partir de E18,5), ces souris présentent une réduction du nombre de cellules pancréatiques endocrines chez l'adulte (Du et al., 2009).

Ces résultats montrent donc l'implication d'*isl1* dans la différenciation des différents sous types cellulaires pancréatiques endocriniens.

3.B Les facteurs de transcription impliqués dans le choix des sous types endocriniens

En parallèle à ces gènes pan-endocriniens, deux gènes ont été caractérisés comme régulant le choix du sous type cellulaire pancréatique endocrinien : *Pax4* et *Arx* (Fig. 6).

3.B.a PAX4

Le premier d'entre eux à avoir été caractérisé est *Pax4*, codant pour un facteur de transcription à domaine « paired box ». Il a ainsi été mis en évidence que la perte de fonction de *pax4* induit une perte spécifique de la différenciation des cellules β et δ , lié à une augmentation des cellules α (Sosa-Pineda et al., 1997, Collombat et al., 2005). Il a également été montré que la perte de fonction de *Pax4* induit une augmentation de la différenciation des cellules ε (Wang et al., 2008). Récemment, Collombat et al. ont montré que l'expression ectopique de *Pax4* au cours du développement pancréatique embryonnaire ou dans les îlots de Langerhans redirige les cellules α en cours de différenciation vers la différenciation en cellule β (Collombat et al., 2009).

3.B.b ARX

Le deuxième de ces deux gènes est *Arx*, codant pour un facteur de transcription à « homeodomain ». Il a en effet été montré par l'étude de la perte de fonction d'ARX, que l'expression d'ARX est requise pour la différenciation des cellules α et, que l'absence d'expression d'*Arx* induit une augmentation de la différenciation des cellules β et δ , et une réduction de l'expression du transcrite du PP (Collombat et al., 2003 ; Collombat et al., 2007). De plus, l'expression ectopique d'*Arx* selon le patron d'expression de *Pdx1* induit une perte de la différenciation des cellules β et δ , et une augmentation de la différenciation des cellules α et PP (Collombat et al., 2007).

3.B.b Arx et Pax4

Collombat et al. ont également mis en évidence que ARX et PAX4 régulent réciproquement leur expression en se fixant sur leur promoteur et en inhibant leur expression (Collombat et al., 2003 ; Collombat et al., 2005; Collombat et al., 2003). Il a en outre été montré que chez les souris mutantes pour *Pax4* et *Arx*, les lignages α et β étaient perdus en faveur d'une augmentation de la différenciation en cellules δ ou PP (Collombat et al., 2005).

L'ensemble de ces résultats montre donc l'importance d'ARX et de PAX4 pour la régulation du choix du sous type cellulaire pancréatique endocrinien. Ils montrent sans équivoque que PAX4 et ARX sont requis respectivement pour la différenciation en cellules sécrétant l'insuline ou la somatostatine, ou pour la différenciation en cellules sécrétant le glucagon. L'induction de la différenciation de cellules sécrétant la ghréline est inhibée par PAX4, suggérant donc que l'expression d'ARX pourrait être nécessaire pour leur différenciation. L'induction de la différenciation des cellules sécrétant le PP reste à confirmer car bien qu'elle semblerait être régulée par ARX, l'absence de variation du nombre de cellules PP chez le mutant *Arx* empêche de conclure définitivement quand au rôle d'*Arx* dans la différenciation des cellules PP. De plus, l'observation d'une augmentation de la différenciation des cellules PP et δ chez les doubles mutants *Arx* et *Pax4* suggère que les cellules PP et δ pourraient se différencier également indépendamment du contrôle d'ARX et de PAX4.

3.C Les facteurs de transcription impliqués dans la différenciation terminale des cellules pancréatiques endocrines

La différenciation du pancréas nécessite une dernière classe de facteurs de transcription qui sera requis pour la différenciation et la maturation d'un ou plusieurs sous types cellulaires endocriniens. On peut notamment y remarquer les protéines MAFA, MAFB, HLXB9, PDX1, NKX6.1, NKX6.2, NKX2.2 et BRN4 (Fig. 6).

3.C.a MAFA et MAFB

MAFA et MAFB sont deux facteurs de transcription à motif « basic Leucine Zipper ». MAFA a été le premier à avoir été caractérisé pour être exprimé dans le pancréas, où son expression est restreinte aux cellules β (Olbrot et al., 2002 ; Matsuoka et al., 2003). Son expression est régulée par PDX1, NKX2.2 et FOXa2 qui sont tous exprimés dans les cellules β (Raum et al., 2006). La perte de fonction de *MafA* a mis en évidence que la fonction de MAFA est limitée à la maturation des cellules β et notamment la sécrétion de l'insuline puisque aucune malformation n'est observée dans le pancréas en développement (Zhang et al., 2005). De plus, MAFA régule l'expression de PDX1 qui est nécessaire à la maturation des cellules β (Samaras et al., 2003 ; Zhang et al., 2005).

L'expression de MAFB varie au cours du développement pancréatique. Il est dans un premier temps exprimé dans une partie des progéniteurs pancréatiques exprimant NEUROG3. Son expression va ensuite se restreindre aux cellules α et β immatures. Finalement, MAFB sera spécifiquement exprimé dans les cellules α dans les îlots de Langerhans (Gu et al., 2004 ; Artner et al., 2006 ; Nishimura et al., 2006). La perte de fonction de *MafB* a montré que MAFB est nécessaire au développement des cellules α et β , et notamment pour leur différenciation finale. En effet, chez les mutants *MafB*, de nombreuses cellules expriment les facteurs de transcription observés chez les cellules α et β matures tels qu'ISL1, PAX6 et NEUROD1. Toutefois, leur différenciation va s'arrêter et ces cellules n'exprimeront ni le glucagon, ni l'insuline (Artner et al., 2007). En revanche, la formation des cellules δ et PP n'est pas affectée (Artner et al., 2007).

3.C.b HLXB9

HLXB9, facteur de transcription à « homeodomain », est lui impliqué dans la différenciation des cellules β . Chez les souris mutantes pour *Hlxb9*, le bourgeon dorsal pancréatique ne se développe pas, et le pancréas se développant à partir du bourgeon ventral montre une forte diminution du nombre de cellules β dans les îlots de Langerhans. De plus, ces cellules β semblent être immatures (Harrison et al., 1999). Cette différence entre les différents sous types cellulaires endocriniens et l'absence de maturation des cellules β suggère que HLXB9 est nécessaire à la maturation des cellules sécrétant l'insuline.

3.C.c PDX1

Tout comme HLXB9 qui est impliqué dans le développement précoce du pancréas, PDX1, un facteur de transcription à « homeodomain », est également impliqué dans la maturation des cellules β , et notamment dans leur réponse au glucose. En effet, il a été montré que l'activité de PDX1 est régulée selon la concentration de glucose sanguin, et que PDX1 régule l'expression de l'insuline (MacFarlane et al., 1994 ; Marshak et al., 1996). De plus, plusieurs études conduisant à la perte d'expression de *Pdx1* spécifiquement dans les cellules β en différenciation ont montré la nécessité de l'expression de *Pdx1* pour la différenciation normale des cellules β (Ahlgren et al., 1998 ; Lottmann et al., 2001 ; Thomas et al., 2001 ; Holland et al., 2002 ; Gannon et al., 2008). L'étude de Gannon et al. a également mis en évidence que PDX1 est requis pour la régulation du nombre de cellules des différents sous types endocriniens (Gannon et al., 2008).

3.C.d NKX6.1 et NKX6.2

NKX6.1 et NKX6.2 sont des facteurs de transcription à « NK-homeodomain ». Ils sont exprimés dans l'endoderme pancréatique, puis leur expression va se restreindre au lignage endocrine pour NKX6.1, alors que l'expression de NKX6.2 est également détectée dans l'épithélium pancréatique exocrine (Oster et al., 1998 ; Sander et al., 2000 ; Henseleit et al., 2005). La perte de fonction pour *Nkx6.1* a montré que NKX6.1 est requis pour le développement des cellules β (Sander et al., 2000). Le double mutant *Nkx6.1/Nkx6.2* a mis en évidence que NKX6.2 peut compenser partiellement la perte de fonction de *Nkx6.1* en engendrant une perte plus importante de la différenciation des cellules β (Henseleit et al., 2005). Le double mutant *Nkx6.1/Nkx6.2* est également caractérisé par une perte partielle des cellules α , suggérant donc que *Nkx6.2* est nécessaire à la formation des cellules α . Toutefois, le fait que le simple mutant *nkx6.2* ne permette pas d'observer d'effet sur la différenciation des cellules α montre que NKX6.1 peut compenser totalement la perte de fonction de *Nkx6.2* (Henseleit et al., 2005). Une étude récente a également confirmé que NKX6.2 peut compenser la perte de fonction de *Nkx6.1* (Nelson et al., 2007). Mais

étonnement, ce sauvetage du lignage β n'est possible que si *Nkx6.2* ou *Nkx6.1* est exprimé sous le contrôle du promoteur *Pdx1*, et non *Neurog3* (Nelson et al., 2007).

3.C.e NKX2.2

Nkx2.2 est également un facteur de transcription à « NK-homeodomain », et contient de même un domaine répresseur de type « engrailed ». NKX2.2 est exprimé dès E9,5 dans l'épithélium pancréatique. Puis au cours de la différenciation, son expression va se restreindre au lignage endocrine pour être détectée finalement dans les cellules α , β et PP (Rudnick et al., 1994 ; Sussel et al., 1998). La perte de fonction de *Nkx2.2* a montré que NKX2.2 est requis pour la différenciation des cellules α , β et PP (Sussel et al., 1998). Prado et al. ont également montré que les mutants *Nkx2.2* sont caractérisés par une augmentation du nombre de cellules ε , suggérant que NKX2.2 serait impliqué dans la différenciation des cellules α , β et PP, et la répression de la différenciation des cellules ε (Prado et al., 2004). Avec son rôle dans la formation des cellules β , NKX2.2 peut induire l'expression de *MafA* et de l'insuline (Raum et al., 2006 ; Cissell et al., 2003). De plus, il a été montré récemment que *Nkx2.2* est requis pour le maintien de la structure des îlots, la sécrétion de l'insuline et l'homéostasie sanguine du glucose (Doyle et al., 2007b). Doyle et al. ont également montré que le domaine de répression de NKX2.2 est nécessaire à l'induction de la différenciation des cellules α , et à la répression de la différenciation des cellules ε (Doyle et al., 2007a). Le double mutant *NeuroD1* et *Nkx2.2* a de plus été généré récemment. Il a ainsi été montré que la perte de fonction de *NeuroD1* avec celle de *Nkx2.2* permet le sauvetage de la différenciation de la majorité des cellules α et PP normalement perdues chez le mutant *Nkx2.2*, et réprime la différenciation des cellules ε . (Chao et al., 2007). Ces résultats montrent donc que NEUROD1 peut réguler la fonction de NKX2.2 dans le choix de la différenciation du lignage endocrine.

3.C.f BRN4

BRN4 est un facteur de transcription à domaine Pou. Il a été caractérisé comme étant spécifiquement exprimé dans les cellules α (Heller et al., 2004). Toutefois, la perte de fonction de *Brn4* ne montre aucun phénotype pancréatique, suggérant qu'un autre facteur

de transcription à domaine Pou exprimé dans le pancréas pourrait compenser la perte d'expression de *Brn4* (Heller et al., 2004).

L'ensemble de ces signaux et l'expression de ces différents gènes seront nécessaires au développement du pancréas endocrine en parallèle duquel le pancréas exocrine se développera. Suite à une structuration en îlots des cellules endocrines et en acini des cellules acinaires, le pancréas formera chez l'individu adulte un organe qui aura pour fonction de permettre l'homéostasie du glucose sanguin, et de faciliter l'assimilation des nutriments grâce à la sécrétion des sucs pancréatiques. L'ensemble de ces études sur la différenciation pancréatique montre le niveau de complexité des cascades génétiques (Fig. 6). Cependant, des études récentes portant sur la différenciation de cellules souches en cellules pancréatiques endocrines, et basées sur l'ensemble des études faites sur le développement du pancréas mettent en évidence le manque de connaissances pour développer des cellules sécrétant l'insuline en fonction de leur environnement (D'Amour et al., 2006). Ainsi, bien que depuis 15 ans notre connaissance sur le développement du pancréas endocrine a fortement progressé, d'autres études seront nécessaires pour définir des thérapies cellulaires pour le diabète de type 1.

D) La famille des gènes à domaine SERTA

La famille des gènes à domaine SERTA contient cinq gènes chez la souris et l'homme. Le premier gène à avoir été caractérisé est le gène *Sertad1* (Sugimoto et al., 1999). Les autres gènes de la famille à domaine SERTA chez les mammifères sont : *Sertad2*, *Sertad3*, *Cdca4* et *Sertad4* (Hsu et al., 2001 ; Cho et al., 2000 ; Abdullah et al., 2001 ; Bennetts et al., 2006). Un homologue a également été caractérisé chez la drosophile : *Tara* (Calgaro et al., 2002)

En plus du domaine SERTA présent dans l'ensemble de ces protéines, un domaine de fixation à la cycline A est trouvé sur l'ensemble de ces protéines murines, humaines et de drosophile. Un domaine d'interaction avec les protéines à PHD bromodomaine, ainsi qu'un

domaine en C-terminal sont conservés chez l'ensemble de ces protéines à l'exception de la protéine SERTAD4 (Fig. 7 ; Bennetts et al., 2006).

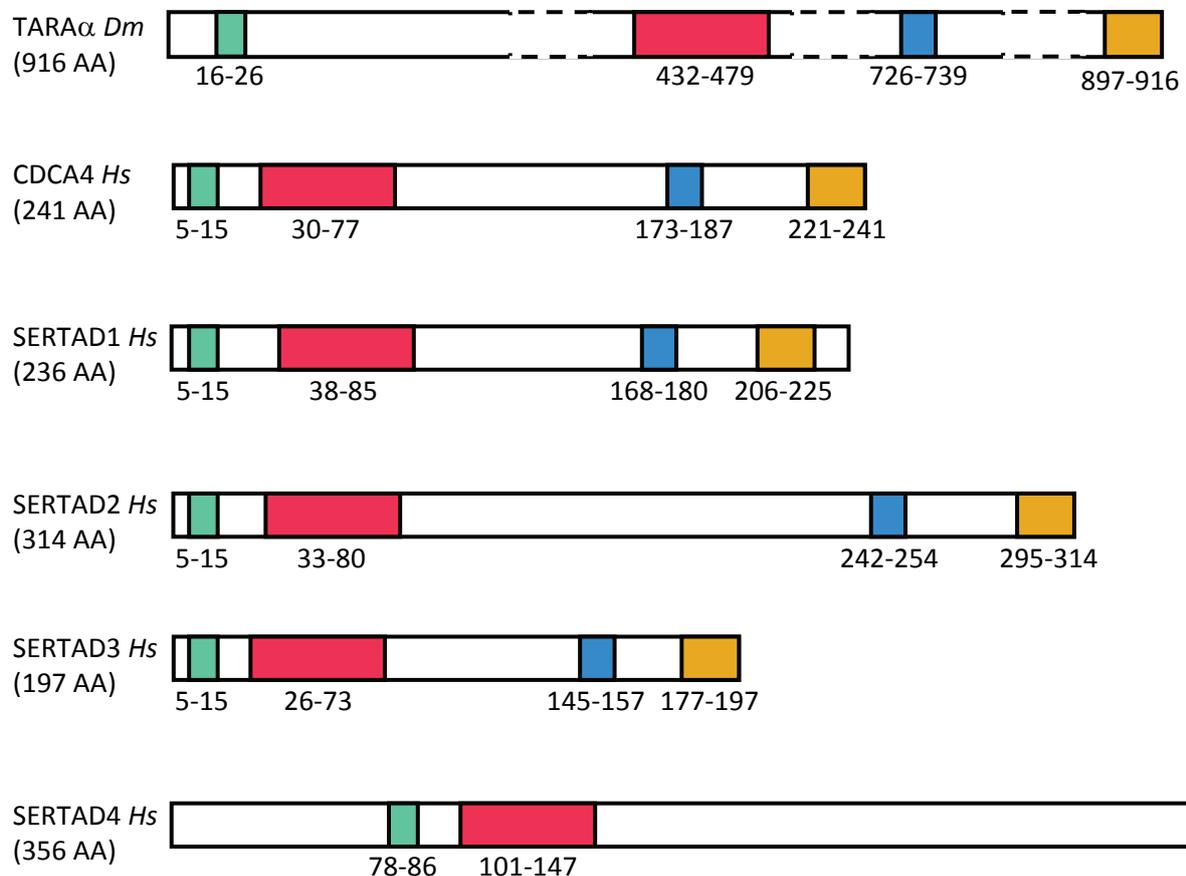


Figure 7 : Représentation des protéines à domaine SERTA humaines et de drosophile. Les protéines de la famille SERTA sont caractérisées par la présence du domaine SERTA (■). Elles comportent également toutes un domaine de fixation à la cycline A (■). A l'exception de SERTAD4, elles comportent également un domaine d'interaction avec les protéines à PHD bromodomaine (■) et un domaine C-terminal conservé (■). Les références des séquences sont : TARAA *Dm* NP_524994, CDCA4 *Hs* NP_060425, SERTAD1 *Hs* NP_037508, SERTAD2 *Hs* NP_055570, SERTAD3 *Hs* NP_037500, SERTAD4 *Hs* NP_062551 (adaptée de Calgaro et al., 2002; Bennetts et al., 2006).

Bien que la fonction de ces gènes soit encore mal définie au cours du développement embryonnaire, *Sertad1*, *Sertad2* et *Sertad3* ont été impliqués dans la régulation du cycle cellulaire. En effet, la perte de fonction de *Sertad1* et *Sertad2* dans des cultures cellulaires induit une réduction de la croissance cellulaire et une réduction du nombre de colonies cellulaires formées (Sugimoto et al., 1999 ; Hsu et al., 2001 ; Sim et al., 2006). L'activité normale de SERTAD1 et SERTAD2 passe par l'activation du complexe Cdk4/Cycline D, qui à

son tour va induire l'expression de la Cycline E. Ensuite en complexe avec Cdk2, la Cycline E permettra la progression du cycle cellulaire (Sim et al., 2006). Toutefois, Watanabe-Fukunaga et al. ont mis en évidence que SERTAD1 serait également impliqué dans l'inhibition de la croissance cellulaire de manière dépendante ou indépendante de P53 (Watanabe-Fukunaga et al., 2005). L'ensemble de ces résultats montre le rôle complexe des protéines à domaine SERTA, et nécessitera de nouvelles études pour déterminer le rôle de *Sertad4* et *Cdca4* dont les fonctions n'ont pas été étudiées.

En parallèle de leur implication dans la régulation du cycle cellulaire, l'expression de gènes de la famille à SERTA domaine a été caractérisée comme altérée dans de nombreux cancers (Xi et al., 2006; Cheong et al., 2009). SERTAD1 dans le cancer colorectal est un bon facteur de pronostic étant donné que son expression décroît dans les tumeurs colorectales (Xi et al., 2008). A l'inverse, *Sertad2* est surexprimé dans de nombreux cancers et son expression est liée à un mauvais pronostic dans les adénocarcinomes hépatiques. De plus, il a été montré que sa simple inhibition dans des lignées cellulaires de carcinome colorectal est suffisante pour inhiber la croissance cellulaire (Cheong et al., 2009).

Récemment, *Tara*, le gène à domaine SERTA chez la drosophile a été identifié comme interagissant avec la voie Notch. De plus, la perte de fonction de *Tara* montre une malformation des ailes similaire à celle observée chez certains mutants pour la Cycline E, qui a également été caractérisée comme interagissant avec la voie Notch (Kankel et al., 2007).

E) La famille des gènes *Rfx*

La famille de gènes *Rfx* (Regulatory factor X) est une famille de 7 gènes (Aftab et al., 2008). Elle a été pour la première fois décrite il y a 20 ans par la propriété de ces protéines à se fixer sur le domaine X de l'ADN, une séquence de 14 nucléotides (Reith et al., 1988 ; Dorn et al., 1987). L'ensemble de ces facteurs de transcription sont caractérisés par la présence d'un domaine de fixation au domaine X (Reith et al., 1990, Fig. 8). Ce domaine de fixation à l'ADN est de type « winged helix » (Gajiwala et al., 2000a; Gajiwala et al., 2000b). Chez les protéines, RFX-1, -2, -3, -4 et -7, on retrouve également trois autres domaines conservés : un domaine de dimérisation ou domaine D, un domaine B et un domaine C (Emery et al., 1996). Ces trois domaines participent à la dimérisation de ces protéines (Katan-Khaykovich et al.,

1998). Toutefois, il est à noter que les protéines RFX sont fonctionnelles à l'état de monomère. Finalement chez les protéines RFX-1, -2 et -3, on retrouve également un domaine d'activation ou domaine A (Fig. 8 ; Emery et al., 1996).

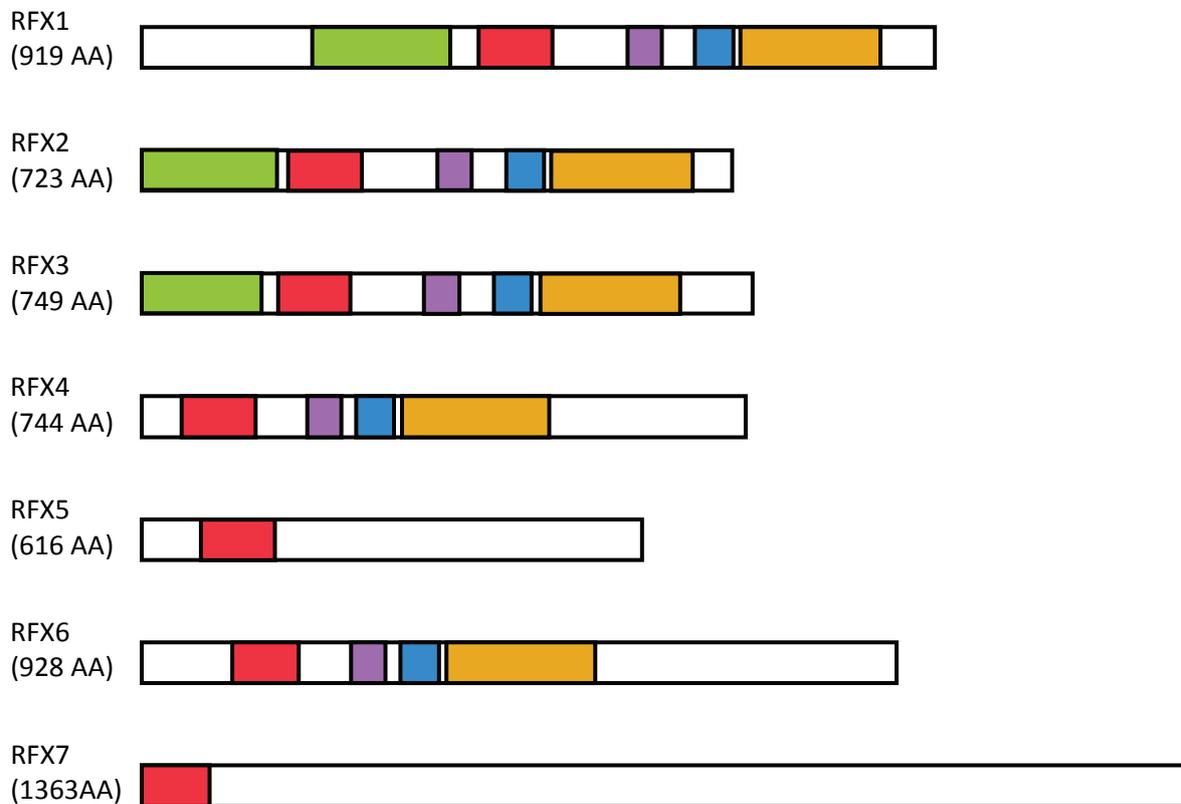


Figure 8 : Représentation des protéines humaines de la famille RFX. Les protéines de la famille RFX sont caractérisées par la présence d'un domaine de fixation à l'ADN (■). Les protéines RFX-1, -2, -3, -4 et -6 sont également caractérisées par la présence d'un domaine de dimérisation (■) et de deux domaines (B, ■ et C, ■) participant à la dimérisation des Rfx. Les protéines RFX-1, -2, -3, contiennent également un domaine d'activation (A, ■). Les références des séquences sont ENSP00000254325 pour RFX1, ENSP00000306335 pour RFX2, ENSP00000371434 pour RFX3, ENSP00000350552 pour RFX4, ENSP00000357864 pour rFX5, ENSP00000332208 pour RFX6 et ENSP00000313299 pour RFX7 (adaptée de Aftab et al., 2008).

1. *Rfx1*

Rfx1 a été le premier gène de la famille *Rfx* à être décrit (Reith et al., 1990). *Rfx1* a été caractérisé pour son action sur l'expression des gènes du complexe majeur d'histocompatibilité de classe II (Reith et al., 1990). Il a ensuite été caractérisé comme étant nécessaire à l'expression des gènes du virus de l'hépatite B (Siegrist et al., 1993). Récemment, *Rfx1* a été montré intervenant dans la réponse aux lésions de l'ADN (Lubelsky

et al., 2005). Par cet aspect, la fonction de *Rfx1* serait similaire à *Crt1*, un gène de la famille *Rfx* chez *S. cerevisiae*. *Crt1* a en effet été impliqué dans la régulation de la réplication de l'ADN suite aux dommages subis par l'ADN (Huang et al., 1998). Finalement, la perte de fonction de *Rfx1* a été étudiée chez la souris. Il a ainsi été montré que *Rfx1* est requis très tôt au cours de l'embryogénèse puisque les mutants *Rfx1* meurent à E2,5 (Feng et al., 2009).

2. *Rfx2*

Rfx2 est un des gènes de la famille *Rfx* le moins étudié depuis sa caractérisation (Reith et al., 1994). Seule une fonction lui a été directement attribuée. Il a ainsi été caractérisé comme étant principalement exprimé dans les testicules, et plus spécifiquement dans les spermatocytes et les spermatides (Reith et al., 1994 ; Horvath et al., 2004). Une étude sur le développement des testicules a montré que *Rfx2* se fixe sur le promoteur de l'histone H1t, une variante de l'Histone H1, suggérant que *Rfx2* régule l'expression de H1t (Horvath et al., 2004). Une récente étude sur l'expression de gènes exprimés dans les tissus fortement ciliés tels que l'épithélium olfactif, la trachée et les poumons, a montré que *Rfx2* serait exprimé dans les tissus ciliés (McClintock et al., 2008).

3. *Rfx3*

Comme suggéré pour *Rfx2*, *Rfx3* est exprimé dans de nombreux tissus ciliés. Il a ainsi été montré que la perte de fonction de *Rfx3* induit un défaut d'asymétrie droite gauche des organes (Bonnafe et al., 2004). Bonnafe et al. ont mis en évidence que ce défaut était dû à une malformation des cils du nœud embryonnaire, qui sont indispensables pour le placement des organes chez l'individu (Bonnafe et al., 2004). Une autre étude a mis en évidence que l'expression de *Rfx3* est spécifiquement perdue chez le mutant *Noto*, chez lequel les cils du nœud embryonnaire sont malformés et ne sont plus fonctionnels (Beckers et al., 2007).

La perte de fonction de *Rfx3* a également été liée à une réduction du nombre de cils dans l'épendyme des ventricules cérébraux, bien que la structure des cils ne semble pas affectée. Ce défaut des cils dans le cerveau en développement conduit à une hydrocéphalie

corrélée avec une malformation des plexus choroïdes et une agénésie de l'organe subcommissural (Baas et al., 2006).

Finalement, *Rfx3* a été récemment impliqué dans la différenciation du pancréas endocrine. Il a ainsi été montré chez les souris mutantes pour *Rfx3*, que la différenciation du pancréas endocrine est réduite avec une réduction importante du nombre de cellules α et β , et une augmentation du nombre de cellules PP (Ait-Lounis et al., 2007). En parallèle de cette réduction de la différenciation du pancréas endocrine, l'observation des cils sur les cellules endocrines ont montré une réduction de la taille des cils et de leur nombre dans les îlots de Langerhans (Ait-Lounis et al., 2007). Cependant, des travaux du groupe de M. Hebrok montrent que la perte des cils au cours du développement du pancréas endocrine ou dans les cellules β n'induit ni malformation pancréatique, ni problème fonctionnel (Cano et al., 2006). Ces résultats suggèrent donc que la réduction de la différenciation du pancréas endocrine observée chez les mutants *Rfx3* n'est pas liée à un défaut des cils.

4. *Rfx4*

Rfx4 a la particularité de présenter six isoformes différentes (Matsushita et al., 2005). Trois d'entre elles sont exprimées dans les testicules, une (*rfx4 v3*) est exprimée dans le cerveau et deux autres ont été observées dans des tumeurs gliales (Matsushita et al., 2005).

L'étude de la perte de fonction de *Rfx4 v3* a mis en évidence que *rfx4 v3* est nécessaire au développement du cerveau. En effet, Blackshear et al. ont montré que la perte de fonction de *Rfx4 v3* induisait une hydrocéphalie comme pour le mutant pour *Rfx3* (Blackshear et al., 2003). Comme chez la souris mutée pour *Rfx3*, l'organe subcommissural n'est pas développé, affectant la circulation du fluide cérébrospinal. Il a également été montré que la perte de fonction de *Rfx4 v3* induit une malformation des cils de l'épendyme bien que l'agénésie de l'organe subcommissural semble être le principal responsable du phénotype du mutant *Rfx4 v3* (Zhang et al., 2007). *Rfx4* a également été impliqué dans la formation du cortex cérébral (Zarbalis et al., 2004). Dans le cerveau, RFX4 a finalement été observé dans les noyaux suprachiasmatiques, l'horloge circadienne du cerveau ; où son expression est spécifiquement induite par la lumière (Araki et al., 2004)

Comme évoqué précédemment, *Rfx4* a été à l'origine caractérisé comme étant exprimé dans les testicules (Morotomi-Yano et al., 2002). Récemment, il a été montré que *Rfx4* est spécifiquement exprimé dans les spermatoocytes dans les testicules, comme *Rfx2*. De plus, il a été montré que RFX4 se fixe directement sur le promoteur de l'histone H1t, et peut interagir avec RFX2 en formant un dimère (Wolfe et al., 2006; Wolfe et al., 2008). Ces résultats suggèrent donc que RFX4 et RFX2 interagissent afin de réguler l'expression de l'histone H1t.

5. *Rfx5*

Rfx5 a été le plus étudié de tous les gènes *Rfx*. Ceci est dû à son implication dans le syndrome des lymphocytes dénudés (*bare*), qui est une maladie immunitaire héréditaire monogénique autosomale récessive. Elle est caractérisée par l'absence d'expression du complexe majeur d'histocompatibilité (CMH) de classe II. Par conséquent, les personnes atteintes souffrent d'infections récurrentes, conduisant fréquemment à la mort pendant l'enfance. Le développement du syndrome des lymphocytes dénudés (*bare*) est dû à la perte d'expression d'un des quatre gènes formant le complexe permettant l'expression du CMH de classe II dont *Rfx5* (Reith et al., 2001).

En plus de ce premier rôle, *Rfx5* a été impliqué dans la régulation de l'expression du collagène (Sengupta et al., 2002). Il a ainsi été montré que RFX5 en complexe avec l'histone deacétylase 2 et un répresseur transcriptionnel inhibe l'expression du collagène $\alpha 2(1)$ (Xu et al., 2006).

Il faut également noter que RFX5 diffère des autres protéines *Rfx* puisqu'elle ne contient que le domaine de fixation à l'ADN et aucun autre domaine conservé (Emery et al., 1996).

6. *Rfx6* et *Rfx7*

Rfx6 et *Rfx7* sont les deux derniers gènes à avoir été caractérisés par une approche bioinformatique (Aftab et al., 2008). D'après les travaux d'Aftab et al., les séquences de *Rfx6*

et *Rfx7* sont similaires à celle de *Rfx4* et *Rfx5* respectivement. Cependant, aucune étude fonctionnelle n'a permis de déterminer leur fonction.

Les gènes de la famille *Rfx* sont donc impliqués dans différents processus au cours du développement. On peut y retenir leur implication dans l'expression du CMH de classe II pour lequel ils ont été initialement décrits. On peut aussi remarquer que plusieurs d'entre eux sont impliqués dans des processus développementaux, notamment dans le cerveau et les testicules. On peut aussi noter que la perte de fonction de certains gènes *Rfx* est liée à une malformation ou à une réduction du nombre des cils dans différents organes dont le pancréas. Pour ce dernier aspect, on peut aussi remarquer que la perte de fonction de *dRfx* et de *Daf19*, les homologues respectifs de la famille des gènes *Rfx* chez la drosophile et chez *Caenorhabditis elegans*, induit une malformation des cils sur les neurones senseurs (Dubruille et al., 2002 ; Swoboda et al., 2000).

OBJECTIFS

Le pancréas, qui se développe à partir de l'endoderme est constitué des îlots de Langerhans, des cellules acinaires et canalaies. Les îlots, agrégats cellulaires de forme sphérique, sont principalement constitués de quatre types cellulaires endocriniens : les cellules α , β , δ et PP, sécrétant respectivement les hormones glucagon, insuline, somatostatine et polypeptide pancréatique. Ces îlots ont pour rôle principal la régulation de la glycémie. La destruction des cellules β par réaction auto-immune est à l'origine du diabète de type 1. Afin de pallier aux effets secondaires dus à l'injection quotidienne d'insuline chez les diabétiques et au manque de greffons de pancréas, le développement de cellules β à partir de cellules souches est aujourd'hui étudié. Nos connaissances sur la différenciation de l'endoderme et du pancréas endocrine permettent ainsi de développer *in vitro* à partir de cellules souches, des cellules sécrétant l'insuline. Cependant ces cellules ne peuvent réguler leur sécrétion en fonction du taux de glucose. Mes travaux de thèse se sont implantés dans cette problématique afin de déterminer les cascades génétiques de différenciation des cellules pancréatiques endocrines, dont celle nécessaire à la différenciation des cellules sécrétant l'insuline de façon régulée.

Au cours du développement embryonnaire, le facteur de transcription à domaine bHLH NEUROG3 est requis pour la différenciation des quatre principaux sous types endocriniens (α , β , δ et PP). Il a en effet été montré que l'expression du facteur de transcription proendocrine NEUROG3 est nécessaire et suffisante pour induire la différenciation endocrine. Il a notamment été montré que NEUROG3 est nécessaire à l'expression des facteurs de transcription tels *Arx*, *Insm1*, *NeuroD1*, *Pax4*, *Pax6*, qui sont requis pour une différenciation normale du pancréas endocrine.

Notre hypothèse de recherche est par conséquent que NEUROG3 contrôlerait les programmes génétiques génériques de différenciation communs à l'ensemble des cellules du pancréas endocrine, et les programmes génétiques de différenciation spécifique de chacun des sous types cellulaires pancréatiques endocriniens. Cependant, malgré le rôle essentiel de *Neurog3* au cours de la différenciation du pancréas endocrine, seul un gène cible de NEUROG3 était connu au début de mon doctorat : *NeuroD1*. Mais seul, ce gène qui est impliqué dans la différenciation de l'ensemble des sous types cellulaires pancréatiques endocriniens, ne permet pas d'expliquer les différentes cascades génétiques induites par NEUROG3 et impliquées dans la différenciation endocrine pancréatique.

L'objectif de mon doctorat a donc été de caractériser le transcriptome pancréatique induit par NEUROG3 au cours de la différenciation pancréatique endocrine et d'étudier l'expression et la fonction des potentiels nouveaux gènes cibles de NEUROG3 mis en évidence, gènes qui pourraient être nécessaires à la différenciation de cellules souches en cellules β .

RESULTATS

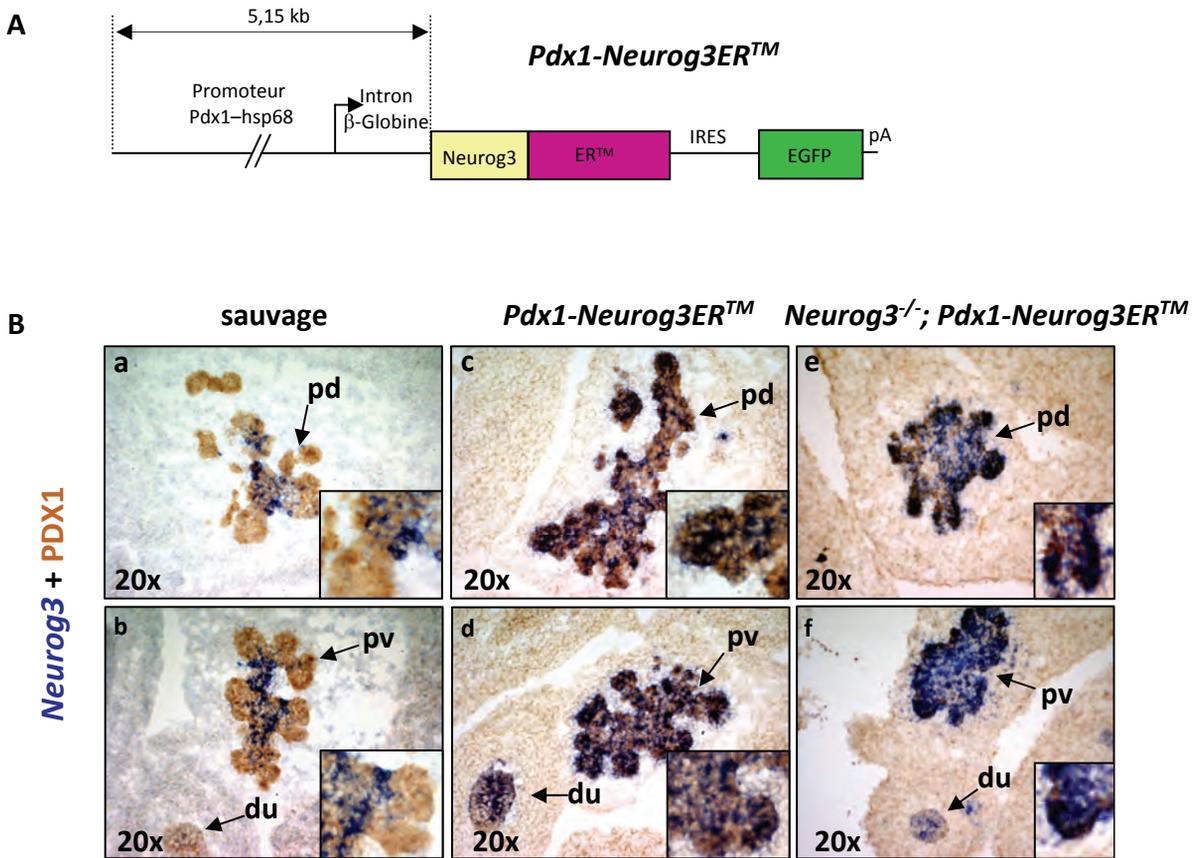


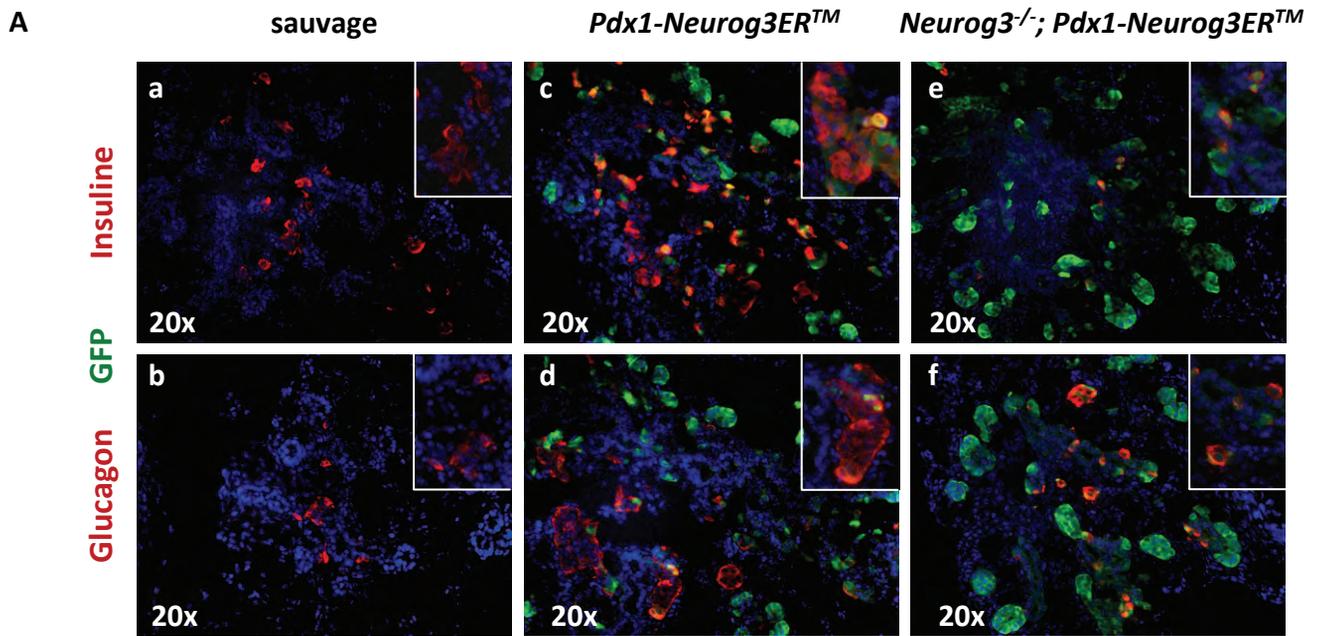
Figure 9 : Caractérisation de l'expression de *Neurog3* contrôlée par les séquences régulatrices du gène *Pdx1* dans la lignée de souris transgénique *Pdx1-Neurog3ERTM*. (A) Représentation schématique du transgène. Un fragment de 5.15 kb contenant les séquences génomiques régulatrices de l'expression tissu spécifique de *Pdx1*, le promoteur minimal du gène *Hsp68* et l'intron du gène de la β -globine ont été clonés en amont d'un gène de fusion entre la phase codante de *Neurog3* et du domaine de liaison des ligands du récepteur des œstrogènes, modifié pour reconnaître spécifiquement le 4-hydroxytamoxifène (*ERTM*). L'activité transcriptionnelle de la protéine résultante *NEUROG3ERTM* est dépendante de l'administration de tamoxifène. La séquence IRES-EGFP permet une expression de la GFP sous le contrôle du promoteur de *Pdx1*. EGFP, Enhanced Green Fluorescent Protein; IRES, Internal Ribosomal Entry Site; pA, signaux de polyAdenylation. (B) Hybridation *in situ* pour *Neurog3* (bleu) et immuno-histochimie pour la protéine PDX1 (brun) sur des sections d'embryons sauvages (a,b), transgéniques *Pdx1-Neurog3ERTM* (c,d), mutants pour *Neurog3* et transgéniques *Pdx1-Neurog3ERTM* (e,f) à E12.5. Dans les embryons transgéniques et transgéniques mutants, l'expression de *Neurog3* est plus large et reflète l'expression de PDX1. Pv, pancréas ventral; pd, pancréas dorsal; du, duodénum; inserts : 40x.

A) Caractérisation de la souris transgénique *Pdx1-Neurog3ERTM*

1. La lignée de souris transgénique *Pdx1-Neurog3ERTM*

Afin de déterminer de nouveaux gènes cibles de NEUROG3, nous avons utilisé la souris transgénique *Pdx1-Neurog3ERTM* (Fig. 9A), fournie par le Dr Anne Grapin-Botton (EPFL, Lausanne ; Johansson et al., 2007). Précédemment, des travaux ont montré que l'expression ectopique de NEUROG3 selon le patron d'expression de *Pdx1* induit une différenciation précoce des progéniteurs pancréatiques principalement en cellules α dès E10,5 (Apelqvist et al., 1999 ; Schwitzgebel et al., 2000). Une étude récente a également montré que les progéniteurs pancréatiques ont différentes périodes de compétence pour se différencier dans les différents sous types endocriniens (Johansson et al., 2007). Il a ainsi été montré que les progéniteurs pancréatiques acquièrent la compétence de se différencier en cellules β à partir de E12,5. Par conséquent, afin d'induire l'activité de NEUROG3 dans une période de compétence des progéniteurs endocrine à se différencier en cellules β , un ADNc codant pour la protéine de fusion NEUROG3ERTM a été cloné en aval des séquences régulatrices de l'expression spatio-temporelle du facteur de transcription *Pdx1* (Fig. 9A ; Gannon et al., 2001). La partie ERTM est constituée du domaine de liaison du ligand du récepteur des œstrogènes, modifié pour reconnaître spécifiquement le 4-hydroxytamoxifène (Hayashi et al., 2002). En absence de 4-hydroxytamoxifène, la protéine de fusion est normalement séquestrée dans le cytoplasme, suite à l'interaction avec la protéine HSP90. L'administration de 4-hydroxytamoxifène a pour conséquence d'entraîner la libération de la protéine de fusion NEUROG3ERTM d'HSP90 entraînant sa translocation nucléaire, induisant ainsi son activité transcriptionnelle (Picard, 1994). La construction est également bicistronique. L'expression du gène EGFP est ainsi régulée comme celle de Neurog3ERTM, ce qui permet de suivre les cellules exprimant le transgène.

Afin d'éviter que la protéine endogène NEUROG3 exprimée dès E9,5 induise la différenciation des progéniteurs pancréatiques en cellules endocrines, les souris transgéniques ont également été croisées sur un fond *Neurog3* mutant, qui est caractérisé par l'absence de cellules pancréatiques endocrines. Par conséquent, les phénotypes observés chez les souris mutantes transgéniques (*Neurog3^{-/-} ; Pdx1-Neurog3ERTM*) seront dus



B

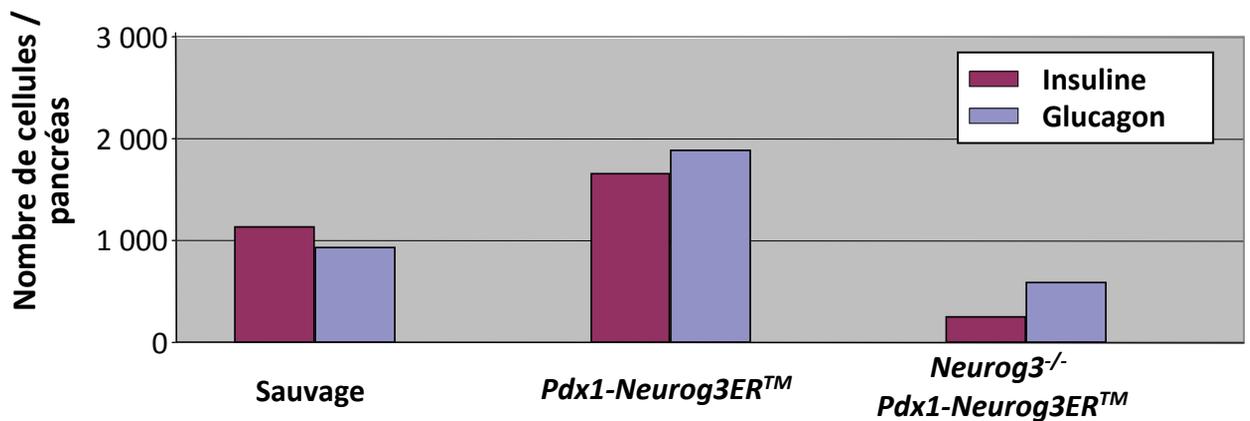


Figure 10: Le transgène *Pdx1-Neurog3ERTM* restaure les lignages α et β dans le mutant *Neurog3*. (A) Immunofluorescence pour l'insuline et la GFP (a, c, e) et, le glucagon et la GFP (b,d, f) sur des sections d'embryon sauvage (a, b), transgénique (c, d) et transgénique mutant (e, f) à E14,5, 2 jours après l'injection de 4-hydroxytamoxifène ; inserts : 40x. (B) 2 jours après l'injection de 4-hydroxytamoxifène, quantification du nombre de cellules exprimant l'insuline ou le glucagon chez les embryons sauvage, transgénique et transgénique mutant à E14,5. Chez les souris transgéniques, le nombre de cellules insuline et glucagon sont respectivement 1,5 fois et 2 fois plus nombreuses par rapport à un embryon sauvage. Chez le mutant *Neurog3*, l'activation de la protéine de fusion a permis de restaurer la différenciation des cellules α et β .

à l'activation de la protéine de fusion NEUROG3ERTM, indépendamment de l'expression de la forme endogène de NEUROG3.

2. Expression ectopique de *Neurog3* dans les cellules épithéliales exprimant PDX1 à E12,5

La première étape de ce travail a été de vérifier que le transgène permettait une expression de l'ARNm du gène *Neurog3ERTM* selon le patron d'expression de *Pdx1*. Pour cela, une hybridation *in situ* (ISH) pour *Neurog3* suivie d'une immuno-histochimie pour PDX1 ont été réalisées sur des sections d'embryons à E12,5 sauvages, transgéniques et, mutants pour *Neurog3* et transgéniques (Fig. 9B). Dans le pancréas dorsal et ventral, ainsi que dans le duodénum, le nombre de cellules exprimant *Neurog3* est très nettement supérieur chez la souris transgénique (Fig. 9Bc-d) par rapport à la souris sauvage (Fig. 9Ba-b) ; d'autre part, les cellules exprimant *Neurog3* couvrent presque complètement le territoire d'expression de PDX1. Les souris mutantes et transgéniques (Fig. 9Be-f) expriment également *Neurog3*, selon le patron d'expression de PDX1. Cette lignée de souris transgénique sous fond *Neurog3* mutant permet donc une expression ectopique de *Neurog3ERTM* dans les cellules de l'épithélium pancréatique exprimant PDX1 à E12,5.

3. Le transgène Pdx1-Neurog3ERTM restaure les lignages α et β dans le mutant *Neurog3*

L'expression ectopique de *Neurog3ERTM* chez les souris *Neurog3^{-/-}* ; *Pdx1-Neurog3ERTM* devrait nous permettre de restaurer les programmes de différenciation nécessaires au développement des cellules pancréatiques endocrines; programmes qui sont induit par NEUROG3 (Grapin-Botton et al., 2001) et qui sont perdus chez les mutants *Neurog3* entraînant une perte totale de la différenciation du pancréas endocrine (Gradwohl et al., 2000).

Afin d'induire la différenciation des cellules β sous l'effet de la protéine de fusion active, nous avons choisis d'activer la protéine de fusion par une injection de 2 mg de 4-hydroxytamoxifène à E12,5. 2 jours après l'injection, j'ai déterminé le nombre de cellules α et β chez des embryons de souris sauvage, transgénique et, mutante transgénique par immunofluorescence (Fig. 10 A).

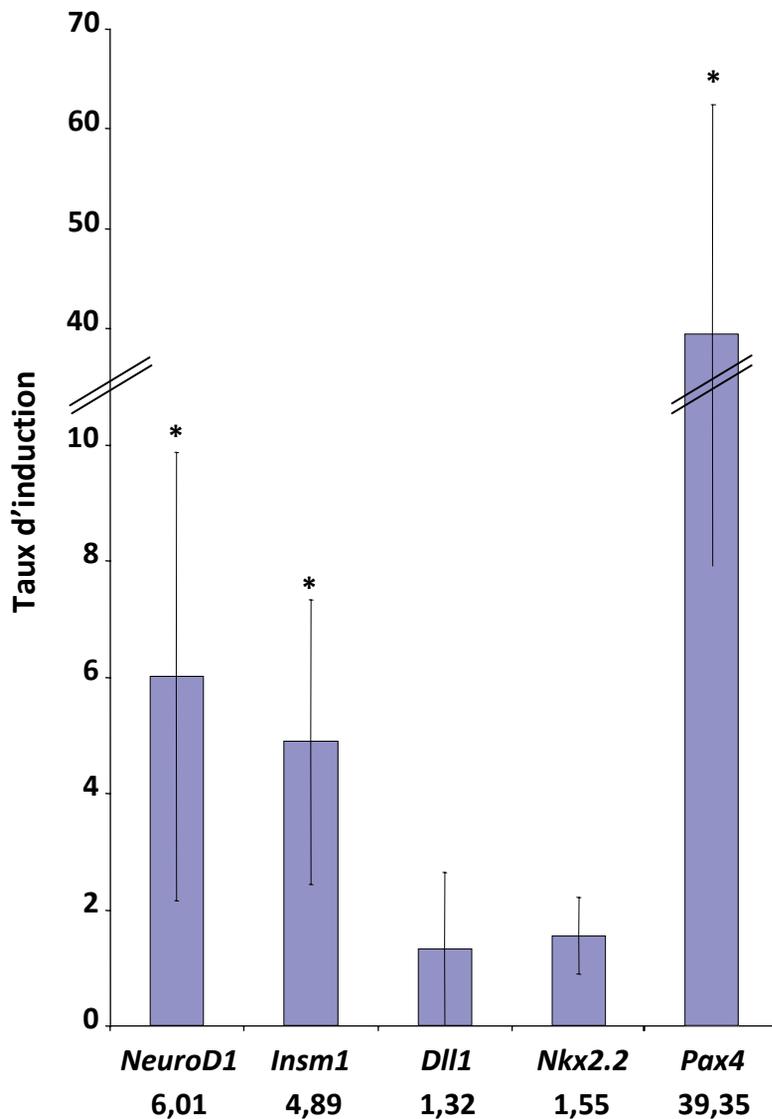


Figure 11 : L'activation de NEUROG3ERTM conduit à l'induction des gènes impliqués dans la différenciation des cellules endocrines. Analyse par RT-qPCR de l'induction de l'expression de *Neurod1*, *Insm1*, *Dll1*, *Nkx2.2* et *Pax4* à E13,5 chez des embryons transgéniques mutants injectés avec du 4-hydroxytamoxifène à E12,5 par rapport à des embryons transgéniques mutants non injectés. L'activation de NEUROG3ERTM induit l'expression des gènes cibles de NEUROG3 (*NeuroD1* et *Insm1*) et de *Pax4* (gène impliqué dans la différenciation précoce des cellules β). *Dll1* et *Nkx2.2*, qui ne sont pas impliqués dans la différenciation précoce des cellules β , ne sont pas induits par l'activation de la protéine de fusion. * : $p < 0,02$; $n = 4$.

Chez les embryons transgéniques, l'injection de 4-hydroxytamoxifène a induit une augmentation de la différenciation des cellules sécrétant l'insuline et le glucagon de 1,5 fois et 2 fois respectivement (Fig. 10 B). Cette augmentation de la différenciation endocrine montre donc que l'activation de la protéine de fusion NEUROG3ERTM permet d'induire une différenciation ectopique des cellules pancréatiques endocrines. L'analyse de la différenciation du pancréas endocrine chez les mutants *Neurog3*, qui sont caractérisés par une perte totale de la différenciation du pancréas endocrine, montre que la protéine NEUROG3ERTM active permet d'induire un programme de différenciation du pancréas endocrine et par conséquent de restaurer 23% et 63% de la différenciation du lignage β et α respectivement par rapport à un embryon sauvage au même stade de développement (Fig 10 B).

4. L'activation de Neurog3ERTM conduit à l'induction des gènes impliqués dans la différenciation des cellules endocrines

Dans cet objectif, afin de valider ce système expérimental pour l'identification de nouveaux gènes cibles de NEUROG3, nous avons ensuite déterminé si la différenciation des lignages α et β chez les souris mutantes transgéniques se fait via l'induction d'un programme génétique normal. Pour cela, nous avons étudié l'expression de gènes déjà impliqués dans la différenciation du pancréas endocrine : *NeuroD1* (Naya et al., 1997), *Insm1* (Mellitzer et al., 2006), *Pax4* (Sosa-Pineda et al., 1997), *Nkx2.2* (Sussel et al., 1998) et *Dll1* (Apelqvist et al., 1999). Dans cet objectif, j'ai réalisé des RT-qPCR sur des ARN de pancréas d'embryons mutants transgéniques à E13,5, injectés ou non avec 2 mg de 4-hydroxytamoxifène à E12,5 (Fig. 11).

L'expression d'*Insm1*, *NeuroD1* et *Pax4* sont induites respectivement 5 fois, 6 fois et 39 fois dans les pancréas des embryons transgéniques mutants injectés avec le 4-hydroxytamoxifène par rapport aux embryons non injectés (Fig. 11). En revanche, l'expression de *Dll1* et *Nkx2.2* ne sont pas induites (Fig. 11). Ces différences peuvent s'expliquer par les différentes fonctions de ces gènes. *NeuroD1* et *Insm1* sont deux gènes cibles de NEUROG3, qui sont impliqués dans la différenciation de l'ensemble des sous types pancréatiques endocriniens ; et *Pax4* est impliqué dans l'engagement des progéniteurs endocrines lors de la différenciation en cellules β ou δ (Naya et al., 1997; Mellitzer et al.,

Nom de la sonde Affymetrix	Référence UniGene	Nom du gène	Symbole du gène	Taux moyen d'induction	pValue
Facteur de transcription, et régulation transcriptionnelle					
1426438_at	Mm.476906	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, Y-linked	Ddx3y	125,07	0,38
1457945_at	Mm.218024.1	Eukaryotic translation initiation factor 2, subunit 3, structural gene Y-linked	Eif2s3y	77,86	0,32
1450042_at	Mm.275547	aristaless related homeobox gene (Drosophila)	Arx	11,42	0,00
1419271_at	Mm.3608.1	paired box gene 6	Pax6	5,01	0,00
1426412_at	Mm.4636.1	neurogenic differentiation 1	Neurod1	3,53	0,00
1455865_at	Mm.379070	insulinoma-associated 1	Insm1	3,03	0,08
1426413_at	Mm.4636.1	neurogenic differentiation 1	Neurod1	2,95	0,00
1457613_at	Mm.132545.1	Regulatory factor X domain containing 1	Rfxdc1	2,12	0,03
1425886_at	Mm.150496.1	FEV (ETS oncogene family)	Fev	2,03	0,07
1438042_at	Mm.130271.1	short stature homeobox 2	Shox2	1,52	0,11
hormone, et sécrétion					
1424865_at	Mm.46248.1	peptide YY	Pyy	6,59	0,00
1418149_at	Mm.4137.1	chromogranin A	Chga	6,42	0,00
1449908_at	Mm.5115.1	gastric inhibitory polypeptide	Gip	5,45	0,00
1435297_at	Mm.188875.1	gap junction membrane channel protein alpha 9	Gja9	3,04	0,01
1423510_at	Mm.415.1	islet amyloid polypeptide	Iapp	2,77	0,02
1448312_at	Mm.1247.1	proprotein convertase subtilisin/kexin type 2	Pcsk2	2,49	0,00
1450389_s_at	Mm.1945.1	phosphatidylinositol-4-phosphate 5-kinase, type 1 alpha	Pip5k1a	2,06	0,24
1450708_at	Mm.5038.1	secretogranin II	Scg2	1,83	0,01
1433947_at	Mm.44462.1	RAB37, member of RAS oncogene family	Rab37	1,75	0,13
1460081_at	Mm.41354.1	Synaptotagmin VII (Syt7), transcript variant 1, mRNA	Syt7	1,74	0,26
1457969_at	Mm.42330.1	RAB interacting factor (Rabif), mRNA	Rabif	1,64	0,19
Cytosquelette et motilité					
1450276_a_at	Mm.2416.1	scinderin	Scin	1,83	0,08
1425558_at	Mm.213403.1	kinesin light chain 3	Klc3	1,60	0,16
1451684_a_at	Mm.132482.1	bicaudal D homolog 1 (Drosophila)	Bicd1	1,56	0,49
Autres fonctions					
1439607_at	Mm.285366	solute carrier family 7 (cationic amino acid transporter, y+ system), member 14	Slc7a14	4,69	0,00
1456068_at	Mm.37606.1	neurofascin	Nfasc	3,17	0,11
1449907_at	Mm.117119.1	beta-carotene 15,15'-monooxygenase	Bcmo1	2,84	0,11
1424767_at	Mm.17817.1	cadherin 22	Cdh22	2,00	0,13
1418191_at	Mm.27498.1	ubiquitin specific peptidase 18	Usp18	1,81	0,30
1445877_at	Mm.266515.1	Rieske (Fe-S) domain containing (Rfcsd), transcript variant 1	Rfcsd	1,75	0,72
1445565_at	Mm.170587.1	Histone 1, H1e (Hist1h1e), mRNA	Hist1h1e	1,69	0,19

Table 1 : Liste des potentiels gènes cibles de NEUROG3.

Liste des 57 sondes étant : présentes et induites au moins 1,5 fois chez les souris mutantes transgéniques après injection de 4-hydroxytamoxifène par rapport aux souris témoins non injectées; absentes dans le transcriptome des pancréas des souris *Neurog3* mutantes; et enrichies dans le transcriptome dans les progéniteurs pancréatiques endocriniens.

Nom de la sonde Affymetrix	Référence UniGene	Nom du gène	Symbole du gène	Taux moyen d'induction	pValue
Fonction non caractérisée					
1455397_at	Mm.70989.1	CASK interacting protein 1	Caskin1	7,71	0,26
1460244_at	Mm.25557.1	ureidopropionase, beta	Upb1	5,22	0,03
1439752_at	Mm.324008.1	membrane bound O-acyltransferase domain containing 4	Mboat4	4,65	0,13
1456387_at	Mm.26405.1	nucleolar protein 4	Nol4	3,19	0,24
1436100_at	Mm.44613.1	SH2 domain containing 5	Sh2d5	2,77	0,26
1435404_at	Mm.39785.1	dispatched homolog 2 (Drosophila)	Disp2	1,89	0,02
1431856_a_at	Mm.34776.2	C1q and tumor necrosis factor related protein 6	C1qtnf6	1,67	0,17
1433950_at	Mm.41284.1	immunoglobulin superfamily, member 21	Igsf21	1,53	0,10
1447811_s_at	Mm.44915.2	adhesion molecule with Ig like domain 1	Amigo1	1,51	0,20
Est					
1455085_at	Mm.25951.1	RIKEN cDNA 1700086L19 gene	1700086L19Rik	45,98	0,06
1435939_s_at	Mm.4441.4	expressed sequence AI987662	AI987662	5,98	0,00
1432556_a_at	Mm.148740.1	RIKEN cDNA 3100002J23 gene	3100002J23Rik	3,41	0,00
1433727_at	Mm.27909.1	cDNA sequence BC038479	BC038479	3,12	0,09
1456192_x_at	Mm.46095.1	RIKEN cDNA 1700006H02 gene	1700006H02Rik	3,02	0,02
1427509_at	Mm.180687.2	gene model 937, (NCBI)	Gm937	1,97	0,01
1452326_at	Mm.46251.1	RIKEN cDNA 1810010G06 gene	1810010G06Rik	1,91	0,06
1440460_at	Mm.60809.1	RIKEN cDNA 4930504O13 gene	4930504O13Rik	1,88	0,23
1457050_at	Mm.71206.1	RIKEN cDNA 1700129O19 gene	1700129O19Rik	1,80	0,07
1451382_at	Mm.35083.1	RIKEN cDNA 1810008K03 gene	1810008K03Rik	1,71	0,17
1434341_x_at	Mm.22138.2	RIKEN cDNA 1110020P15 gene	1110020P15Rik	1,70	0,15
1436656_at	Mm.99790.1	cDNA sequence BC062109	BC062109	1,68	0,40
1434715_at	Mm.100145.1	RIKEN cDNA 1600014C10 gene	1600014C10Rik	1,58	0,56
1452473_at	Mm.41984.1	RIKEN cDNA E130201N16 gene	E130201N16Rik	1,53	0,72

Table 1 (Suite) : Liste des potentiels gènes cibles de NEUROG3.

Liste des 57 sondes étant: présentes et induites au moins 1,5 fois chez les souris mutantes transgéniques après injection de 4-hydroxytamoxifène par rapport aux souris témoins non injectées; absentes dans le transcriptome des pancréas des souris *Neurog3* mutantes; et enrichies dans le transcriptome dans les progéniteurs pancréatiques endocriniens.

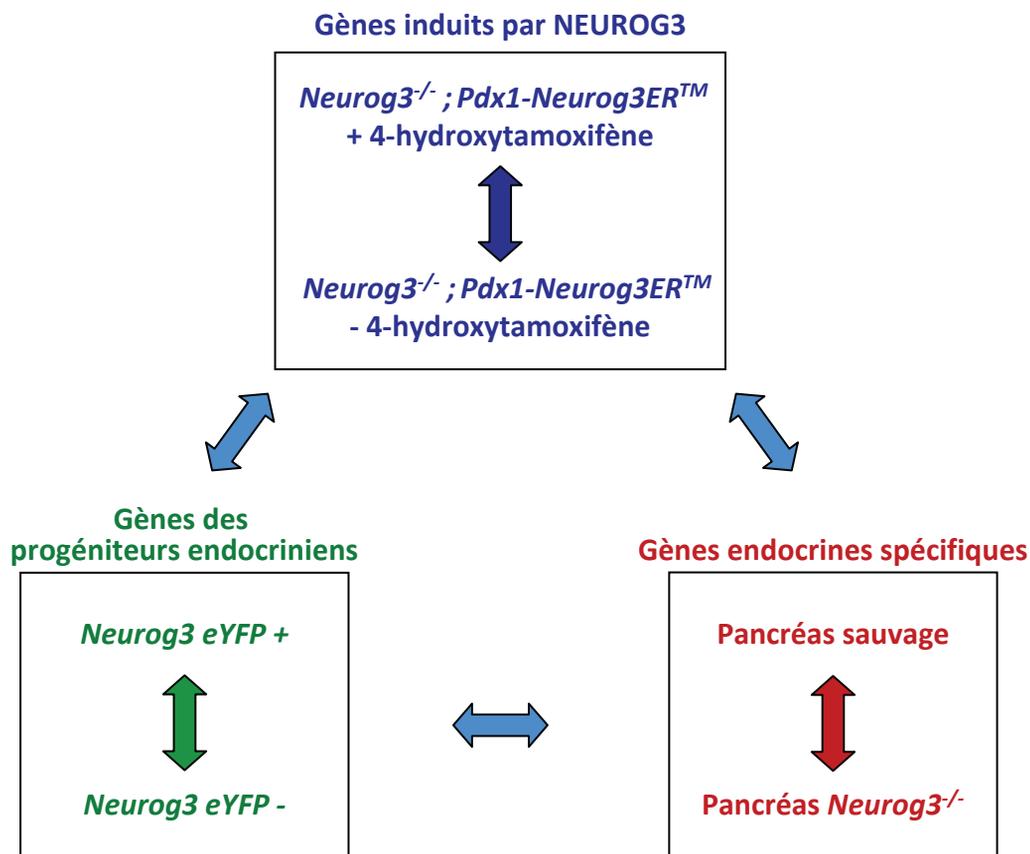
2006 ; Sosa-Pineda et al., 1997). En revanche, *Nkx2.2* est impliquée dans la différenciation finale des cellules β (Sussel et al., 1998) ; et *Dll1* permet de réguler le nombre de progéniteurs pancréatiques, mais n'est pas caractérisé comme un gène cible de NEUROG3 (Apelqvist et al., 1999). Ainsi, 24h après l'activation de la protéine de fusion Neurog3ERTM, seul les gènes cibles directs de NEUROG3 ou les gènes impliqués dans l'engagement de la différenciation des différents sous types endocriniens semblent être induits.

L'ensemble de ces expériences a donc montré que le transgène *Pdx1-Neurog3ERTM* est exprimé dans l'ensemble des cellules exprimant *Pdx1* ; que la protéine de fusion activée par l'injection de 4-hydroxytamoxifène induit l'expression des gènes déjà impliqués dans la différenciation des cellules β ; et que la protéine fusion activée restaure pour la première fois le lignage α et β chez les mutants *Neurog3*. L'ensemble de ces résultats suggère que l'étude du transcriptome pancréatique des embryons *Neurog3* mutants et transgéniques pourrait permettre de définir de nouveaux gènes cibles de NEUROG3 potentiellement impliqués dans la différenciation des cellules pancréatiques endocrines.

B) Identification des potentiels gènes cibles de Neurogenin3

Dans l'objectif de définir de nouveaux gènes cibles de NEUROG3, nous avons caractérisé le transcriptome des pancréas de souris *Neurog3* mutantes transgéniques. Pour cela, 2 mg de 4-hydroxytamoxifène ont été injectés aux souris *Neurog3* mutantes transgéniques à E12,5. Puis l'ARN de 4 pancréas de souris a été extrait à E13,5. L'ADNc issu de ces ARN a été hybridé sur des puces à ADN Affymetrix 430 2.0 couvrant l'ensemble du génome murin, avec 45037 sondes couvrant plus de 34000 gènes et variants. En parallèle, Quatre autres puces ont été hybridées avec des ADNc obtenus à partir d'ARN issus de 4 pancréas de souris mutantes transgéniques à E13,5 non injectées avec le 4-hydroxytamoxifène. Afin de définir une liste de potentiels gènes cibles directs de NEUROG3 impliqués dans la différenciation des cellules β , ces résultats ont ensuite été comparés avec des données de puces à ADN précédemment réalisées au laboratoire en

A



B

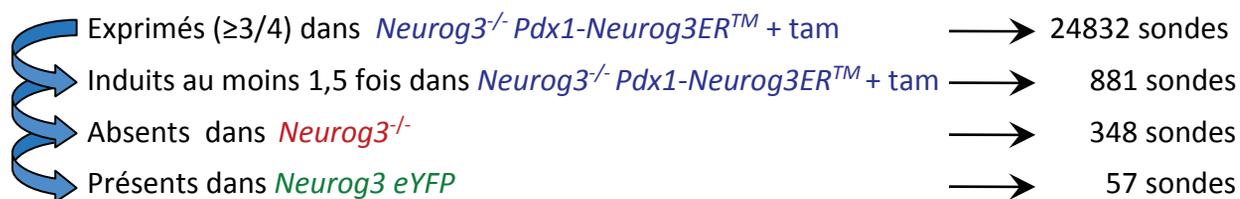


Figure 12 : Filtres et critères de sélections des potentiels gènes cibles de Neurogenin3. (A) La sélection des potentiels gènes cibles de NEUROG3 a été réalisée en comparant le transcriptome des progéniteurs pancréatiques (*Neurog3 eYFP +*), celui du lignage endocrine pancréatique (*Neurog3^{-/-}*), et celui induit par NEUROG3 chez les souris *Neurog3^{-/-} ; Pdx1-Neurog3ERTM* traitées avec le 4-hydroxytamoxifène. (B) Afin de sélectionner de potentiels gènes cibles de NEUROG3, nous avons choisi les gènes:

- _ exprimés dans au moins 3 des 4 puces à ADN réalisées à partir d'ARN de souris *Neurog3^{-/-} ; Pdx1-Neurog3ERTM + tam*;
- _ induits au moins 1,5 fois chez les souris *Neurog3^{-/-} ; Pdx1-Neurog3ERTM + tam*;
- _ absents chez les souris *Neurog3^{-/-}*;
- _ enrichis dans les progéniteurs pancréatiques (*Neurog3 eYFP*).

duplicata (Mellitzer et al, résultats non publiés ; Fig. 12 A). Ainsi, nous avons comparé mes résultats de puce avec le transcriptome de pancréas de souris sauvages et mutantes pour *Neurog3* caractérisée pour leur perte du lignage pancréatique endocrine (Gradwohl et al., 2000) à E15,5. Nous avons également comparé mes résultats avec le transcriptome des progéniteurs pancréatiques endocrines. Celui-ci a été obtenu grâce à l'utilisation des souris « knock add on » *Neurog3-eYFP* (Mellitzer et al., 2004), qui expriment la eYFP sous le contrôle du promoteur de *Neurog3* par une insertion de la séquence de la eYFP dans la partie 3' non codante de la séquence de *Neurog3*. Le transcriptome des progéniteurs pancréatiques endocrines a été obtenu à partir de cellules triées YFP positives issues de pancréas à E15,5. Le transcriptome des cellules pancréatiques YFP négatives à E15,5 a également été caractérisé.

La sélection des potentiels gènes cible s'est déroulée en quatre étapes (Fig. 12B). Nous avons premièrement sélectionné les sondes qui sont exprimées dans au moins trois des quatre pancréas des souris mutantes transgéniques injectées avec le 4-hydroxytamoxifène. Sur les 45037 sondes composant la puce à ADN, 24832 sondes répondent à ce critère. Puis, afin de choisir les gènes dont l'expression est induite par NEUROG3, nous avons sélectionné les gènes dont l'expression est induite en moyenne au moins 1,5 fois chez les souris mutantes transgéniques injectées avec le 4-hydroxytamoxifène par rapport aux souris mutantes transgéniques non injectées. Seules 881 sondes des 24832 répondent à ce second critère. Ensuite, afin de sélectionner les gènes dont l'expression est spécifiquement induite dans le pancréas par NEUROG3, nous avons sélectionné parmi ces 881 sondes, uniquement celles dont le signal est perdu dans les deux puces à ADNc hybridées avec des ARNm de pancréas des souris *neurog3* mutantes, soit 348 sondes. Finalement, afin de déterminer parmi ces sondes lesquelles pouvaient correspondre à de potentiels gènes cibles directs précoces de NEUROG3, nous avons sélectionné celles qui étaient exprimées dans les deux puces à ADNc hybridées avec des ARNm des progéniteurs pancréatiques endocrines, les cellules eYFP positives dans les souris « knock add on » *Neurog3-eYFP*. Après ces quatre étapes de sélection réalisées sur Excel, j'ai sélectionné 57 sondes (Table 1). Afin de déterminer la valeur statistique de l'induction des gènes chez les souris mutantes transgéniques après injection du 4-hydroxytamoxifène, les P-values de ces 57 sondes ont été déterminées via le logiciel ZOE (Dembele, 2007).

Parmi ces 57 sondes, on trouve *NeuroD1* et *Insm1*, qui sont des gènes cibles de NEUROG3, et qui dans les études préliminaires de la souris transgénique mutante, avaient été caractérisés comme étant induits après l'injection de 4-hydroxytamoxifène (Fig. 11). De plus, l'analyse par puce à ADN montre que l'expression de *Dll1* et *Nkx2.2* ne sont pas induites, comme observé lors de l'étude préliminaire de la souris transgénique mutante (Fig. 11).

Parmi ces 57 sondes, on trouve également des gènes qui ont été caractérisés comme nécessaires au développement ou au fonctionnement des cellules pancréatiques endocrines matures tels que *Arx*, *pax6*, *ChromograninA*, la proprotéine convertase subtilisin/kexin type 2. On trouve également parmi ces 57 sondes, plus de 20 gènes dont la fonction n'était pas encore caractérisée.

Pour l'ensemble des gènes précédemment impliqués dans la différenciation du pancréas, il est à noter que leur P-Value est comprise entre 0,001 et 0,079. Pour la suite de mon étude, nous avons donc choisi deux gènes avec une P-Value inférieure à 0,079. Ces gènes sont le gène *Rfx6* (appelé également *Rfxdc1*), et le gène « *RIKEN cDNA 1700086L19* » (*Riken L19*). J'ai choisi *Rfx6*, car son expression est induite 2,1 fois chez les souris transgéniques mutantes injectées avec le 4-hydroxytamoxifène, perdue chez les mutants *Neurog3* et enrichie 21 fois chez les progéniteurs pancréatiques eYFP positifs. De plus, nous avons choisi *Rfx6* car il est membre de la famille des gènes *Rfx*, qui sont caractérisés par leur domaine de fixation à l'ADN (Emery et al., 1996). Les gènes *Rfx2*, *Rfx3* et *Rfx4* (Wolfe et al., 2006 ; Ait-Lounis et al., 2007 ; Zhang et al., 2007) sont également impliqués dans différents processus liés au développement embryonnaire. Le choix de *Riken L19* a été fait car son expression est induite 46 fois chez les souris transgéniques mutantes injectées avec le 4-hydroxytamoxifène, est perdue chez les mutants *Neurog3* et est induite 122 fois chez les progéniteurs pancréatiques eYFP positifs. De plus, la séquence peptidique présomptive de *RIKEN L19* contient la séquence du domaine SERTA. Bien que la fonction du domaine SERTA soit inconnue, les protéines contenant ce domaine sont impliquées dans la régulation du cycle cellulaire, et le remodelage de la chromatine (Bennetts et al., 2006).

```

          10          20          30          40          50          60          70
-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|
GCATTTAAAAAGATAGATTTTATTTCCTTTTGAGAAGCTTAATACAGGTATGCAATGTGTTTCGACCATA

+1  A F K K I D F I C L L R T * Y R Y A M C F D H I
+2  H L K R * I L F A F * E L N T G M Q C V S T I
+3  I * K D R F Y L P F E N L I Q V C N V F R P Y

          80          90          100          110          120          130          140
-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|
TTCATTTCCCACTCCCGGGACCACTATCCCTGGAAACCCCTACATCTTTCTTCCAACCTTCATGTCATTTG

+1  H F P L P G P L S L E T P T S F F Q L H V I C
+2  F I S H S R D H Y P W K P L H L S S N F M S F V
+3  S F P T P G T T I P G N P Y I F L P T S C H L

          150          160          170          180          190          200          210
-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|
TGTGATTATTTTTTATAGAACATTTTCATATTAATTGAAGGGTTTATTCTTCTCTCATTAGATGAAGTCT

+1  V I I F Y R T F H I N * R V Y S S L I * M K S
+2  * L F F I E H F I L I E G F I L L S F R * S L
+3  C D Y F L * N I S Y * L K G L F F S H L D E V *

          220          230          240          250          260          270          280
-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|
AATGGTATAGAGGCAGAACACTAGACCATGGGAACACCCCATGGAAACACTCCCCCTCGTCTACCCGGTG

+1  N G I E A E H * T M G T P H G N T P P R L P G A
+2  M V * R Q N T R P W E H P M E T L P L V Y P V
+3  W Y R G R T L D H G N T P W K H S P S S T R C

          290          300          310          320          330          340          350
-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|
CCATCCTCCCGTTTTTAGGAGGTAGCCATATAGGGTGCCCAAAGGAATTAGACTCAGGATTGAGTTTCAT

+1  I L P F L G G S H I G C P K E L D S G L S F M
+2  P S S R F * E V A I * G A Q R N * T Q D * V S W
+3  H P P V F R R * P Y R V P K G I R L R I E F H

          360          370          380          390          400          410          420
-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|
GGATTAATATTTCCCTCTTTCAAGTCTGTCAATTGCTCAAACCTGCTTGTAGTAACAGGAACACAGTGGG

+1  D * Y F P L S S L S I A Q T C L * * Q E H S G
+2  I N I S L F Q V C Q L L K P A C S N R N T V G
+3  G L I F P S F K S V N C S N L L V V T G T Q W D

          430          440          450          460          470          480          490
-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|
ACACAGTCCCGGAGCCAAAGGATCCCTCACACCTTTAATGCAGTTCACCTCTGTCCCATTCAACTTC

+1  T Q S R S Q R I P H T F K C S S T S V P I Q L P
+2  H S P G A K G S L T P L N A V P P L S P F N F
+3  T V P E P K D P S H L * M Q F H L C P H S T S

          500          510          520          530          540          550          560
-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|
CATGGGGACACCAATCCAGCTCCAGTGAGTGACCGTTTTCTGGGACGCGAGACTCCTCTTAACGCTCAG

+1  W G H Q S S L Q * V T V F W D A R L L L T L R
+2  H G D T N P A S S E * P F S G T R D S S * R S G
+3  M G T P I Q P P V S D R F L G R E T P L N A Q

```

Figure 13 : Séquence putative du gène *Riken cDNA 1700086L19*. Séquence nucléique déduite d'après les bases de données du NIH, et traduction dans les 3 phases directes (+1, +2, +3). CCGAAA : succession exon-exon putatif; ILK : domaine SERTA putatif

```

          570          580          590          600          610          620          630
-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|
GCCCGCGACGCTCCCGCTCCCGTGTGCGCCAGCTCTCCTGGACCAAGTCAGTGTGACCCGTGGCTGGACCC

+1 P A T L P L P C R Q L S W T K S V * P V A G P
+2 P R R S R S R V A S S P G P S Q C D P W L D P
+3 A R D A P A P V S P A L L D Q V S V T R G W T P

          640          650          660          670          680          690          700
-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|
CTCGTTCTCAGCCTTCTTCCCGAAAAGCCGCGCGCCCGCCTCGCGTTCTGAGCCGGAGCGAGAGGCGCT

+1 L V L S L P S R K A A R P P R V L S R S E R R F
+2 S F S A F L P E K P R A R L A F * A G A R G A
+3 R S Q P S F P K S R A P A S R S E P E R E A L

          710          720          730          740          750          760          770
-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|
TCAGAGAAGGAGGGACCGACCTGCAGGAGCCGGCCGCCCGGAGCTGCAGTGCCTGCCGCCGCGCTG

+1 R E G G T D L Q E P A A P P S C S R C R R R C
+2 S E K E G P T C R S R P P R R A A V A A A A A V
+3 Q R R R D R P A G A G R P A E L Q S L P P P L

          780          790          800          810          820          830          840
-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|
TCGCTGTGCTGTGTCAGAACCCGTCCGAAGAGTGACACCTTCGCGAAAAACAGGAGCTGTGAAACATTTG

+1 R C R C Q N P S E E * H L R E K T G A V K H L
+2 A V A V R T R P K S D T F A K K Q E L * N I C
+3 S L S L S E P V R R V T P S R K N R S C E T F A

          850          860          870          880          890          900          910
-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|
CACACGCAGTCTTCGGATCCCAGGTCCCATTTCATGAGCCGAGGCTGTGAGACATGCTCCATGTGCCAA

+1 H T Q S S D P R S P F M S P R L * D M L H V P R
+2 T R S L R I P G P H S * A R G C E T C S M C Q
+3 H A V F G S Q V P I H E P E A V R H A P C A K

          920          930          940          950          960          970          980
-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|
GGGGAGTTCTCACCACATCTTCCGAGGCTGTCTAGACATACTCTTCGGTCTGAGTCAGCATTACCAG

+1 G V L T T S S E A V L D I H S S V * V S I T S
+2 G E F S P H L P R L S * T Y T L R S E S A L P A
+3 G S S H H I F R G C P R H T L F G L S Q H Y Q

          990          1000          1010          1020          1030          1040          1050
-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|
CTTATCCAAGAGAAGTCCTATCATATTCATCAACAACGTCGAAAGTAGCCTACTTTAAGAGAAAATATGC

+1 L S K R S P I I F I N N V E S S L L * E K I C
+2 Y P R E V L S Y S S T T S K V A Y F K R K Y A
+3 L I Q E K S Y H I H Q Q R R K * P T L R E N M Q

          1060          1070          1080          1090          1100          1110          1120
-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|
AGAAGAAGAAGATCTACATCAAGATTCCCATGGGTATTTTCCAAAACACCTGATATTACCGGAAGACAGG

+1 R R R R S T S R F P W V F S K T P D I T G R Q D
+2 E E E D L H Q D S H G Y F P K H L I L P E D R
+3 K K K I Y I K I P M G I F Q N T * Y Y R K T G

```

Figure 13 (suite) : Séquence putative du gène *RIKEN cDNA 1700086L19*. Séquence nucléique déduite d'après les bases de données de NIH, et traduction dans les 3 phases directes (+1, +2, +3). CCGAAA : succession exon-exon putatif; LK : domaine SERTA putatif

```

          1130      1140      1150      1160      1170      1180      1190
-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|
ACTTGTATCCTCAAACCTTCCCTAGAGAAGCTCAGGTTTCTTGAAGACCCAGAAACCTACCTAAGAAGGT

+1  L Y P Q T F P R E A Q V S * R P R N L P K K V
+2  T C I L K L S L E K L R F L E D P E T Y L R R S
+3  L V S S N F P * R S S G F L K T Q K P T * E G

          1200      1210      1220      1230      1240      1250      1260
-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|
CTGTGTTAATTAACAATTTGCTGAGGAAGATACACCTAGAGACGGAGAAAGAGAGCTATGAATACTTCAA

+1  C V N * Q F A E E D T P R D G E R E L * I L Q
+2  V L I N N L L R K I H L E T E K E S Y E Y F K
+3  L C * L T I C * G R Y T * R R R R K R A M N T S R

          1270      1280      1290      1300      1310      1320      1330
-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|
GGAAGCTCCCTGCTATAAGACTGCTTATTCTGATACGAGAAAACGGCTGAAGTTTATGGTACAGGAGTGC

+1  G S S L L * D C L F * Y E K T A E V Y G T G V L
+2  E A P C Y K T A Y S D T R K R L K F M V Q E C
+3  K L P A I R L L I L I R E N G * S L W Y R S A

          1340      1350      1360      1370      1380      1390      1400
-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|
TGTTCCAGTCTCTCTACTATGAAGAGCTGCACTCCTACCACATTGTGCCTTACACTTCGGAGAGTACCA

+1  F P V S L L * R A A L L P H C A L H F G E Y H
+2  C S Q S L Y Y E E L H S Y H I V P Y T S E S T I
+3  V P S L S T M K S C T P T T L C L T L R R V P

          1410      1420      1430      1440      1450      1460      1470
-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|-----:-----|
TCTATGGGATGGGCTACACTGGTGACCCTTGGAGCAAATTCTCAGTTGCTCATTATAAAATGAATTA

+1  L W D G L H W * P L G A K F S V A H L * N E L
+2  Y G M G Y T G D H L E Q N S Q L L I Y K M N *
+3  S M G W A T L V T T W S K I L S C S F I K * I K

          1480      1490
-----:-----|-----:-----|-----
AAAAGTAACTGTTAAGTTCATTGA

+1  K S N C * V H *
+2  K V T V K F I X
+3  K * L L S S L X

```

Figure 13 (suite) : Séquence putative du gène *RIKEN cDNA 1700086L19*. Séquence nucléique déduite d'après les bases de données du NIH, et traduction dans les 3 phases directes (+1, +2, +3). CCGAAA : succession exon-exon putatif; LK : domaine SERTA putatif

C) Etude du gène *RIKEN cDNA 1700086L19*

1. Caractérisation de la séquence de l'ARNm du gène *Riken L19*

L'étude du gène *Riken L19* a commencé par la caractérisation de sa séquence. Pour cela, je me suis basé sur la séquence couverte par la sonde *Riken L19* sur la puce Affymetrix, sur la séquence putative (XM_358420) de l'ARNm décrit sur le site du NCBI et sur des comparaisons avec les bases de données d'EST. Nous avons ainsi construit la séquence décrite dans la figure 13. La partie 5' de la séquence nucléique est supportée par trois EST (BB_175955, BF_455503 et BM_939330) issues de banque d'EST d'hypothalamus ou d'hippocampe, qui couvrent les nucléotides 1 à 658. Deux EST (BU936920 et BE456732) issues de banque d'EST de testicule couvrent les nucléotides 653 à 1494. La séquence utilisée pour la sonde sur la puce Affymetrix correspond aux nucléotides 927 à 1440 (Fig. 13).

En comparant cette séquence avec la séquence génomique murine, j'ai montré que cette séquence est divisée en 5 exons sur le chromosome 12 : les nucléotides 1 à 653 (exon 1), 654 à 856 (exon 2), 857 à 980 (exon 3), 981 à 1096 (exon 4) et 1097 à 1494 (exon 5).

Afin de vérifier que cette séquence est exprimée dans le pancréas embryonnaire, j'ai réalisé des RT-PCR à partir d'ARN totaux de pancréas à E13,5 avec différents couples d'oligonucléotides répartis sur l'ensemble de la séquence (Fig. 13). Nous avons ainsi montré que la séquence couvrant les exons 2, 3, 4 et 5 est exprimée dans le pancréas embryonnaire avec un couple d'oligonucléotides couvrant les nucléotides 694 à 1340. En revanche, je n'ai pu amplifier par RT-PCR de produit spécifique à l'exon 1.

Ces résultats suggèrent donc que l'exon 1 qui a été mis en évidence à partir de banque d'EST d'hypothalamus n'est pas exprimé dans le pancréas embryonnaire. Ainsi l'ARNm pancréatique de *Riken L19* contiendrait uniquement les exons 2 à 5.

Cependant la traduction de cette séquence ne permet pas de former dans aucune des phases de lecture, une séquence contenant la séquence nucléique du domaine SERTA (Fig. 13).

Nous avons donc tenté une approche de RACE afin de définir la partie 5' de l'ARNm de *Riken L19*. Toutefois, cette approche n'a pas permis la caractérisation de la partie 5' suite à des problèmes techniques.

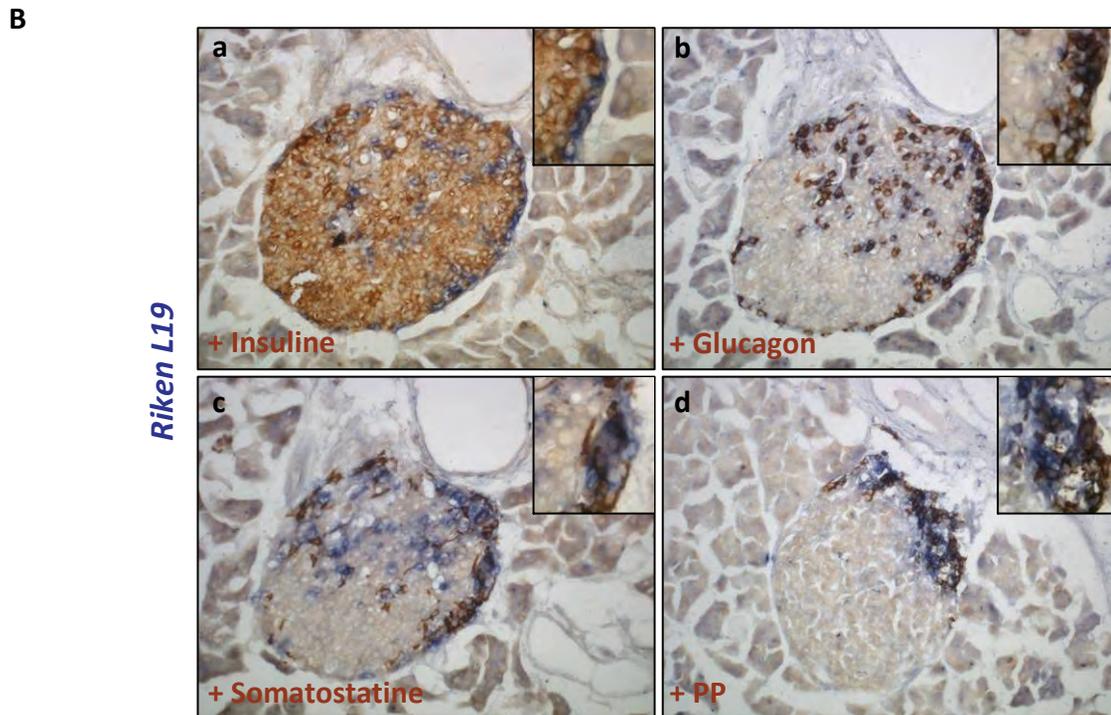
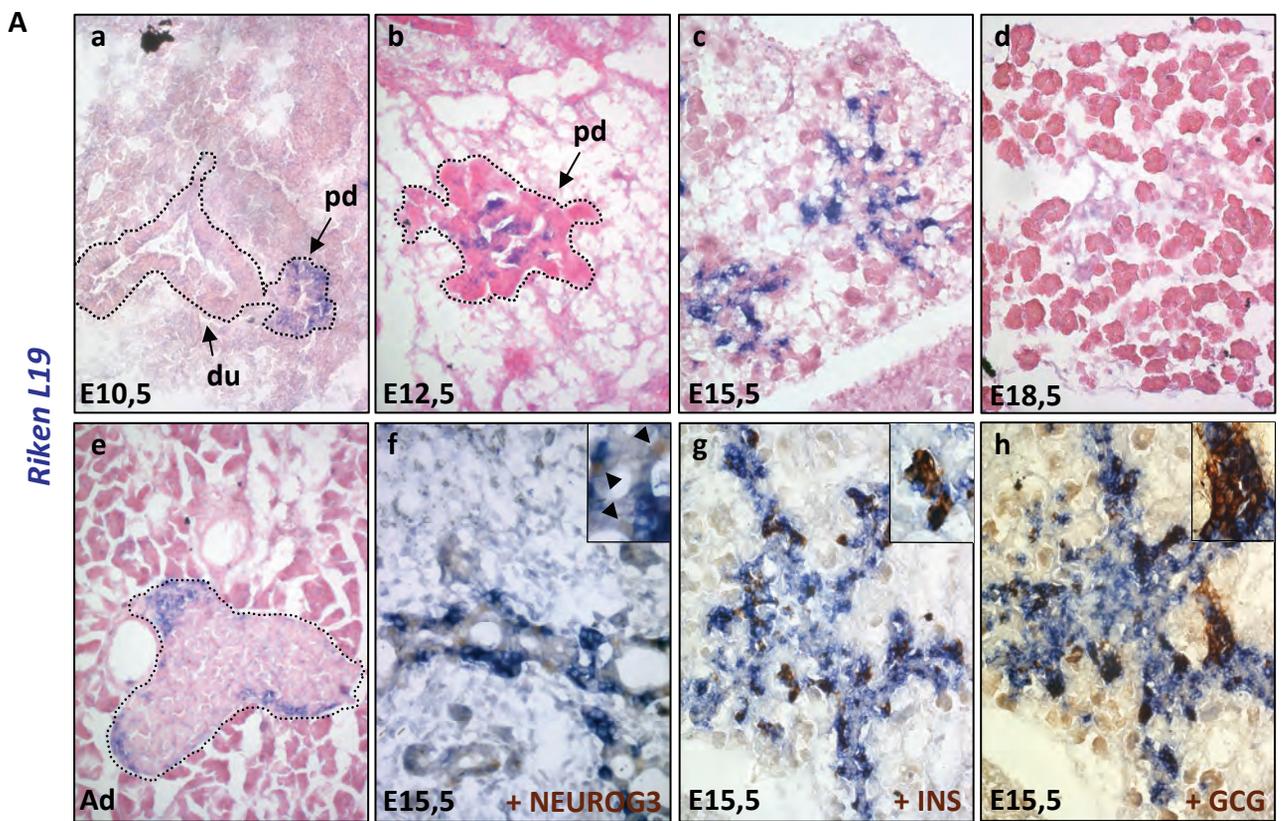


Figure 14 : *Riken L19* est exprimé au cours du développement pancréatique embryonnaire et dans les îlots de Langerhans à l'exception des cellules β . (A) Hybridation *in situ* pour *Riken L19* (bleu, a-h) et immuno-histochimie (brun) pour la protéine NEUROG3 (f), insuline (g) et glucagon (h) sur des sections d'embryons sauvages à E10,5 (a), E12,5 (b), E15,5 (c, f-h), E18,5 (d) et de pancréas adulte (e); Tête de flèche: cellules NEUROG3 positives et *Riken L19* négatives. Comme l'expression de *Neurog3*, l'expression de *Riken L19* est observée à E10,5 puis s'étend pour atteindre son maximum à E15,5 (a-c). *Riken L19* est partiellement coexprimé avec les cellules insulines (g) ou glucagons (h) à E15,5. A E18,5, son expression n'est plus détectée. *Riken L19* est ensuite réexprimé dans la périphérie des îlots de Langerhans. (B) Hybridation *in situ* pour *Riken L19* (bleu, a-d) et immuno-histochimie (brun) pour l'insuline (a), le glucagon (b), la somatostatine (c) et le polypeptide pancréatique (d) sur des sections de pancréas adultes. *Riken L19* est partiellement exprimé dans les cellules α , δ et PP.; pd, pancréas dorsal; du, duodénum; Grossissement: 20X (sauf Af 40x), inserts : 40x (sauf Af zoom).

Par ces différentes approches, je n'ai donc pu caractériser la séquence de l'ARNm pancréatique de *Riken L19*.

Toutefois, récemment, de nouvelles séquences de *Riken L19* ont été prédites chez le macaque (XM_001096508), le gorille (XM_001153624) et le cheval (XM_001493274). Ces séquences diffèrent de notre séquence dans leur partie 5' et la suppression de l'exon 3. Ces séquences permettent l'expression de protéine chez le macaque (XP_001096508), le gorille (XP_001153624) et le cheval (XP_001493324) contenant la séquence du domaine SERTA. Une nouvelle étude de l'ARNm chez la souris en se basant sur ces séquences pourrait peut être permettre de définir la séquence de l'ARNm murin de *Riken L19*.

2. Expression du gène *Riken L19*

Bien que nous n'ayons pu caractériser le transcrite du gène *Riken L19*, j'ai caractérisé l'expression de ce gène au cours du développement pancréatique et dans le pancréas adulte (Fig. 14) par hybridation *in situ*, suivie ou non d'immuno-histochimie pour *NEUROG3* et les hormones pancréatiques. La sonde pour les Hybridations *in situ* a été transcrite à partir d'un ADNc contenant les nucléotides 694 à 1340 de la séquence décrite à la figure 13. Cette séquence comprend la séquence contenant les acides nucléiques codant pour le putatif domaine SERTA.

Nous avons ainsi mis en évidence que l'expression de *Riken L19* est détectable dès E10,5 dans le bourgeon pancréatique dorsal. Puis, son expression s'étend dans l'épithélium pancréatique pour atteindre son maximum à E15,5, comme l'expression de *Neurog3* (Fig. 14Aa-c). A E15,5, *Riken L19* n'est pas exprimé dans les cellules exprimant *NEUROG3* mais semble être exprimé dans une partie des cellules glucagon et insuline (Fig. 14A f-h). A E18,5, *Riken L19* ne semble plus être exprimé dans l'épithélium pancréatique (Fig. 14Ad).

En revanche, au stade adulte, le gène *Riken L19* est réexprimé dans une partie des cellules des îlots de Langerhans, stade auquel *Neurog3* n'est pas exprimé (Fig. 14A e). Dans les îlots de Langerhans, *Riken L19* est exprimé dans les cellules sécrétant le glucagon, la somatostatine et le polypeptide pancréatique ; mais n'est pas exprimé dans les cellules sécrétant l'insuline (Fig. 14B).

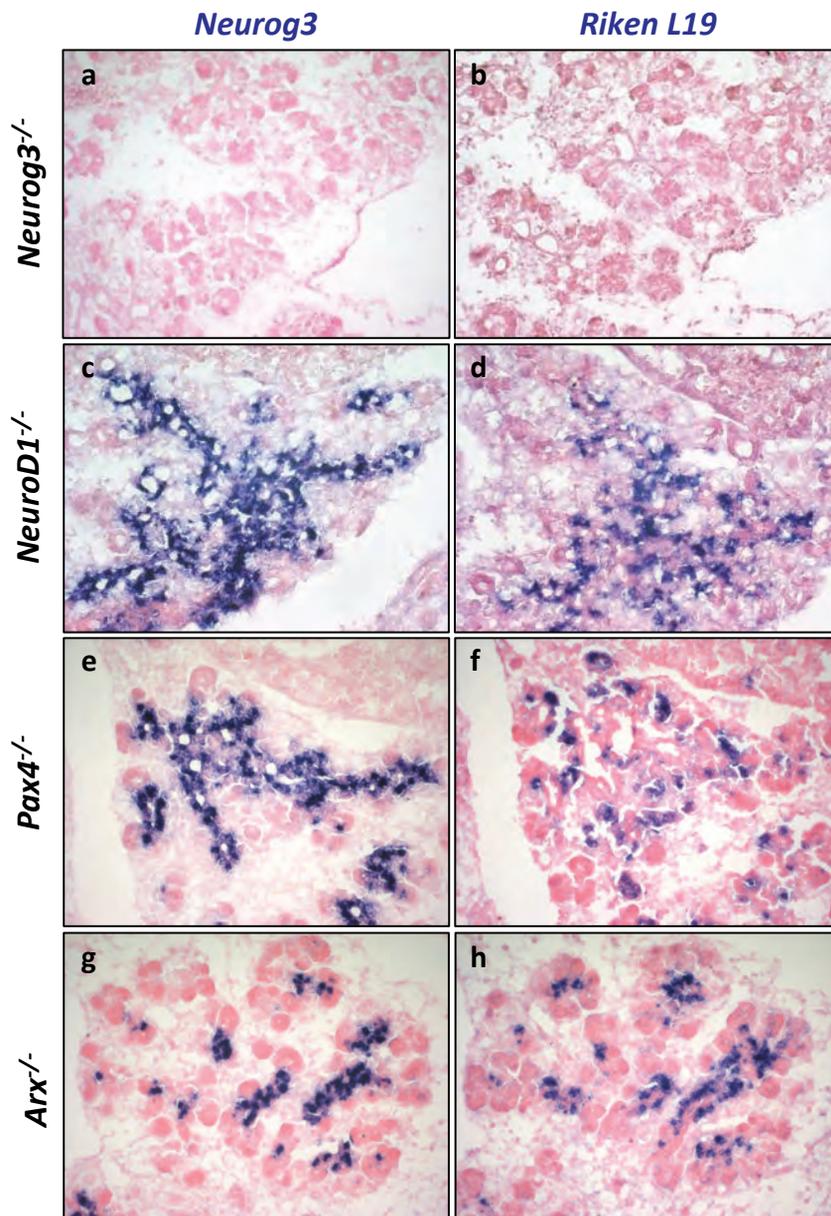


Figure 15 : *Riken L19* est spécifiquement exprimé dans les cellules du pancréas endocrine. Hybridation *in situ* pour *Neurog3* (a,c, e, g) et *Riken L19* (b, d, f, h) sur des sections adjacentes de souris *Neurog3*^{-/-} à E15,5 (a, b), *NeuroD1*^{-/-} (c,d), *Pax4*^{-/-} (e,f) et *Arx*^{-/-} (g,h) à E14,5 (c, h). L'expression de *Riken L19* est perdue chez le mutant *Neurog3*. Mais, son expression est indépendante de celle de *NeuroD1*, *Pax4* et *Arx*, trois gènes exprimés en aval de *Neurog3* et impliqués dans le différenciation endocrine. Grossissement: 20X.

Afin de déterminer si *Riken L19* est spécifiquement impliqué dans la formation du lignage endocrine, j'ai réalisé des hybridations *in situ* sur des sections de pancréas *Neurog3*^{-/-} qui sont caractérisés par la perte du lignage endocrine (Gradwohl et al., 2000 ; Fig. 15b). De plus, afin de déterminer si *Riken L19* est spécifiquement requis pour la différenciation des cellules α (sécrétant le glucagon) ou des cellules β (sécrétant l'insuline), nous avons étudié l'expression de *Riken L19* chez des embryons mutants *Pax4* ou *Arx*, qui sont respectivement caractérisés par une perte de la différenciation des cellules β sécrétant l'insuline, et une perte de la différenciation des cellules α sécrétant le glucagon (Sosa-Pineda et al., 1997 ; Collombat et al., 2003 ; Fig. 15f, h). Finalement, afin de déterminer si l'expression de *Riken L19* est sous le contrôle de *NEUROD1*, le seul gène cible connu pour *NEUROG3* au début de cette étude, nous avons étudié l'expression de *Riken L19* chez des embryons mutants pour *NeuroD1*, dont la perte de fonction induit une réduction de la différenciation des différents sous-types cellulaires endocriniens (Huang et al., 2000 ; Naya et al., 1997 ; Fig. 15d)

Nous avons ainsi montré que l'expression de *Riken L19* est spécifique du lignage endocrine par la perte de son expression dans la souris *Neurog3* mutante (Fig. 15b). En revanche, son expression ne semble pas altérée chez les souris mutantes pour *NeuroD1*, *Pax4* ou *Arx* (Fig. 15d, f, h), et serait donc indépendante de celle de *Arx*, *Neurod1* et *Pax4*.

L'ensemble de ces résultats montre que *Riken L19* est spécifiquement exprimé dans le lignage endocrine à partir de E10,5 en aval de *Neurog3*, de façon indépendante des autres facteurs de transcription étudiés déjà impliqués dans la différenciation du pancréas endocrine (*NeuroD1*, *Pax4* et *Arx*). De plus, l'absence de coexpression de *Neurog3* et *Riken L19* laisse suggérer que *Riken L19* n'est pas une cible directe de *Neurog3*. Il faut également noter que dans le pancréas adulte *Riken L19* est exprimé dans les cellules α , δ et PP, mais pas dans les cellules β .

Toutefois, malgré cet intéressant patron d'expression de *Riken L19*, nous n'avons pas pu définir de séquence pour ce gène permettant d'exprimer une protéine contenant le domaine SERTA, et n'ai donc pas poursuivi l'étude de ce gène.

```

      10      20      30      40      50      60      70
----:----|----:----|----:----|----:----|----:----|----:----|
ACACAGAGCTCCTTCAAGCCCGGGCGGCGTGCCGAGATCCGGGCGGTGGGAAGGATGGCTAAGGTCCGGG

                                     M A K V R E

      80      90      100     110     120     130     140
----:----|----:----|----:----|----:----|----:----|----:----|
AACTGGAGGAAGCTTTTCGTGCAGGAGCAGCCTTCGCCCAACTTCCCTCCGAGATCGCGGAGGAGTGCTG

L E E A F V Q E Q P S P Q L P S E I A E E C C

      150     160     170     180     190     200     210
----:----|----:----|----:----|----:----|----:----|----:----|
CGCGCAGCTGCTTGGCAAGGGCTTGCTGGTCTACCCTGAGGACAGTGCTTACCTGCTGGCTGAAACAGCC

A Q L L G K G L L V Y P E D S A Y L L A E T A

      220     230     240     250     260     270     280
----:----|----:----|----:----|----:----|----:----|----:----|
GCGGGCGCTCGGAGCAGCGGTGAGAAAGGCGGAGATCCCGGGCTGCAGGTGGGAGTTAAATCAGAAATGC

A G A R S S G E K G G D P G L Q V G V K S E M Q

      290     300     310     320     330     340     350
----:----|----:----|----:----|----:----|----:----|----:----|
AGTTAAACAACGGCAACTTTTCTCGGAAGAAGAGGATGCGGACACCCAGGAGAGCAAGACCAAAGCAGC

L N N G N F S S E E E D A D T Q E S K T K A A

      360     370     380     390     400     410     420
----:----|----:----|----:----|----:----|----:----|----:----|
CGACCCACAGCTCTCTCAGAAGAAGAGCATCACGCAGATGATGAAGGATAAGAAGAAGCAGACTCAGCTC

D P Q L S Q K K S I T Q M M K D K K K Q T Q L

      430     440     450     460     470     480     490
----:----|----:----|----:----|----:----|----:----|----:----|
ACCCTGCAGTGGCTTGAAGACAATTACATAGTGTGTGAAGGAGTTTGCCTCCACGGTGCATTCTTTATG

T L Q W L E D N Y I V C E G V C L P R C I L Y A

      500     510     520     530     540     550     560
----:----|----:----|----:----|----:----|----:----|----:----|
CTCACTACCTAGATTTCTGCAGGAAAGAGAAACTGGAGCCAGCTTGTGCAGCCACCTTTGGAAAGCAAT

H Y L D F C R K E K L E P A C A A T F G K T I

      570     580     590     600     610     620     630
----:----|----:----|----:----|----:----|----:----|----:----|
TCGCCAAAATTTCCCCTCCTAACAACTCGAAGACTTGAACAAGAGGACATTCAAAGTACCATTACTAT

R Q K F P L L T T R R L G T R G H S K Y H Y Y

      640     650     660     670     680     690     700
----:----|----:----|----:----|----:----|----:----|----:----|
GGGATTGGCATCAAAGAAAGCAGTGCATATTACCACTCTGTTTATTCTGGAAAGGGCTTGACAAGGTTTT

G I G I K E S S A Y Y H S V Y S G K G L T R F S

```

Figure 16 : Séquence nucléique et peptidique de *Rfx6*. Séquence nucléique exprimée dans le pancréas embryonnaire, et séquence correspondante en acides aminés. CCGAAA : succession exon-exon ; Domaines peptidiques conservés: **Domaine de fixation à l'ADN**, **région B**, **région C** et **domaine de dimérisation**

710 720 730 740 750 760 770
 -----|-----|-----|-----|-----|-----|-----|
 CGGGGAGCAAACCTGAAGAATGAGGGTGGTTTTACACGTAAATACTCACTTAGCTCAAAAACAGGAACACT
 G S K L K N E G G F T R K Y S L S S K T G T L
 780 790 800 810 820 830 840
 -----|-----|-----|-----|-----|-----|-----|
 TCTTCCAGAATTTCCAGCGCTCAACATCTTGTGTACCAAGGATGCATTTCTAAGGACAAGGTTGATACT
 L P E F P S A Q H L V Y Q G C I S K D K V D T
 850 860 870 880 890 900 910
 -----|-----|-----|-----|-----|-----|-----|
 CTCATAATGATGTACAAAACCTCACTGCCAGTGCATACTCGACAATGCCATCAATGGGAACTTTGAAGAGA
 L I M M Y K T H C Q C I L D N A I N G N F E E I
 920 930 940 950 960 970 980
 -----|-----|-----|-----|-----|-----|-----|
 TCCAGCATTTCTTACTACACTTTTGGCAAGGAATGCCAGACCATCTCCTTCCCCTGCTTGAAAATCCTGT
 Q H F L L H F W Q G M P D H L L P L L E N P V
 990 1000 1010 1020 1030 1040 1050
 -----|-----|-----|-----|-----|-----|-----|
 TATCATTGATATTTCTGTGTCTGTGACTCAATTCTTTATAAGGTTCTTACAGATGTCCTCATTCTCTGCA
 I I D I F C V C D S I L Y K V L T D V L I P A
 1060 1070 1080 1090 1100 1110 1120
 -----|-----|-----|-----|-----|-----|-----|
 ACAATGCAAGAAATGCCCTGAAAGTTTGTTAGCAGATATAAGAAATTTTGCATAAAATTTGGGAACAGTGGG
 T M Q E M P E S L L A D I R N F A K N W E Q W V
 1130 1140 1150 1160 1170 1180 1190
 -----|-----|-----|-----|-----|-----|-----|
 TTGTTTCATCCTTGAAAACCTTGCCAGAAGCCCTCATTGATAAGAAAATCCCCATTTTGCGAAGATTTGT
 V S S L E N L P E A L I D K K I P I L R R F V
 1200 1210 1220 1230 1240 1250 1260
 -----|-----|-----|-----|-----|-----|-----|
 ATCTTCCCCTGAAGCGACAGACATCTTTCTTGCATCTTGCTCAGATTGCCAGACCAGCTCTCTTTGACCAG
 S S L K R Q T S F L H L A Q I A R P A L F D Q
 1270 1280 1290 1300 1310 1320 1330
 -----|-----|-----|-----|-----|-----|-----|
 CATGTGGTGAATGCCATGGTATCTGATATTGAAAAGGTTGACTTAAATAGTATTGGGTCTCAGGCTCTTC
 H V V N A M V S D I E K V D L N S I G S Q A L L
 1340 1350 1360 1370 1380 1390 1400
 -----|-----|-----|-----|-----|-----|-----|
 TTACCATATCCAACAGCACAGACACGGAATCTGACATCTACAGTGAACATGACTCTATTACTGTGTTCCA
 T I S N S T D T E S D I Y S E H D S I T V F Q

Figure 16 (suite) : Séquence nucléique et peptidique de *Rfx6*. Séquence nucléique exprimée dans le pancréas embryonnaire, et séquence correspondante en acides aminés. CCGAAA : succession exon-exon ; Domaines peptidiques conservés: **Domaine de fixation à l'ADN**, **région B**, **région C** et **domaine de dimérisation**

1410 1420 1430 1440 1450 1460 1470
 -----|-----|-----|-----|-----|-----|-----|
 AGAACTGAAAGATCTCCTTAAGAAGAATGCTACAGTGGAGGCATTTATTGAATGTTGGACACTGTGGTA
 E L K D L L K K N A T V E A F I E W L D T V V

1480 1490 1500 1510 1520 1530 1540
 -----|-----|-----|-----|-----|-----|-----|
 GAGCAGAGAGTTATTAAGATGAGCAAACAAAATGGAAGATCTCTGAAGAAGAGGGCTCAAGACTTTCTGCTC
 E Q R V I K M S K Q N G R S L K K R A Q D F L L

1550 1560 1570 1580 1590 1600 1610
 -----|-----|-----|-----|-----|-----|-----|
 TCAAATGGAGCTTTTTTGGTGCCCGCGTGATGCATAATCTTACCTTGAACAACGCATCAAGTTTTGGCTC
 K W S F F G A R V M H N L T L N N A S S F G S

1620 1630 1640 1650 1660 1670 1680
 -----|-----|-----|-----|-----|-----|-----|
 TTTCCATCTGATCCGAATGCTTCTGGATGAGTACATTCTCCTGGCCATGGAGACTCAATTTAACAATGAC
 F H L I R M L L D E Y I L L A M E T Q F N N D

1690 1700 1710 1720 1730 1740 1750
 -----|-----|-----|-----|-----|-----|-----|
 AAAGAGCAGGAAC TACAGAATTTATTGGACAAGTATATGAAGAACTCGGATGCGAGTAAAGCTGCCTTCA
 K E Q E L Q N L L D K Y M K N S D A S K A A F T

1760 1770 1780 1790 1800 1810 1820
 -----|-----|-----|-----|-----|-----|-----|
 CAGCTTCCCCGAGCTCTTGCTTTCTGGCCAACCGAAATAAGGCTAGCTCACTTGCCAGTGACACTGTGAA
 A S P S S C F L A N R N K A S S L A S D T V K

1830 1840 1850 1860 1870 1880 1890
 -----|-----|-----|-----|-----|-----|-----|
 GAACGAAAGCCACGTGGAGACATCCTATGTCCCTCTGCCTTCCAGCCAGCCTGGAGCCATACCCCCTGCT
 N E S H V E T S Y V P L P S S Q P G A I P P A

1900 1910 1920 1930 1940 1950 1960
 -----|-----|-----|-----|-----|-----|-----|
 CTGCACCCATTCTCAACTGAGGACACTGATAACATGCCACTCCCAGGTCAAATAGAGCTTTCACAAAGTA
 L H P F S T E D T D N M P L P G Q I E L S Q S T

1970 1980 1990 2000 2010 2020 2030
 -----|-----|-----|-----|-----|-----|-----|
 CTGGCCATCTGATGACACCACCGATTTCTCCAGCCATAGCAAGCAGAGGAAGTGTATTAAACCAAGGGCC
 G H L M T P P I S P A I A S R G S V I N Q G P

2040 2050 2060 2070 2080 2090 2100
 -----|-----|-----|-----|-----|-----|-----|
 AATGGCGAGCAGACCCCCGAGCGTGGGCACAGTTCTCTCAGCTCCAACACATTGCTCAACATATGCAGAA
 M A S R P P S V G T V L S A P T H C S T Y A E

Figure 16 (suite) : Séquence nucléique et peptidique de *Rfx6*. Séquence nucléique exprimée dans le pancréas embryonnaire, et séquence correspondante en acides aminés. CCGAAA : succession exon-exon ; Domaines peptidiques conservés: **Domaine de fixation à l'ADN**, **région B**, **région C** et **domaine de dimérisation**

2110 2120 2130 2140 2150 2160 2170
 -----|-----|-----|-----|-----|-----|-----|
 CCAATTTATCCTACGCTCTCTCCAGCCAACCACGACTTTTATGGGACCAACTCTAACTATCAGACTATGT
 P I Y P T L S P A N H D F Y G T N S N Y Q T M F

2180 2190 2200 2210 2220 2230 2240
 -----|-----|-----|-----|-----|-----|-----|
 TTAGGACACAGTCTCACCCCTGCATCAAGCCTCTATGCTCACCGTGCAGAGCATGGGCGGTGCATGGCCTG
 R T Q S H P A S S L Y A H R A E H G R C M A W

2250 2260 2270 2280 2290 2300 2310
 -----|-----|-----|-----|-----|-----|-----|
 GACTGAACAGCAGCTTTCTAGAGACTTCTTTGGTGGCAGTTGTGCTGGGTCTCCATATAATTGTAGGCCA
 T E Q Q L S R D F F G G S C A G S P Y N C R P

2320 2330 2340 2350 2360 2370 2380
 -----|-----|-----|-----|-----|-----|-----|
 CCTTCCAGTTATGGACCATCCACACACACAAGAGTCACACAGCATGCAGGTTTTGAACACAGGAAGCT
 P S S Y G P S T H T Q E S H S M Q V L N T G S F

2390 2400 2410 2420 2430 2440 2450
 -----|-----|-----|-----|-----|-----|-----|
 TCAATTTCCCTTAGTAATGCAGGAGCTGGCAGCTGCCAAGGGTCAACATTGCCTTCTAATTCTCCCAATGG
 N F L S N A G A G S C Q G S T L P S N S P N G

2460 2470 2480 2490 2500 2510 2520
 -----|-----|-----|-----|-----|-----|-----|
 ATACTATGGAAACAATATAAACTACTCAGAGGCACATAGGCTTGGATCGATGGTGAACCAACATGTTTCA
 Y Y G N N I N Y S E A H R L G S M V N Q H V S

2530 2540 2550 2560 2570 2580 2590
 -----|-----|-----|-----|-----|-----|-----|
 GTCATCAGCAGTGTGCGCTCCCTGCCTCCCTACAGTGATATTCATGATCCACTTAACATTTTTAGATGACA
 V I S S V R S L P P Y S D I H D P L N I L D D S

2600 2610 2620 2630 2640 2650 2660
 -----|-----|-----|-----|-----|-----|-----|
 GCAGCCGGAAGCAGAACAACCTCGTTTTATGCAGACACATTGTCTCCTGTTGCATGTCGTACTIONTGTAGT
 S R K Q N N S F Y A D T L S P V A C R T T V V

2670 2680 2690 2700 2710 2720 2730
 -----|-----|-----|-----|-----|-----|-----|
 AGCTTCCAACCTTGCAAACCCAGATTCCCTTCATCTTCATCCCAGTGTATGTATGGAACCTCCAATCAGTAT
 A S N L Q T Q I P S S S S Q C M Y G T S N Q Y

2740 2750 2760 2770 2780 2790 2800
 -----|-----|-----|-----|-----|-----|-----|
 CCAGTGCAAGATAGTCTGGACTCCAATGCAGCAAGCAACAGAGAAATGGTGTCTCTTTACCACCCATCA
 P V Q D S L D S N A A S N R E M V S S L P P I N

Figure 16 (suite) : Séquence nucléique et peptidique de *Rfx6*. Séquence nucléique exprimée dans le pancréas embryonnaire, et séquence correspondante en acides aminés. CCGAAA : succession exon-exon ; Domaines peptidiques conservés: **Domaine de fixation à l'ADN**, **région B**, **région C** et **domaine de dimérisation**

```

      2810      2820      2830      2840      2850      2860      2870
----:----|----:----|----:----|----:----|----:----|----:----|----:----|
ACACCGTGTTTATGGGGACAGCAGCTGGAGACACTTAAATGAACAATGTAGAATGCAGGTGAAACTCAA

  T  V  F  M  G  T  A  A  G  D  T

      2880      2890      2900      2910      2920      2930      2940
----:----|----:----|----:----|----:----|----:----|----:----|----:----|
CGCTCTGCTGCCTTGATGCCATTATCCACCTCCAGCTGCCACGAGAGAAGAGAGGGTGGAAAGGTGCAATG

      2950      2960      2970      2980      2990      3000      3010
----:----|----:----|----:----|----:----|----:----|----:----|----:----|
GCTGAATGTTATTTTCAGATCACTGCTATGCCATAACAACCTCCATAGTGAATGGAGATACTTGTACTGCAC

      3020      3030      3040      3050      3060      3070      3080
----:----|----:----|----:----|----:----|----:----|----:----|----:----|
ACAAATAGGC TAAAATGTGAAGTCATTATTAATATAACTTCAAATACTATCAAATGAAACCAACGAAAA

      3090      3100      3110      3120      3130      3140      3150
----:----|----:----|----:----|----:----|----:----|----:----|----:----|
TTTAAAGATGCAGACATTACAACGATGTAATATTTCACTTTCTAACTTATTGAAAAAGCTCTATTTTCCT

      3160      3170      3180      3190      3200      3210      3220
----:----|----:----|----:----|----:----|----:----|----:----|----:----|
CCCTCTTTGGTTTCTAAAAGAAGAAAAGTATATTC AATGCCTTCACAATATCTTTAATTTATAATCATTG

      3230      3240      3250      3260      3270      3280      3290
----:----|----:----|----:----|----:----|----:----|----:----|----:----|
TTTACTATCACCTATTTGACCAAAGTCAAATGTCTAAGTTCTACCTCCTGTGGAAATGTTTACAAGTCA

      3300      3310      3320      3330      3340      3350      3360
----:----|----:----|----:----|----:----|----:----|----:----|----:----|
AGACTTTATTGTATGTAGCTAGATCCAAGTGACTTGACCTTCTACCCCTGTGCACTAGTGTGGTGATTTCC

      3370      3380      3390      3400      3410      3420      3430
----:----|----:----|----:----|----:----|----:----|----:----|----:----|
TGTTACATAGTAGTGTCCAATTCTATGTAAC TGAATAGAAATGATAAACACTTTCCTTCTTTATTTAAA

      3440      3450      3460
----:----|----:----|----:----|----
ATGATAAGAGGAAATAAAAAAGAATGTATGG

```

Figure 16 (suite) : Séquence nucléique et peptidique de *Rfx6*. Séquence nucléique exprimée dans le pancréas embryonnaire, et séquence correspondante en acides aminés. CCGAAA : succession exon-exon ; Domaines peptidiques conservés: **Domaine de fixation à l'ADN**, **région B**, **région C** et **domaine de dimérisation**

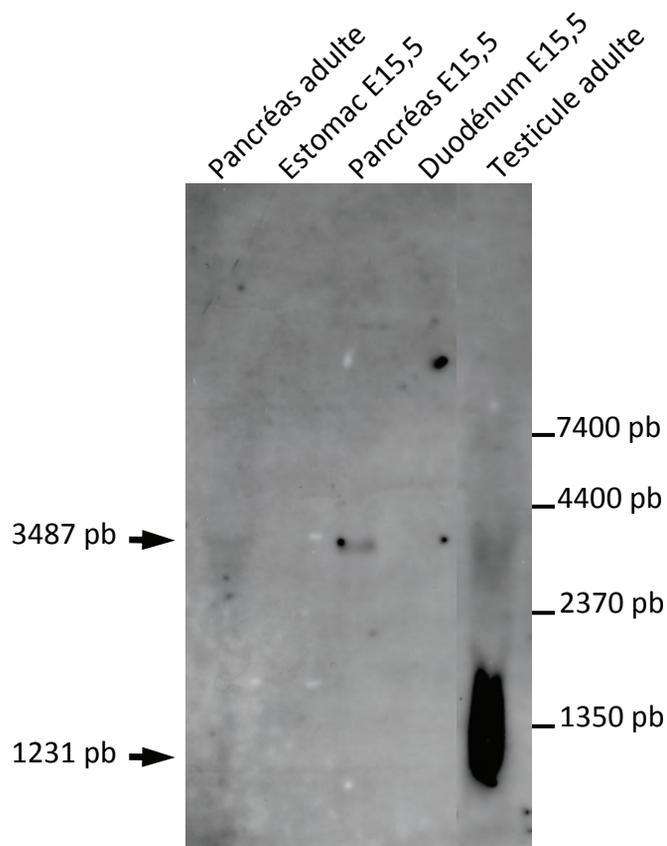


Figure 17 : Le transcrit de *Rfx6* dans le pancréas mesure environ 3487 nucléotides. Northernblot pour *Rfx6* sur des extraits d'ARN totaux de pancréas, de testicules adultes et de tissus embryonnaires à E15,5 (estomac, pancréas, duodénum et tête). Dans le pancréas adulte et embryonnaire, une bande est détectée vers 3487 nucléotides. Dans les testicules adultes, 2 transcrits semblent être présents: un de 3371 nucléotides et un autre de 1231 nucléotides.

D) Etude du gène *Regulatory Factor X 6*

1. Caractérisation de la séquence de l'ARNm du gène *Rfx6*

Le second gène que nous avons étudié pendant mon doctorat est le gène *Rfx6* (initialement appelé *Rfxdc1*). Ce gène était comme pour *Riken L19*, non caractérisé au début de cette étude, aussi bien du point de vue de sa séquence, que de sa fonction.

Nous avons initié l'analyse du gène *Rfx6* par la caractérisation de sa séquence. Car bien que rattaché à la famille des gènes *Rfx*, la séquence définie au début de cette étude (aujourd'hui référencée sur le site du NCBI sous la référence NM_177306) ne contenait pas le domaine de fixation à l'ADN caractéristique des gènes de la famille *Rfx*. Par l'analyse des bases de données du NCBI, j'ai mis en évidence plusieurs EST exprimés dans les îlots de Langerhans (CJ132427, BP765921 et BY741659). Les séquences couvertes par ces trois EST pouvaient potentiellement coder pour le domaine de fixation à l'ADN spécifique de la famille des gènes *Rfx*. Afin de confirmer l'expression de ces séquences dans le pancréas embryonnaire, j'ai réalisé des RT-PCR sur de l'ARN de pancréas issu d'embryon à E13,5, stade auquel avaient été réalisées les études transcriptomiques des souris *Neurog3^{-/-}*; *Pdx1-Neurog3ERTM*. Nous avons ainsi confirmé l'expression de cette séquence dans le pancréas embryonnaire et cloné l'ADNc décrit à la figure 16.

Cette séquence comporte 3461 nucléotides et permet l'expression d'une protéine de 927 acides aminés contenant : le domaine de fixation à l'ADN caractéristique de la famille RFX, les séquences codant pour la « région B » et « région C » ainsi qu'un domaine de dimérisation conservés chez les protéines RFX (Fig. 16 ; Emery et al., 1996).

Pour confirmer que le transcrit que j'avais cloné était complet, j'ai réalisé des northern blots (Fig. 17) sur des ARN d'estomac, de pancréas, de duodénum à E15,5 et de pancréas et de testicules adultes. Pour cela, j'ai utilisé une sonde transcrite à partir d'un clone contenant les nucléotides 1829 à 2685 de la séquence décrite à la figure 16. Nous avons choisi cette séquence pour son homologie entre la séquence courte (NM_177306, 3088 nucléotides) et la séquence que je venais de décrire (3461 nucléotides). Nous avons ainsi mis en évidence que le transcrit de *Rfx6* dans le pancréas embryonnaire et le pancréas adulte faisait environ 3487 nucléotides d'après l'estimation de la taille des bandes sur le northern blot (Fig. 17). Ce résultat montre que l'ADNc cloné correspond à l'ARN complet de *rfx6* avec les parties non

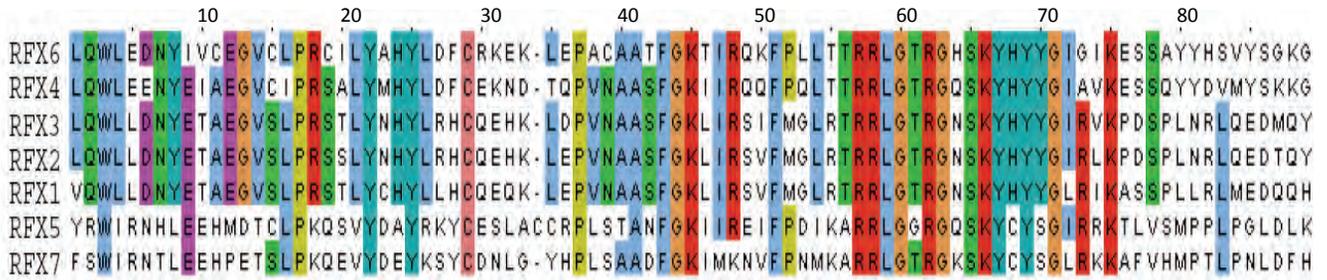
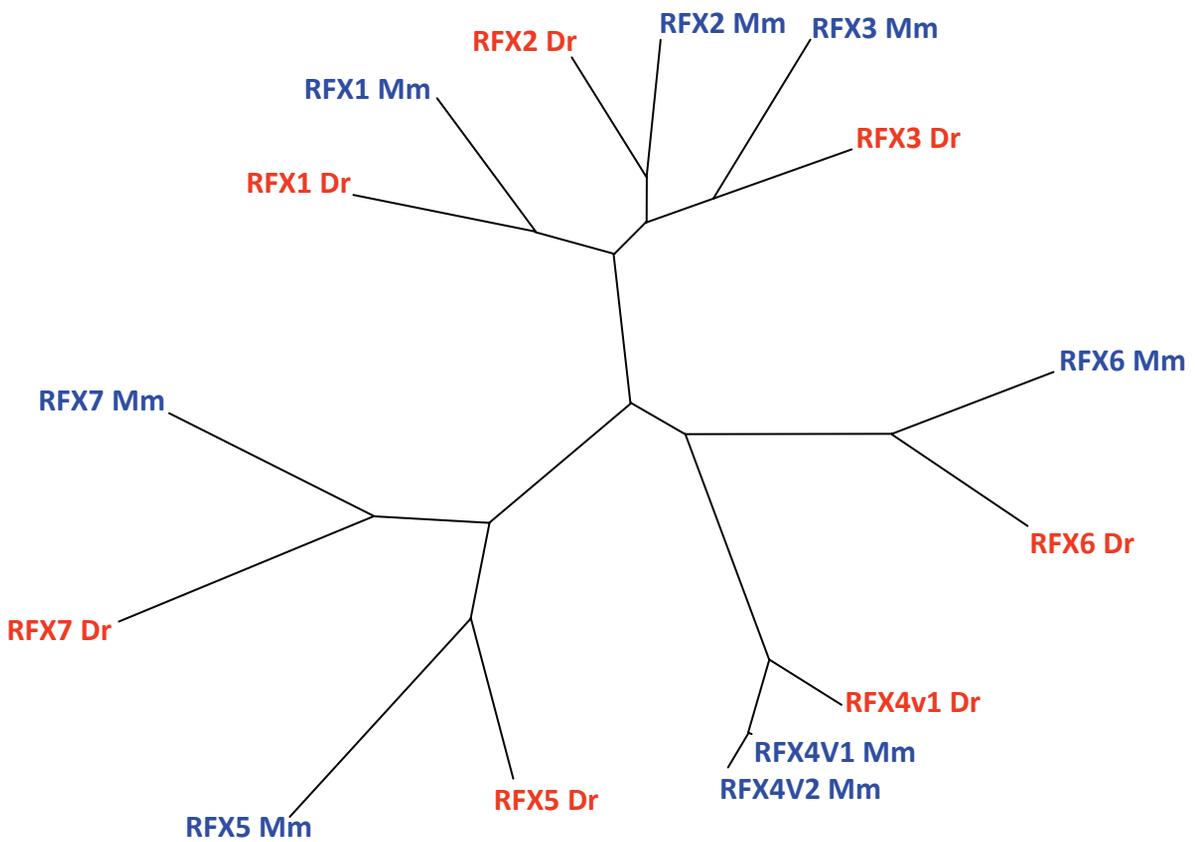
A**B**

Figure 18 : Comparaison des protéines de la famille RFX. (A) comparaison des domaines de fixation à l'ADN des protéines RFX1 à RFX7 murins. (B) comparaison des séquences protéiques de la famille RFX chez la souris et le poisson zèbre. Mm: *Mus musculus*, Dr: *Danio rerio*.

traduites (5' et 3'UTR). De plus, dans les testicules adultes, 2 transcrits semblent être présents : un de 3371 nucléotides et un autre de 1231 nucléotides. Ces résultats avec l'ARN de testicules adultes suggèrent donc la présence d'un transcrit qui n'est pas décrit dans les banques de données. De plus, le transcrit de 3371 nucléotides, plus court que celui observé dans les autres tissus étudiés, pourrait correspondre au second transcrit (variant 2) décrit pour *Rfx6* (NM_177306), plus court que le variant 1 présent dans le pancréas.

En alignant la séquence de l'ARNm de *Rfx6* sur l'ADN génomique, j'ai montré que le gène codant pour *Rfx6* est situé sur le chromosome 10 et que l'ARNm est constitué de 19 exons. Cette séquence correspond à la séquence aujourd'hui décrite sur le site du NCBI pour le variant 1 de *Rfx6* (NM_001159389), à l'exception des sites d'épissage qui diffèrent entre ma séquence et la séquence du NCBI. Cette différence peut s'expliquer par le fait que la séquence du NCBI a été déduite par une approche bioinformatique, alors que mes sites d'épissage ont été déduits de la localisation des sites consensus d'épissage (Kreivi et al., 1996).

De plus, afin de déterminer les similarités entre les différents gènes de la famille *Rfx*, j'ai comparé la séquence peptidique des domaines de fixation à l'ADN des protéines murines : RFX1 (NP_033081), RFX2 (NP_033082), RFX3 (NP_035395), RFX4 (NP_001020089), RFX5 (NP_059091), RFX7 (NP_001028708) et RFX6 (séquence décrite précédemment, Fig. 18A). Nous avons ainsi mis en évidence que la séquence protéique de RFX6 déduite de la séquence nucléique que j'ai clonée contient un domaine de fixation à l'ADN très similaire à celui de RFX1, RFX2, RFX3 et RFX4 avec une pourcentage d'homologie de 64 %, 64 %, 63 % et 69 % respectivement. Cette homologie de séquence entre le domaine de fixation à l'ADN des protéines RFX1, RFX2, RFX3 et RFX4 suggère donc que le domaine de fixation à l'ADN de RFX6 serait fonctionnel. En revanche, on peut noter que le domaine de fixation à l'ADN de RFX5 et RFX7 est très différent de celui des autres protéines RFX murines. En effet, l'homologie avec le domaine de fixation à l'ADN de RFX6 est de 41 % et 36 % respectivement pour RFX5 et RFX7. J'ai également réalisé un arbre phylogénétique à partir des séquences des protéines RFX murines et de poisson zèbre (zRFX1 : ENSDARPO0000039332, zRFX2 : ENSDARPO0000010260, zRFX3 : ENSDARPO0000041235, zRFX4 : ENSDARPO0000032295, zRFX5 : ENSDARPO0000086706, zRFX6 : ENSDART00000061122 ; Fig. 18B). On peut y noter trois sous groupes. Le premier d'entre eux est constitué des protéines RFX5 et RFX7 murines et de poisson zèbre. Le second groupe est constitué des protéines RFX1 à RFX3 murines et

Rfx6

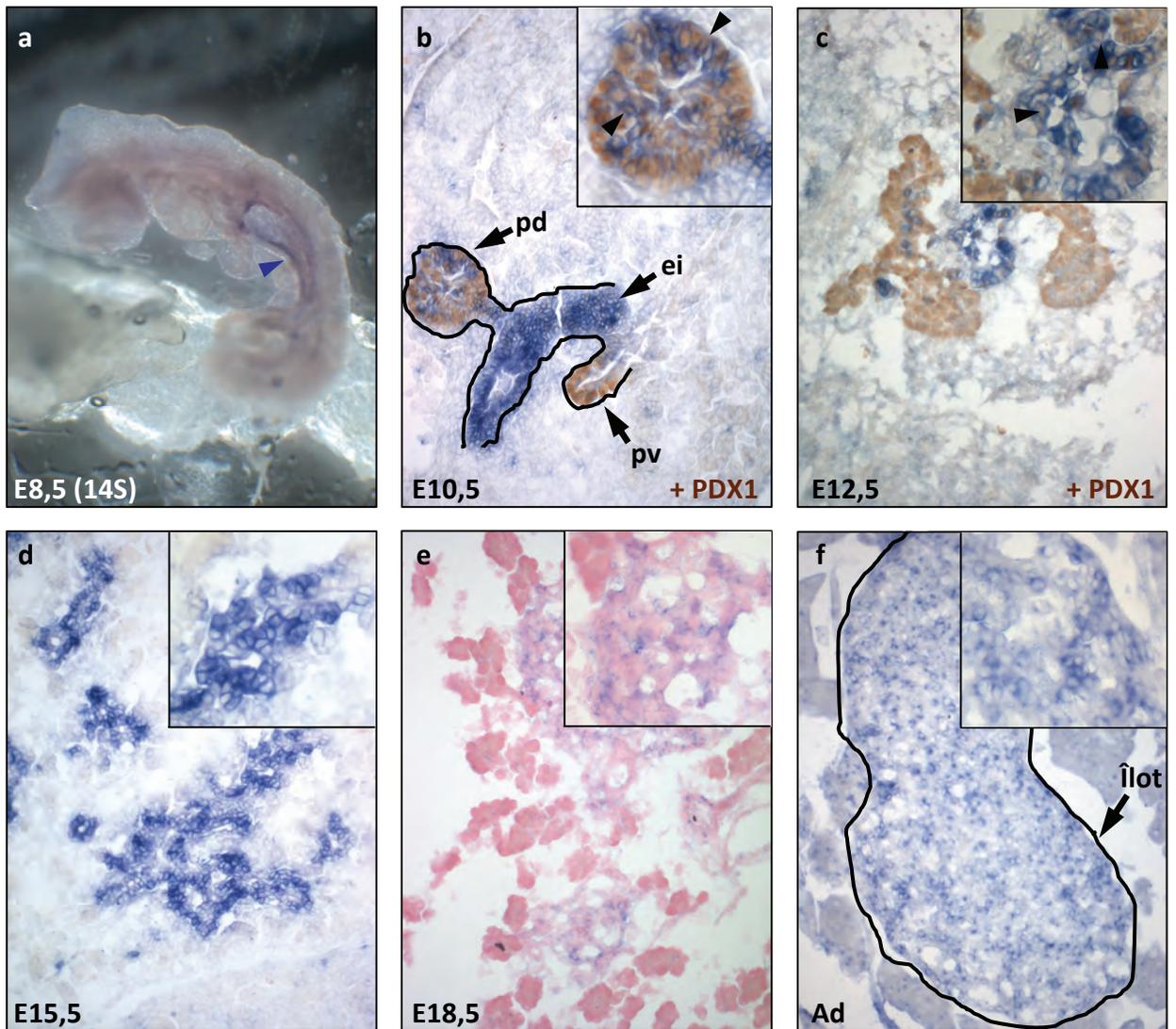


Figure 19 : *Rfx6* est exprimé tout au long du développement pancréatique et dans les îlots de Langerhans. Hybridation *in situ* pour *Rfx6* (bleu, a-f) et immuno-histochimie pour PDX1 (brun, b-c) sur un embryon à E8,5 (14 somites) et sur des sections d'embryons sauvages à E10,5 (b), E12,5 (c), E15,5 (d), E18,5 (e) et de pancréas adulte (f). *Rfx6* est exprimé dans l'endoderme intestinal primitif à E8,5. Son expression se restreint à E10,5 au bourgeon pancréatique ainsi qu'à l'estomac (non montré) et à l'endoderme intestinal. Puis, dans le pancréas, *Rfx6* est exprimé pendant tout le développement embryonnaire, ainsi que dans les îlots de Langerhans. Grossissement: 20X, insert: 40X. Tête de flèche bleue: Cellules exprimant *Rfx6*. Tête de flèche noire: cellule comarquée PDX1 et *Rfx6*. pd : pancréas dorsal, pv : pancréas ventral et ei : endoderme intestinal;

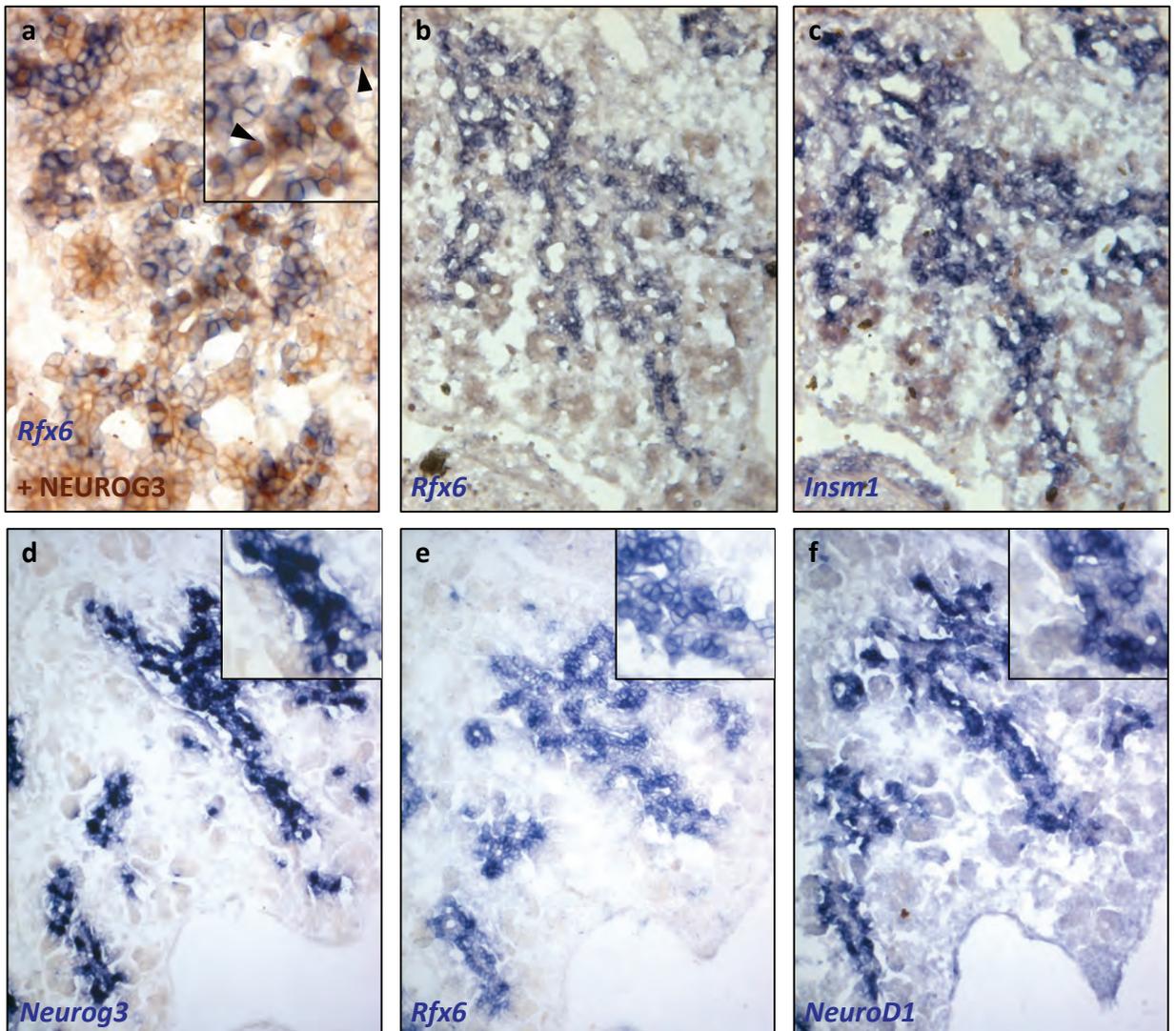


Figure 20 : *Rfx6* est exprimé dans la majorité des cellules exprimant **NEUROG3**. Hybridation *in situ* (bleu) pour *Rfx6* (a,b,e), *Insm1* (c), *Neurog3* (d) et *NeuroD1* (f) et immuno-histochimie pour **NEUROG3** (brun, a) sur des sections adjacentes (b-c, d-f) d'embryon sauvage à E15,5 (a-f). *Rfx6* est exprimé dans la majorité des cellules exprimant **NEUROG3**. Le patron d'expression de *Rfx6* est similaire à celui de *Neurog3* ainsi qu'à deux de ces gènes cibles: *NeuroD1* et *Insm1*. Grossissement: 20X, insert: 40X. Tête de flèche: cellule comarquée **NEUROG3** et *Rfx6*.

de leurs homologues chez le poisson zèbre. Ces protéines sont les seules de la famille RFX à contenir un domaine d'activation transcriptionnelle (Emery et al., 1996). Le dernier groupe est constitué des protéines Rfx4 et Rfx6 murines et de leurs homologues chez le poisson zèbre (Fig. 18B).

L'ensemble de ces résultats montre que la séquence protéique de RFX6 a conservé les domaines protéiques de la famille RFX à l'exception du domaine d'activation transcriptionnelle présent chez les protéines RFX1 à 3.

2. Expression du gène *rfx6*

2.A Cinétique d'expression de *rfx6* au cours du développement pancréatique

Après avoir déterminé la séquence de *Rfx6*, nous avons caractérisé l'expression de *Rfx6* chez les souris en développement et dans le pancréas de souris adulte. Pour cela, j'ai réalisé des hybridations *in situ* sur des embryons entre E8,5 (14 somites) et E18,5 ainsi que sur des sections de pancréas adulte (Fig. 19). J'ai ainsi montré que *Rfx6* est exprimé dès 8,5 jours de développement dans l'endoderme intestinal primitif (Fig. 19a), puis son expression va se restreindre à E10,5 à l'estomac, à l'endoderme intestinal et au pancréas (Fig. 19b). Dans le pancréas à E10,5, son expression est à la fois observée dans des cellules exprimant PDX1 et n'exprimant pas PDX1 (Fig. 19b). Dans le pancréas ventral à E10,5, l'expression de *rfx6* est restreinte à quelques rares cellules. L'expression de *Rfx6* va ensuite s'étendre à E12,5, tout en restant localisée dans la région centrale du pancréas en développement n'exprimant pas PDX1, à l'exception de rares cellules coexprimant PDX1 (Fig. 19c). Puis, tout comme *Neurog3*, son expression va se développer dans l'épithélium pancréatique à E15,5 (Fig. 19d). Par contre, alors que l'expression de *Neurog3* est restreinte à quelques cellules à E18,5 pour être perdue dans le pancréas adulte, l'expression de *Rfx6* persiste à E18,5 et est maintenue dans les îlots de Langerhans dans le pancréas adulte (Fig. 19e-f).

2.B *Rfx6* est spécifiquement exprimé dans le lignage endocrine du pancréas

Au vu de ces premières observations, nous avons souhaité préciser le patron d'expression de *Rfx6* dans le pancréas. Pour cela, j'ai réalisé des comarquages *Rfx6* et NEUROG3, et des hybridations *in situ* sur des sections adjacentes de pancréas à E15,5 pour *Rfx6*, *Neurog3*, et *Insm1* et *NeuroD1*, deux gènes cibles directs de NEUROG3 (Fig. 20). Nous

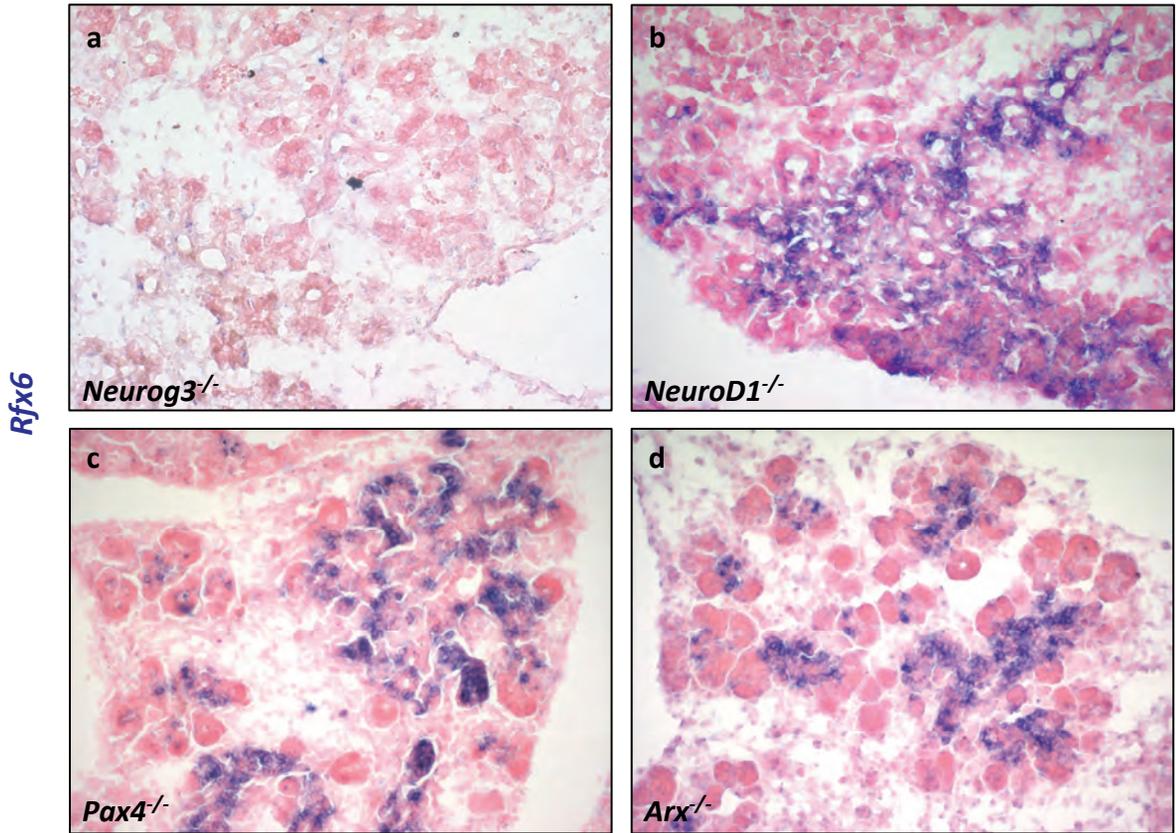


Figure 21 : *Rfx6* est spécifiquement exprimé dans les cellules du pancréas endocrine. Hybridation *in situ* pour *Rfx6* (a-d) sur des sections de souris *Neurog3*^{-/-} (a) à E15,5, *NeuroD1*^{-/-} (b), *Pax4*^{-/-} (c) et *Arx*^{-/-} (d) à E14,5. L'expression de *Rfx6* est perdue chez le mutant *Neurog3*, et indépendante de l'expression de *NeuroD1*, *Pax4* et *Arx*, trois gènes exprimés en aval de *Neurog3* et impliqués dans la différenciation endocrine. Grossissement: 20X.

avons ainsi montré que la majorité des cellules exprimant NEUROG3 expriment *Rfx6* (Fig. 20a). J'ai également mis en évidence que l'expression de *Rfx6* est localisée dans un territoire similaire où sont exprimés *Insm1* et *NeuroD1*, deux gènes qui ont été caractérisés comme étant spécifiques du lignage pancréatique endocrine (Fig. 20d-f). L'ensemble de ces résultats suggère donc que l'expression de *rfx6* est restreinte au lignage endocrine.

Afin d'étudier cette hypothèse, j'ai réalisé des hybridations *in situ* sur des sections de pancréas *Neurog3*^{-/-}, qui sont caractérisés par la perte du lignage endocrine (Gradwohl et al., 2000 ; Fig. 21a). De plus, afin de déterminer si *Rfx6* est spécifiquement requis pour la différenciation des cellules α (sécrétant le glucagon) ou des cellules β (sécrétant l'insuline), nous avons étudié l'expression de *Rfx6* chez des embryons mutants *Pax4* ou *Arx*, qui sont caractérisés respectivement par une perte de la différenciation des cellules β , et une perte de la différenciation des cellules α (Sosa-Pineda et al., 1997 ; Collombat et al., 2003 ; Fig. 21c-d). Finalement, afin de déterminer si l'expression de *Riken L19* est sous le contrôle de NEUROD1, le seul gène cible connu pour NEUROG3 au début de cette étude, nous avons étudié l'expression de *Riken L19* chez des embryons mutants pour *NeuroD1*, dont la perte de fonction induit une réduction de la différenciation des différents sous-types cellulaires endocriniens (Huang et al., 2000 ; Naya et al., 1997 ; Fig. 21b)

Nous avons ainsi montré que l'expression de *Rfx6* est perdue dans le pancréas des souris *Neurog3* mutantes, alors que son expression est conservée dans les mutants *NeuroD1*, *Pax4* et *Arx* (Fig. 21).

L'ensemble de ces résultats montre que *Rfx6* est spécifiquement exprimé dans le lignage endocrine dans le pancréas en aval de *Neurog3*, et de façon indépendante de *NeuroD1*, *Pax4* et *Arx*, trois gènes déjà impliqués dans la différenciation des sous types cellulaires pancréatiques endocriniens. Afin de poursuivre la caractérisation de l'expression de *Rfx6*, un anticorps dirigé contre la partie N-terminale de la protéine RFX6 a été produit chez le lapin et le rat au sein de l'IGBMC.

2.C Expression précoce de RFX6 dans l'endoderme intestinal primitif et l'épithélium pancréatique

Suite à la caractérisation de l'expression de *Rfx6* à E8,5 et à E10,5 (Fig. 19), nous avons continué la caractérisation de l'expression précoce de *Rfx6* entre E8,5 et E10,5 afin de

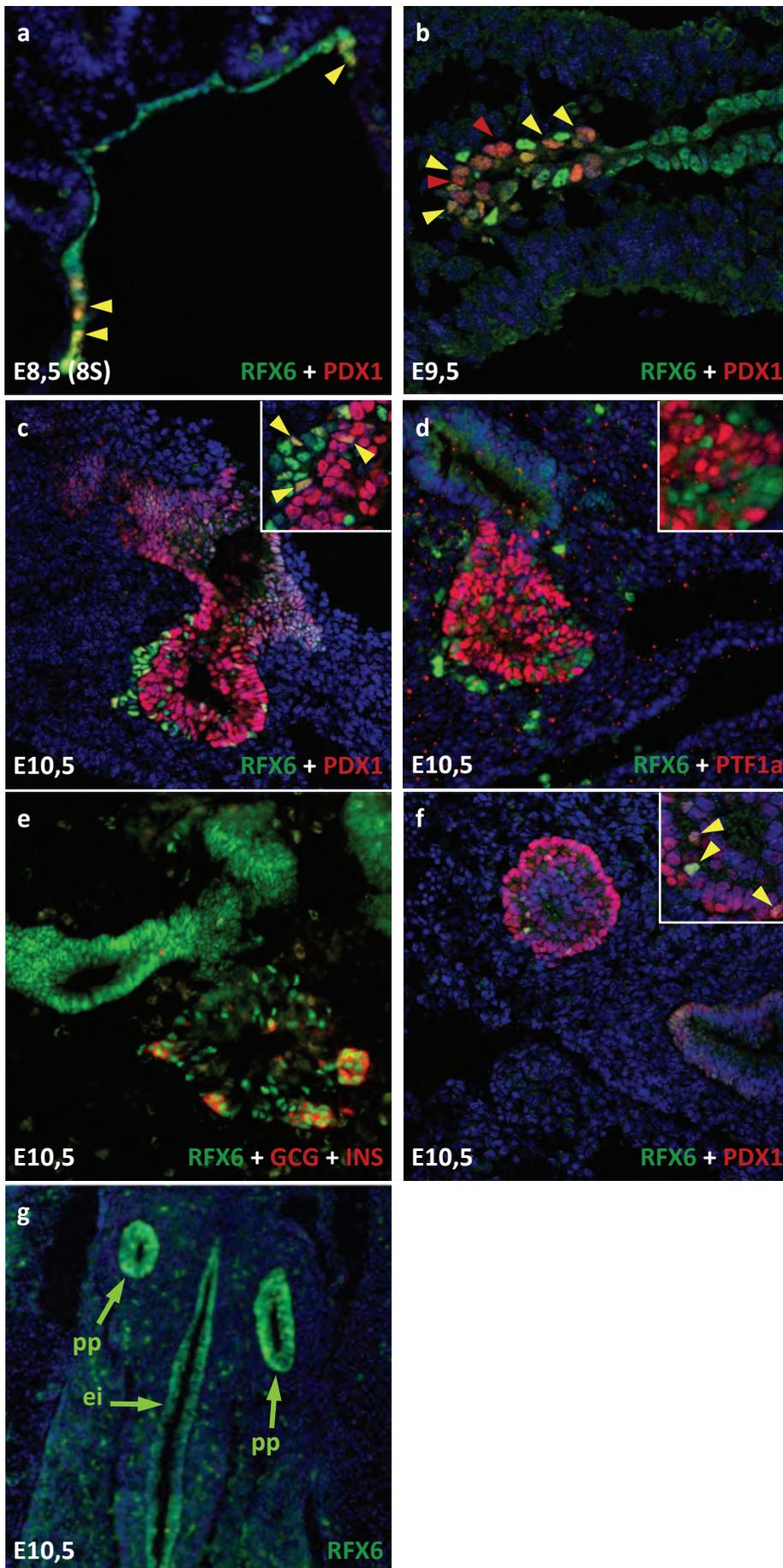


Figure 22 : Expression précoce de RFX6. Immunofluorescence pour RFX6 (vert, a-g) et PDX1 (rouge, a-c, f) ou PTF1a (rouge, d) ou glucagon + insuline (rouge, e) sur des sections d'embryon sauvage à E8,5 (8 somites, a), à E9,5 (b), à E10,5 (c-e, g) et sur des sections d'embryon *neurog3^{-/-}* à E10,5 (f). RFX6 est exprimé dans l'ensemble de l'épithélium intestinal primitif à E8,5. Puis dans les bourgeons pancréatiques, RFX6 n'est pas exprimé dans les progéniteurs pancréatiques (PTF1a positifs). Puis RFX6 est exprimé dans les cellules engagées dans la différenciation endocrine et les cellules exprimant le glucagon et l'insuline. Grossissement: 20X (a, c-f) et 40X (b), insert: 40X. Tête de flèche jaune: cellule comarquée PDX1 et RFX6, tête de flèche rouge: cellule comarquée RFX6 négative et PDX1 positive. ei : endoderme intestinal, pp : primordium pulmonaire.

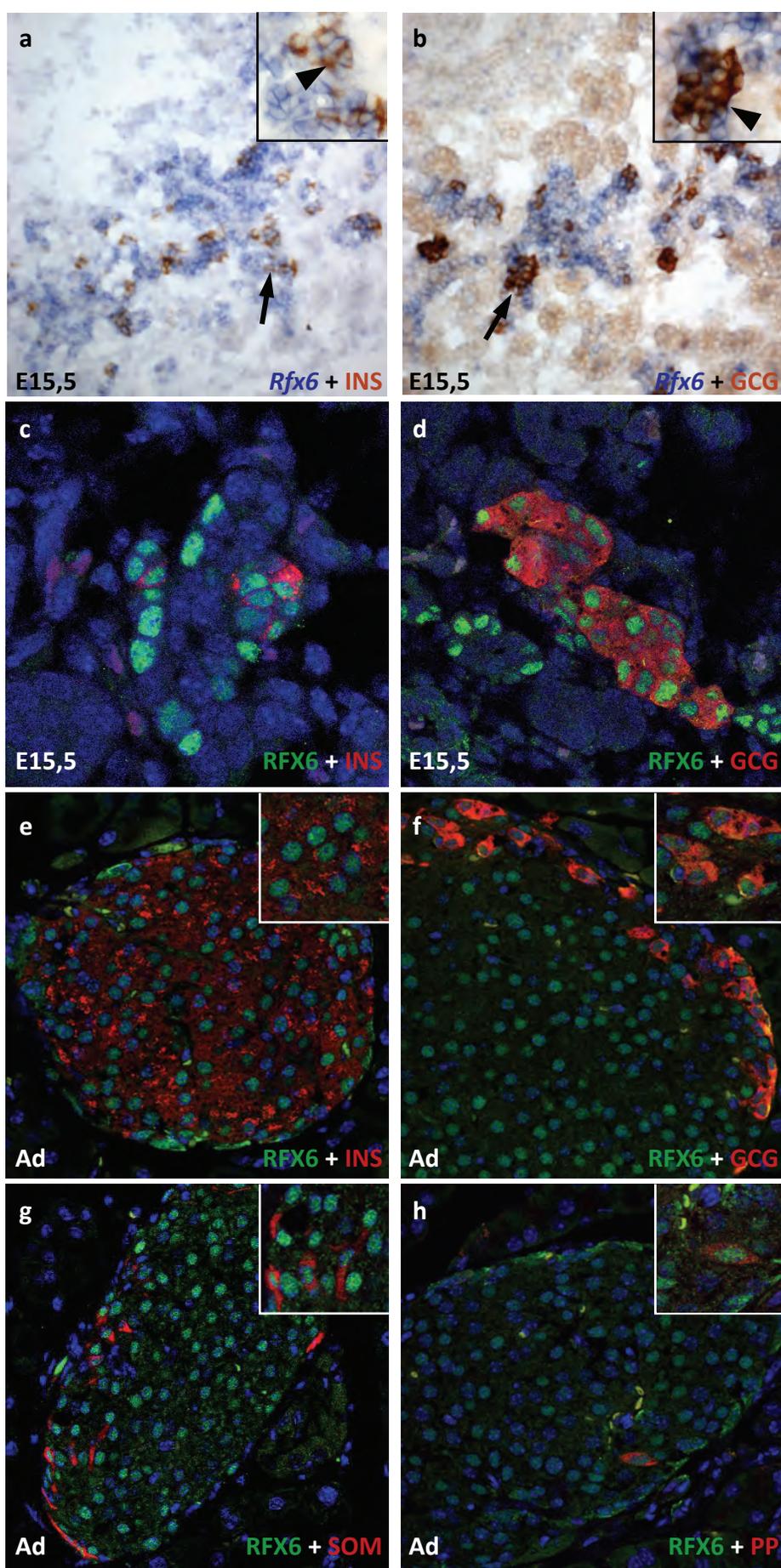


Figure 23 : RFX6 est exprimé dans les différents sous types cellulaires pancréatiques endocriniens. Hybridation *in situ* pour *Rfx6* (bleu, a-b) et immuno-histochimie pour l'insuline (brun, a) ou le glucagon (b) à E15,5. Immunofluorescence pour RFX6 (c-h) et l'insuline (c, e), le glucagon (d, f), la somatostatine (g) et le polypeptide pancréatique (h) sur des sections d'embryon sauvage à E15,5 (a, b) et de pancréas adulte (c, f). RFX6 est exprimé dans les cellules α et β à E15,5. Dans les îlots de Langerhans, RFX6 est exprimé dans les quatre principaux sous types cellulaires endocriniens. Grossissement: 20X (a-b), 63X (c-d) et 40X (e-h), insert: 40X (a-b) et 63X (e-h) ; Ad : adulte.

déterminer précisément dans quelles cellules RFX6 est exprimé à cette période. Pour cela, j'ai réalisé par immunofluorescence des doubles marquages pour RFX6 et PDX1, RFX6 et PTF1a ou RFX6 et le glucagon et l'insuline chez des embryons de souris entre E8,5 (8 somites, 8 S) et E10,5 (Fig. 22). Nous avons ainsi confirmé l'expression précoce de *Rfx6* dans l'endoderme intestinal primitif dès E8,5 (8 S), où son expression est détectée de façon large dans l'endoderme intestinal primitif (Fig. 22a). A ce stade, nous avons également montré que RFX6 est exprimé dans le territoire des futurs bourgeons pancréatiques exprimant PDX1 (Fig. 22a). A E9,5, RFX6 est principalement exprimé dans les cellules PDX1 positives dans le bourgeon pancréatique dorsal (Fig. 22b). Toutefois, on observe également quelques cellules RFX6 positives et PDX1 négatives, ainsi que quelques cellules RFX6 négatives et PDX1 positives (Fig. 22b). Un jour plus tard, à E10,5, la coexpression de RFX6 avec PDX1 est quasiment perdue à l'exception de quelques cellules (Fig. 22c). On peut également noter que les progéniteurs pancréatiques exprimant PTF1a sont tous RFX6 négatif (Fig. 22d). A E10,5 dans le bourgeon pancréatique dorsal, les cellules RFX6 sont soit dispersées, soit groupées en petits agrégats qui expriment le glucagon et l'insuline (Fig. 22e). A E10,5, on peut également noter que RFX6 est exprimé dans l'endoderme intestinal et dans les primordiums pulmonaires (Fig. 22g).

Ces résultats montrent donc que dès E8,5 (8S), RFX6 est exprimé de façon précoce dans l'endoderme intestinal primitif, avant l'expression de *Neurog3* qui n'est pas exprimé avant E9,0 (Apelqvist et al., 1999). Puis, à E10,5, l'expression de RFX6 est exclue des progéniteurs pancréatiques exprimant PTF1a mais est exprimé dans les cellules exprimant le glucagon et l'insuline. On peut également noter que l'expression de RFX6 est maintenue entre E9,5 et E10,5 dans l'endoderme intestinal primitif qui formera l'intestin. Ces résultats suggèrent donc que *Rfx6* aurait une expression précoce dans l'endoderme intestinal primitif indépendante de celle dans les cellules pancréatiques endocrines en différenciation. On peut ainsi supposer que l'expression de RFX6 détectée dans quelques rares cellules à E10,5 chez les mutants *Neurog3* (Fig. 22f) correspond à des cellules qui n'ont pas encore exprimé PTF1a, ayant ainsi encore les caractéristiques de l'endoderme intestinal primitif.

2.D RFX6 est exprimé dans l'ensemble des sous types cellulaires insulaires

Afin de déterminer si RFX6 est exprimé dans les cellules endocrines matures, nous avons ensuite analysé l'expression de RFX6 par rapport à l'expression des hormones

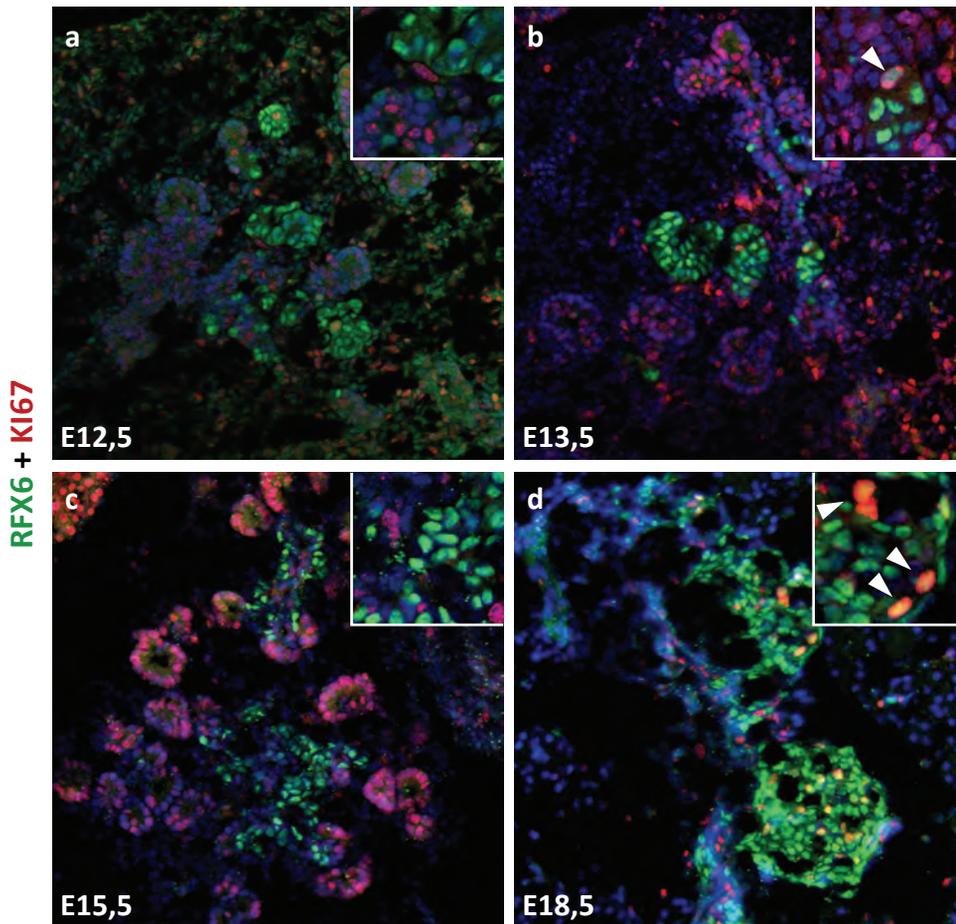


Figure 24 : Les cellules RFX6 ne prolifèrent pas pendant la différenciation des cellules du pancréas endocrine. Immunofluorescence pour RFX6 (rouge, a-d) et KI67 (vert, a-d) sur des sections d'embryon sauvage à E12,5 (a), E13,5 (b), E15,5 (c) et E18,5 (d). RFX6 n'est pas exprimé dans les cellules KI67 positives de E12,5 à E15,5. A E18,5, RFX6 est exprimé dans les cellules KI67 positives dans les îlot en formation. Grossissement: 20X, insert: 40X. Tête de flèche: cellule comarquée KI67 et RFX6.

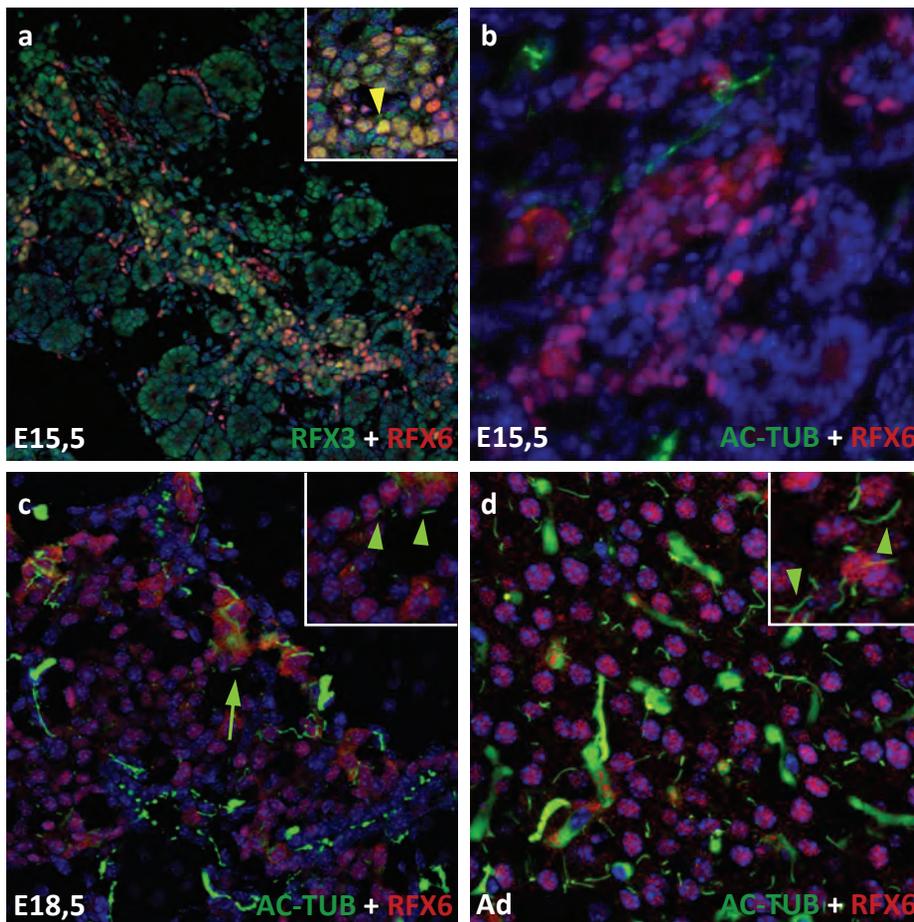


Figure 25 : Les cellules RFX6 dans les îlots de Langerhans sont ciliées. Immunofluorescence pour RFX6 (rouge, a-d) et RFX3 (vert, a) ou AC-TUB (vert, b-d) sur des sections d'embryon sauvage à E15,5 (a-b), E18,5 (c) et de pancréas adulte (d). La majorité des cellules Rfx6 sont Rfx3 positives. A E15,5, aucun cil n'est observé sur les cellules du pancréas. A E18,5, les premiers cils apparaissent et, dans les îlots de Langerhans, l'ensemble des cellules RFX6 sont ciliées. Grossissement: 20X (a-b), 40X (c) et 63X (d), insert: 40X (a), 63X (c) et 100X (d). Tête de flèche jaune: cellule comarquée RFX3 et RFX6; tête de flèche verte: cellule comarquée AC-TUB et RFX6.

pancréatiques au cours du développement pancréatique et dans les îlots de Langerhans. Pour cela, j'ai réalisé des hybridations *in situ* pour *Rfx6* suivies d'immuno-histochimie pour l'insuline ou le glucagon à E15,5 (Fig. 23a-b). J'ai également réalisé des immunofluorescences pour RFX6 et l'insuline ou le glucagon à E15,5, et pour RFX6 et l'insuline, le glucagon la somatostatine ou le polypeptide pancréatique sur des sections de pancréas adulte (Fig. 23c-h). Nous avons montré qu'à E15,5, les cellules exprimant l'insuline (Fig. 23c) ou le glucagon (Fig. 23d) coexpriment RFX6. Toutefois, la majorité des cellules *Rfx6* n'expriment ni l'insuline, ni le glucagon à ce stade (Fig. 23a-b). Dans les îlots de Langerhans, l'ensemble des quatre principaux sous types cellulaires endocriniens exprime *Rfx6* (Fig. 23e-h).

2.E Les cellules *Rfx6* ne prolifèrent pas

Comme les cellules NEUROG3 sont principalement postmitotiques, j'ai réalisé des doubles immunofluorescences pour RFX6 et Ki67, un marqueur des cellules en prolifération en phase G1, M, G2 et S afin de déterminer si les cellules RFX6 prolifèrent (Fig. 24). Entre E12,5 et E15,5, les cellules *Rfx6* sont pour la plupart quiescentes (Fig. 24a-c). Seules quelques très rares cellules sont doubles marquées *Rfx6* et Ki67. (Fig. 24a-c). En revanche, à E18,5, stade auquel commence la formation des îlots de Langerhans, la majorité des cellules exprimant l'antigène KI67 expriment RFX6 (Fig. 24d).

2.F Les cellules *Rfx6* dans les îlots de Langerhans sont ciliées

Des travaux récents ont mis en évidence l'implication du gène *Rfx3* dans la différenciation des cellules insulaires. En effet, Ait-Lounis et al. ont montré que chez les mutants *Rfx3*, la différenciation des cellules α , β et δ est fortement réduite. De plus, les souris *Rfx3*^{-/-} sont caractérisées par une réduction et une malformation des cils primaires présents sur les cellules insulaires (Ait-Lounis et al., 2007). Afin d'étudier l'expression de RFX6 vis-à-vis de RFX3, j'ai réalisé des doubles immunofluorescences pour RFX3 et RFX6 à E15,5 (Fig. 25a). Nous avons ainsi montré que les cellules RFX6 sont pour la plupart comarquées avec RFX3 à E15,5. En revanche, bien que l'anticorps anti RFX3 développe du bruit de fond, le patron d'expression de RFX3 est plus large que celui de RFX6. En effet, bien que l'expression de RFX3 soit plus intense chez les cellules exprimant RFX6, RFX3 est également exprimé dans l'épithélium qui formera le tissu acinaire (Fig. 25a). L'analyse

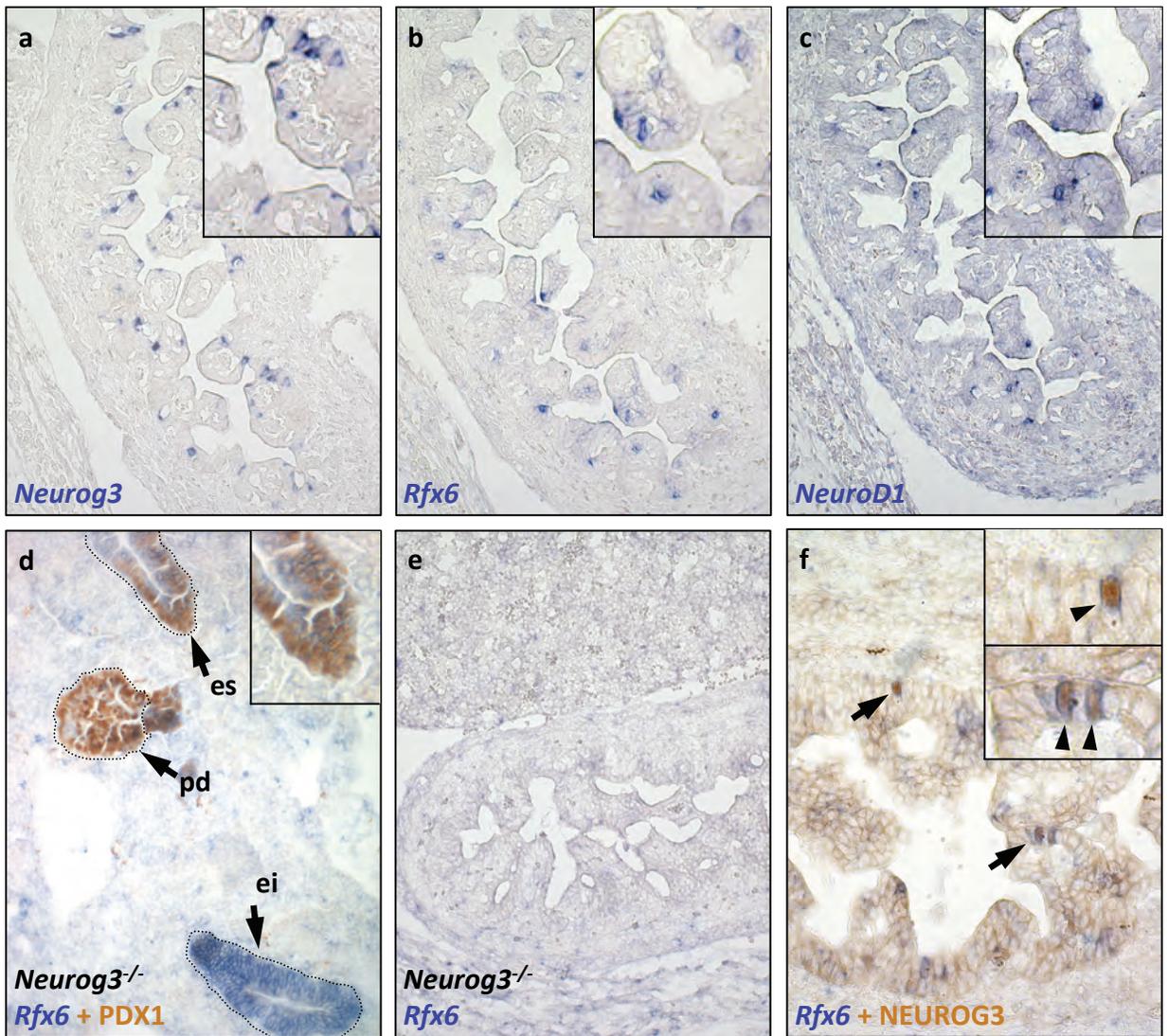


Figure 26 : Expression de *Rfx6* dans l'intestin embryonnaire. Hybridation *in situ* (bleu) pour *Neurog3* (a), *Rfx6* (b, d-f) et *NeuroD1* (c) et immuno-histochimie pour PDX1 (brun, d) et pour NEUROG3 (f) à E10,5 (d) et E15,5 (a-c, e-f) sur des sections d'embryon sauvage (a-c, f) ou *Neurog3*^{-/-}(d-e). A E15,5, *Rfx6* est exprimé dans les cellules intestinales exprimant NEUROG3 et son expression est perdue chez le mutant *Neurog3*. Grossissement: 20X (a-e), 40X (f), insert; 40X (a-c), zoom (f). Tête de flèche: cellule *Rfx6* et NEUROG3. pd : pancréas dorsal, ei : endoderme intestinal et es : estomac.

transcriptomique des mutants *Neurog3* montre également que *Rfx3* est toujours exprimé chez les mutants *Neurog3*. Ces résultats montrent que l'expression de *Rfx3* n'est pas restreinte au lignage pancréatique endocrine.

Nous avons également étudié la présence de cils sur les cellules exprimant RFX6 par des immunofluorescences pour RFX6 et la tubuline acétylée, un marqueur des cils (Fig. 25b-d). La présence de cils sur les cellules RFX6 n'est pas détectée à E15,5. En revanche, quelques cils sont observés à E18,5 (Fig. 25c). Dans les îlots de Langerhans, l'ensemble des cellules insulaires exprimant *Rfx6* sont ciliées (Fig. 25d).

2.G *Rfx6* est exprimé dans l'estomac et l'intestin

Comme montré précédemment, *Rfx6* est exprimé dans l'endoderme primitif intestinal qui formera le tube digestif (Fig. 19b). Afin de déterminer si l'expression de *Rfx6* est comme dans le pancréas, maintenue dans l'épithélium intestinal en différenciation, j'ai réalisé des hybridations *in situ* pour *Neurog3*, *Rfx6* et *NeuroD1* sur des sections sériées d'embryon sauvage à E15, 5 (Fig. 26 a-c). Nous avons ainsi montré que comme *Neurog3*, qui est exprimé dans l'épithélium intestinal à E15,5 (Fig. 26a), *Rfx6* est exprimé dans quelques cellules épithéliales intestinales (Fig. 26b). Son expression semble également similaire à celle de *Neurod1* (Fig. 26c) qui a une expression plus restreinte à E15,5. Nous avons également souhaité définir si comme dans le pancréas, l'expression de *Rfx6* est perdue chez le mutant *Neurog3*. Pour cela, des hybridations *in situ* pour *Rfx6* ont été réalisées sur des sections d'embryons *Neurog3* mutants à E10,5 et E15,5, suivi à E10,5 d'une immuno-histochimie pour PDX1 (Fig. 26d-e). A E10,5, *Rfx6* est exprimé dans l'endoderme intestinal qui formera le tube digestif chez les mutants *Neurog3* (Fig. 26d). Le maintien de l'expression de *Rfx6* dans l'endoderme intestinal chez les mutants *Neurog3* à E10,5 concorde avec les précédentes observations de Jenny et al. montrant que l'expression de *Neurog3* n'est pas détectée avant E12,5 dans l'intestin (Jenny et al., 2002). Ces résultats montrent qu'à E10,5 l'expression de *Rfx6* est indépendante de l'expression de *Neurog3* dans l'intestin. En revanche à E15,5, la perte d'expression de *Rfx6* chez le mutant *Neurog3* montre qu'à ce stade l'expression de *Rfx6* est dépendante de celle de *Neurog3* (Fig. 26e). De plus, par hybridations *in situ* pour *Rfx6* sur des sections d'embryons sauvages à E15,5, suivis d'une immuno-histochimie pour NEUROG3, j'ai montré que les cellules NEUROG3 coexpriment *Rfx6* (Fig. 26f).

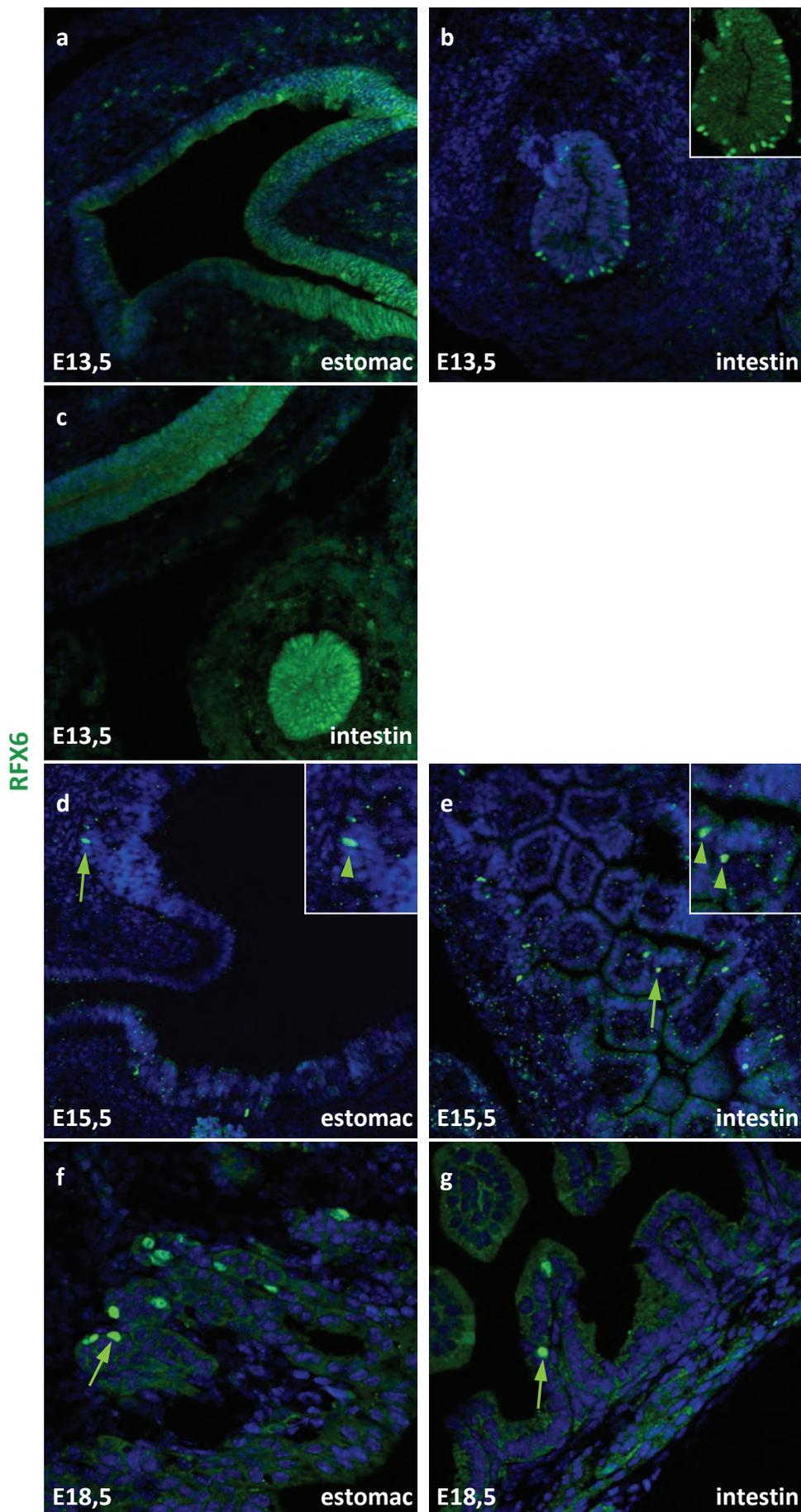
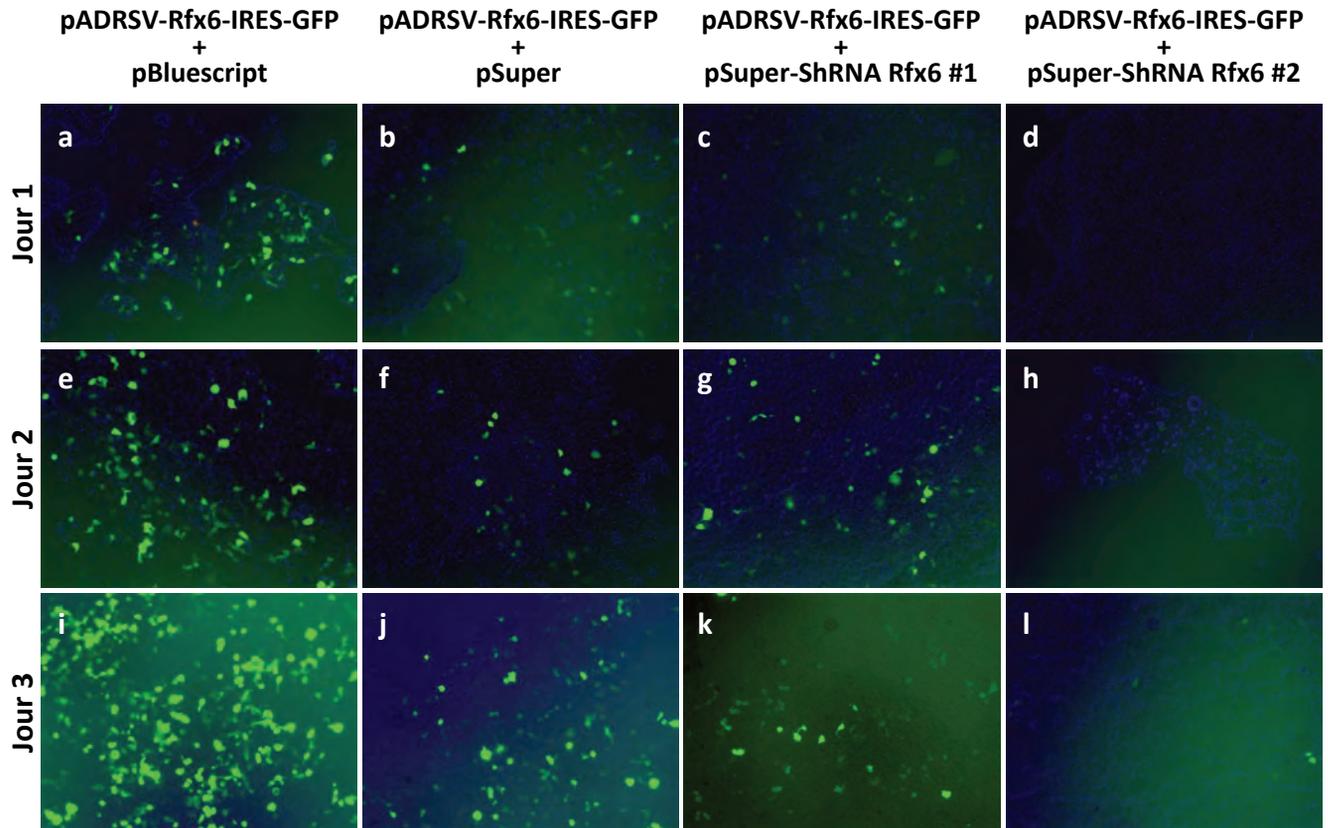


Figure 27 : RFX6 est exprimé dans l'épithélium intestinal. Immunofluorescence pour RFX6 (vert, a-g) sur des sections d'embryon sauvage : estomac (a, d, f) et intestin (b-c, e, g) à E13,5 (a-c), E15,5 (d-e) et E18,5 (f-g). A E13,5, RFX6 est exprimé faiblement dans l'endoderme intestinal pluristratifié, sauf quelques cellules exprimant fortement RFX6. A partir de E15,5, son expression va se restreindre à quelques cellules dans l'épithélium stomacal et intestinal. Grossissement: 20X (a-e), 40X (f-g). Tête de flèche: cellule RFX6, insert in b = marquage RFX6, insert in d-e = zoom.

Afin de préciser ces résultats, j'ai réalisé des immunofluorescences pour RFX6 sur des sections d'estomac et d'intestin entre E13,5 et E18,5 (Fig. 27). A E13,5, stade auquel la différenciation antéropostérieure de l'endoderme intestinal commence (Henning, 1987), on peut observer un faible signal pour RFX6 dans la région du futur duodénum, adjacente à l'estomac, ainsi que l'expression forte de RFX6 dans quelques rares cellules (Fig. 27b). Dans l'estomac et la partie postérieure de l'épithélium intestinal à E13,5, RFX6 est exprimé dans l'ensemble de l'épithélium de manière homogène (Fig. 27a, c). En revanche, à partir de E15,5, l'expression de RFX6 dans l'épithélium gastrique et intestinal est restreinte à quelques cellules (Fig. 27d-g). Ces résultats confirment les résultats précédemment obtenus par hybridation *in situ* (Fig. 26b).

L'ensemble de ces résultats montre que l'expression de *Rfx6* évolue au cours de la différenciation du pancréas et de l'intestin. Pour résumer le patron d'expression de RFX6, RFX6 est dans un premier temps exprimé dans l'endoderme intestinal primitif, et dans les territoires présomptifs du pancréas exprimant PDX1. Puis, son expression va progressivement être perdue dans les cellules PDX1 lors de la formation des bourgeons pancréatiques. Et l'expression de RFX6 sera totalement perdue dans les bourgeons pancréatiques exprimant PTF1a. RFX6 est ensuite réexprimé dans les progéniteurs endocrines (NEUROG3 positifs). Puis son expression est conservée dans les cellules endocrines en différenciations (*Insm1* ou *NeuroD1* positives) et dans les cellules insulaires (glucagon, insuline, somatostatine ou PP positives). De plus, il est à noter que l'expression de *Rfx6* est endocrine spécifique au cours de la formation du pancréas (expression de *Rfx6* perdue chez les mutants *Neurog3*). Et que son expression est indépendante de celle d'Arx et *Pax4*, impliqués respectivement dans la différenciation des sous types cellulaires α et β . Au cours de la différenciation gastrique et intestinale, RFX6 semble être dans un premier temps exprimé de façon indépendante de *Neurog3* dans l'ensemble de l'endoderme qui formera l'intestin. Dans un second temps, RFX6 est exprimé sous le contrôle de NEUROG3 dans quelques cellules de l'épithélium gastrique et intestinal.

A



B

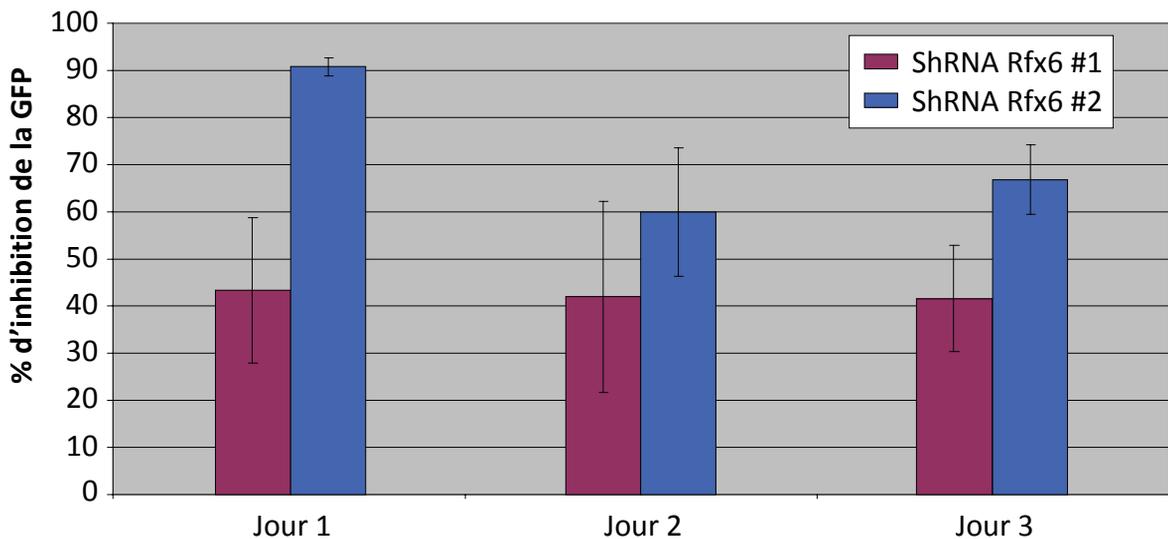


Figure 28 : Le ShRNA Rfx6 #2 induit une dégradation du transcrite *Rfx6-ires-GFP*. (A) Observation directe de la GFP dans les cellules HEK293t transféctées avec le plasmide pADRSV-Rfx6-IRES-GFP et pBluescript (a, e, i), pSuper (b, f, j), pSuper-ShRNA Rfx6 #1 (c, g, k), et pSuper-ShRNA Rfx6 #2 (d, h, l), 1, 2 ou 3 jours après la transfection. (B) Quantification de l'inhibition de la GFP dans les cellules transféctées avec les plasmides pSuper-ShRNA Rfx6 #1 ou #2 par rapport aux cellules transféctées avec le plasmide pSuper-ShRNA. Le ShRNA Rfx6 #2 conduit a une meilleure dégradation de l'ARN bicistrionique *Rfx6-ires-GFP* que le ShRNA Rfx6 #1.

3. Caractérisation de la fonction de *rfx6* au cours du développement pancréatique murin

3.A Caractérisation des ShRNA Rfx6

Au vu de ces résultats, nous avons souhaité caractériser la fonction de Rfx6 au cours du développement pancréatique. Pour cela, nous avons choisi d'infecter des explants d'épithélium pancréatique à E11,5 avec un lentivirus permettant l'expression d'un ARN antisens (ShRNA) dirigé contre *Rfx6*.

La conception des ShRNA contre *rfx6* a été faite en utilisant le logiciel *siDESIGN*[®] center de Dharmacon. La région de l'ARNm utilisée pour définir les ShRNA est comprise entre les nucléotides 1170 et 2685 de la séquence décrite figure 16. Nous avons ainsi choisi deux ShRNA. Le ShRNA Rfx6 #1 et le ShRNA Rfx6 #2 qui ciblent respectivement les séquences comprises entre les nucléotides 1677 à 1699 et les nucléotides 2145 à 2167.

Les séquences codant les ShRNA Rfx6 #1 et le ShRNA Rfx6 #2 ont été clonées dans le vecteur pSuper contenant le promoteur de la polymérase 3 (H1) afin de permettre l'expression du ShRNA. En parallèle, les fragments d'ARNm contenant les nucléotides 1166 à 2085 ou 1829 à 2685 ont été clonés dans le vecteur pADRSV-IRES-GFP. Ce vecteur pADRSV-Rfx6-IRES-GFP permet l'expression d'un ARN bicistronique comprenant l'ARN ciblé par les ShRNArfx6 et l'ARN codant pour la GFP. Afin de caractériser ces ShRNA, les plasmides permettant l'expression des ShRNA ont été cotransfectés dans les cellules HEK293t avec un plasmide pADRSV-Rfx6-IRES-GFP. Ainsi, si le ShRNA est spécifique de la séquence pour laquelle il a été défini, il devrait induire une dégradation de l'ARNm bicistronique, et par conséquent induire une diminution de l'expression de la GFP.

Afin de déterminer la diminution de l'expression de la GFP dans les cellules HEK293t transfectées avec les plasmides pADRSV-Rfx6-IRES-GFP et pSuper-ShRNA Rfx6 (Fig. 28Ac-d, g-h, k-l) par rapport aux cellules transfectées avec les plasmides pADRSV-Rfx6-IRES-GFP et pSuper vide (Fig. 28Ab, f, j), les cellules ont été analysées par cytométrie de flux entre 1 et 3 jours après la transfection (Fig. 28B). Nous avons ainsi montré que le ShRNA Rfx6 #2 induit une plus forte dégradation de la GFP par rapport au ShRNA Rfx6 #1. Un jour après transfection, le ShRNA Rfx6 #2 induit une dégradation de 91 % contre 43 % pour le ShRNA Rfx6 #1 ; et trois jours après transfection, le ShRNA rfx6#2 induit une dégradation de 67 % contre 42 % pour le ShRNA Rfx6 #1 (Fig. 28B).

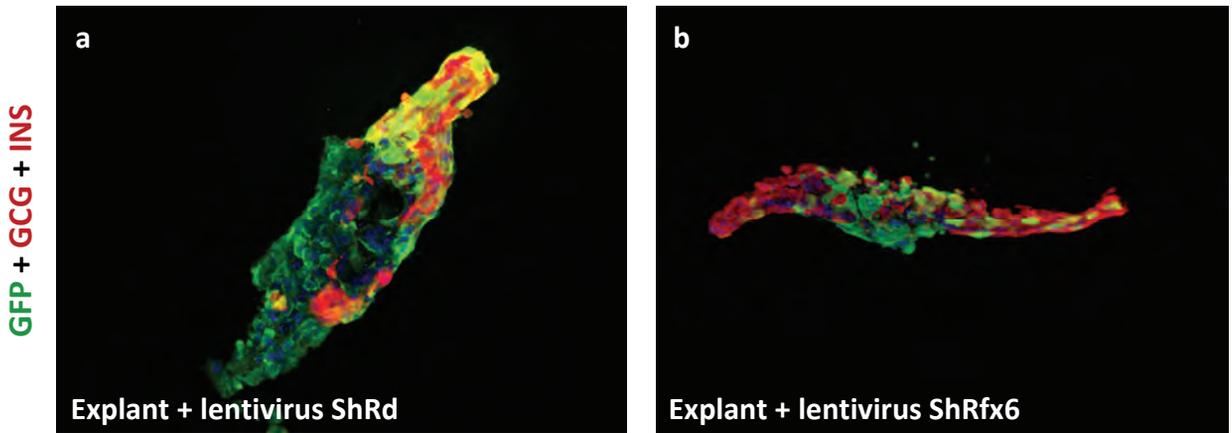
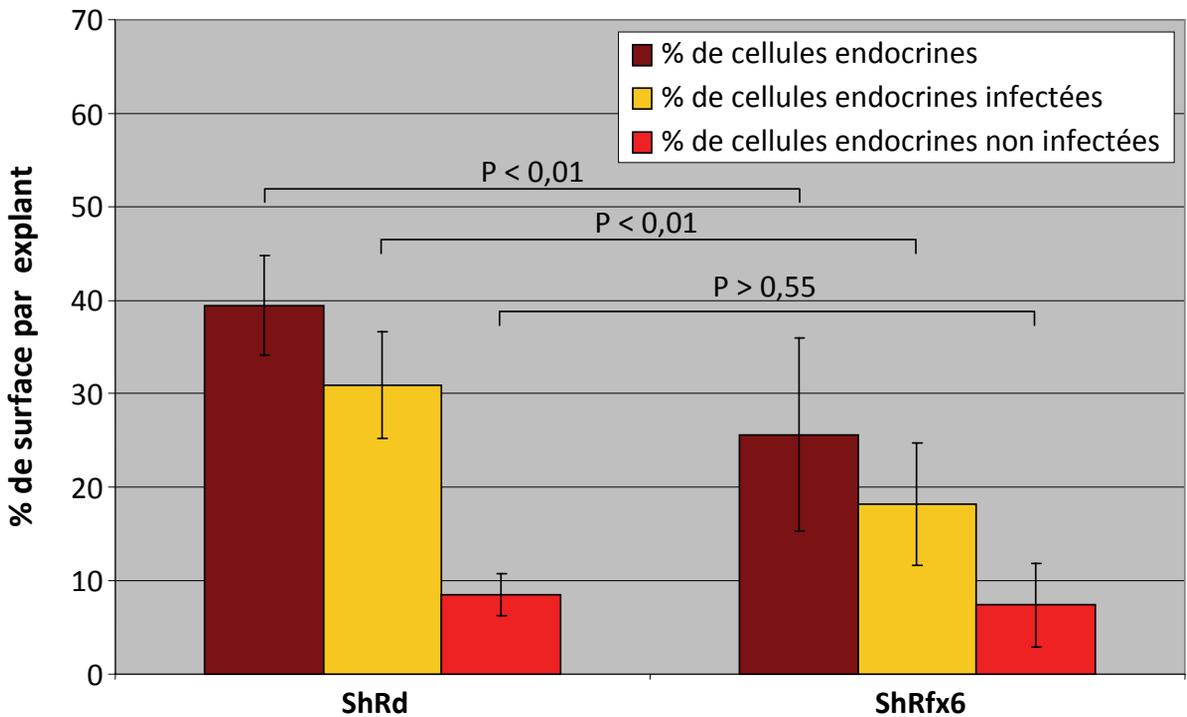
A**B**

Figure 29 : Le Shrfx6 induit une diminution de la différenciation endocrine. (A) 7 jours après l'infection avec le lentivirus exprimant le ShRd (a) ou le ShRfx6 (b) d'explants disséqués à E11,5, immunofluorescence pour l'insuline (INS) et le glucagon (GCG, Rouge) et la GFP (vert) sur des sections d'explants. Grossissement : 20x. (B) 7 jours après l'infection avec le lentivirus exprimant le ShRd (n=5) ou le ShRfx6 (n=8), quantification de la surface de cellules exprimant l'insuline et le glucagon et la GFP dans des explants disséqués à E11,5. Le Shrfx6 induit une diminution de la différenciation des cellules glucagons ou insulines positives par rapport au ShRd chez les explants infectés.

Pour le reste de cette étude, nous avons donc utilisé le ShRNA Rfx6 #2 (ShRfx6) qui permet d'induire une meilleure dégradation de l'ARNm de *Rfx6* pour étudier la perte de fonction de *Rfx6*. Afin de pouvoir utiliser ce ShRNA sur des cultures d'épithélium pancréatique, j'ai sous cloné le ShRfx6 dans un vecteur permettant la production de lentivirus exprimant le ShRfx6.

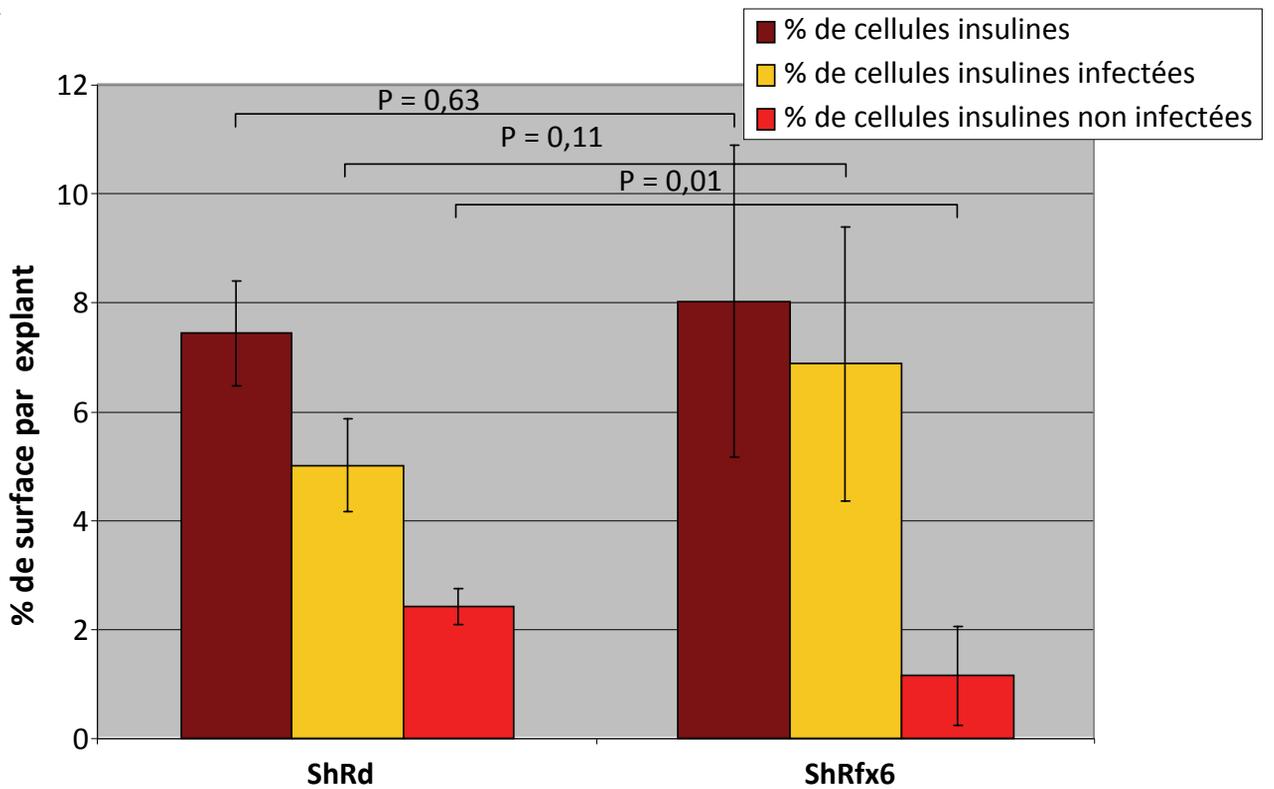
3.B Etude de la perte de fonction de *Rfx6* sur la différenciation endocrine

L'étude de la perte de fonction de *Rfx6* dans le pancréas en développement a été réalisée sur des cultures d'épithélia pancréatiques disséqués à E11,5, puis infectés par un lentivirus. Ces épithélia ont été soit infectés avec un lentivirus permettant l'expression du ShRfx6, soit avec un lentivirus permettant l'expression d'un ShRNA non spécifique (ShRd). De plus, les lentivirus permettant l'expression du ShRd ou ShRfx6 expriment la GFP sous le contrôle d'un promoteur CMV, permettant ainsi de suivre et quantifier le taux d'infection des épithélia.

Nous avons analysé dans des épithélia pancréatiques 7 jours après leur infection, les impacts de l'expression des ShRNA sur la différenciation endocrine. Pour cela, 7 jours après l'infection, les explants sont inclus et coupés afin de réaliser une analyse par immunofluorescence de l'expression du glucagon, de l'insuline et de la GFP (Fig. 29Aa-b). Grâce à une analyse informatique des images, la surface des différents marquages a ensuite été quantifiée (Fig. 29B). Afin de pouvoir étudier un effet dû à l'infection des épithélia par les lentivirus, seuls les explants infectés à plus de 65% de leur surface totale ont été analysés. Cette analyse a été réalisée sur trois expériences différentes. Nous avons ainsi mis en évidence que les épithélia infectés avec le lentivirus exprimant le ShRfx6 sont caractérisés par une baisse de 34 % du nombre de cellules à glucagon ou à insuline par rapport aux épithélia infectés avec le lentivirus exprimant le ShRd (Fig. 29b). De plus, afin de confirmer ce résultat, j'ai analysé l'expression de l'insuline ou du glucagon dans les cellules infectées (GFP positives) et non infectées (GFP négatives). Nous avons ainsi montré que cette diminution est spécifiquement liée à une baisse de l'expression de l'insuline ou du glucagon dans les cellules infectées exprimant la GFP (Fig. 29b).

Ces résultats suggèrent donc que RFX6 serait nécessaire à la différenciation endocrine au cours de la différenciation pancréatique.

A



B

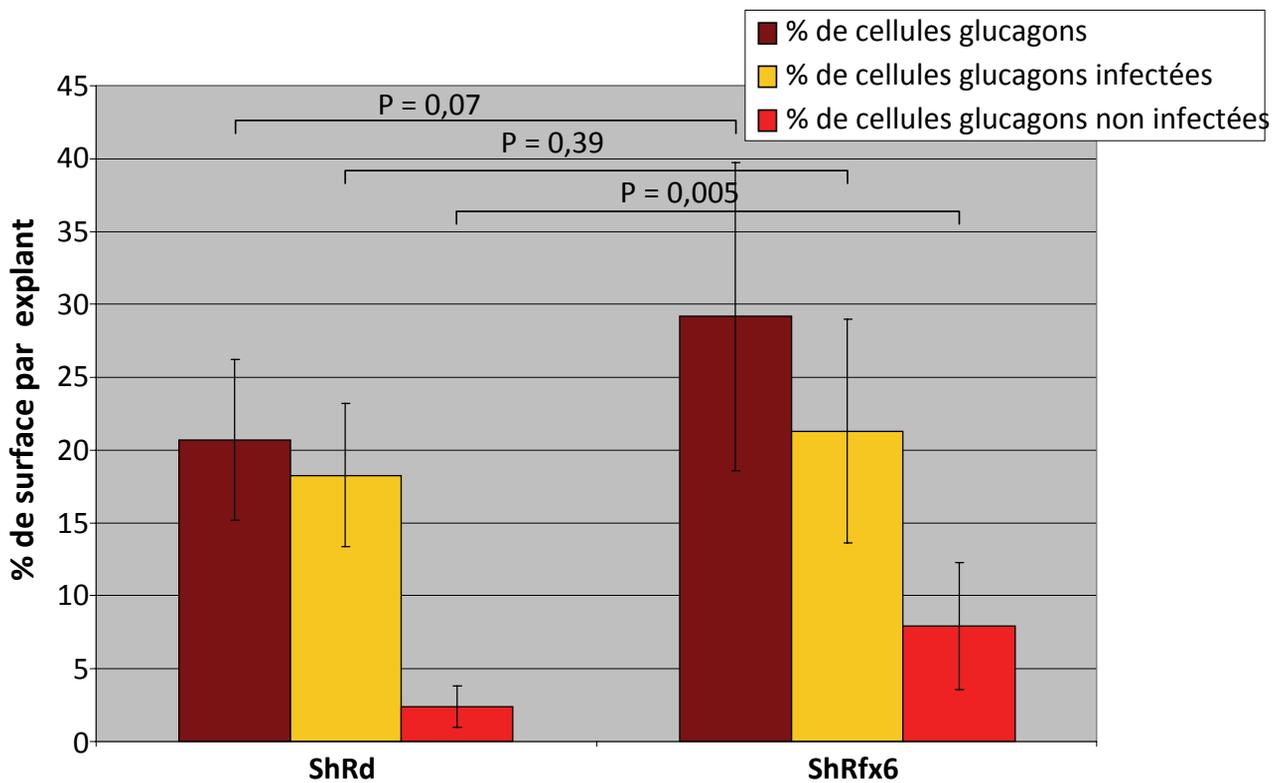


Figure 30 : Aucune diminution du nombre de cellules à insuline ou à glucagon n'est induite par le ShRfx6. Quantification de la surface de cellules exprimant l'insuline et la GFP (A), et le glucagon et la GFP (B) dans des explants disséqués à E11,5, 7 jours après l'infection avec le lentivirus exprimant le ShRd (insuline: n=3, glucagon: n=5) ou le ShRfx6 (insuline: n=7, glucagon: n=9). Le ShRfx6 n'induit aucune diminution de la différenciation des cellules glucagons ou insulines positives par rapport au ShRd dans les explants infectés.

Afin de déterminer quel type de cellules est affecté par la réduction de la différenciation endocrine, j'ai analysé séparément l'expression de l'insuline (Fig. 30A) et du glucagon (Fig. 30B). Comme précédemment, cette analyse a été réalisée dans des épithélia pancréatiques 7 jours après leur infection avec le lentivirus exprimant le ShRfx6 ou le ShRd. Puis, après l'inclusion et la coupe des explants, une immunofluorescence a ensuite été réalisée soit pour l'insuline et la GFP, soit pour le glucagon et la GFP. La surface des différents marquages a ensuite été quantifiée. Cette analyse a été réalisée respectivement sur deux et trois expériences différentes pour l'analyse de l'expression de l'insuline et du glucagon. A l'inverse des résultats précédemment décrits, nous n'avons pu mettre en évidence une diminution de l'expression des cellules à insuline ou à glucagon dans les épithélia infectés avec le lentivirus permettant l'expression du ShRfx6 par rapport aux épithélia infectés avec le lentivirus permettant l'expression du ShRd (Fig. 30).

Ces résultats soulèvent le problème de la validité des précédents résultats montrant une baisse de la différenciation endocrine dans les explants infectés avec le lentivirus permettant l'expression de ShRfx6 (Fig. 29B). L'hypothèse d'une différence de lot de lentivirus ne semble pouvoir être retenue. Car bien que trois lots de virus aient été utilisés au cours de ces différentes infections, le premier lot de virus qui a été utilisé pour deux des quatre expériences analysant l'impact sur la différenciation des cellules α (Fig. 30B), a également été utilisé pour les expériences analysant l'impact sur la différenciation des cellules α et β (Fig. 29B). Afin de mettre en évidence une potentielle mutation du ShRNA, j'ai également séquencé l'ADN que le lentivirus intègre au génome hôte après l'infection de la cellule. Par cette analyse, nous avons montré que la séquence du ShRNA et de son promoteur ne présentent pas de mutation. Une autre hypothèse est que le ShRfx6 ne soit pas capable d'induire une dégradation du transcrit endogène de *Rfx6*.

3.C Le ShRfx6 n'induit pas de dégradation de la forme endogène de l'ARNm de *Rfx6*

Afin de répondre à cette hypothèse, nous avons donc déterminé si le ShRfx6 pouvait induire une dégradation du transcrit *Rfx6* chez des cellules β TC3, lignée d'insulinome murin immortalisé, exprimant de façon endogène *Rfx6* (Fig. 31A). Pour cela, j'ai infecté des cellules β TC3 avec le lentivirus exprimant le ShRd ou le ShRfx6. Afin d'analyser l'expression de *Rfx6*

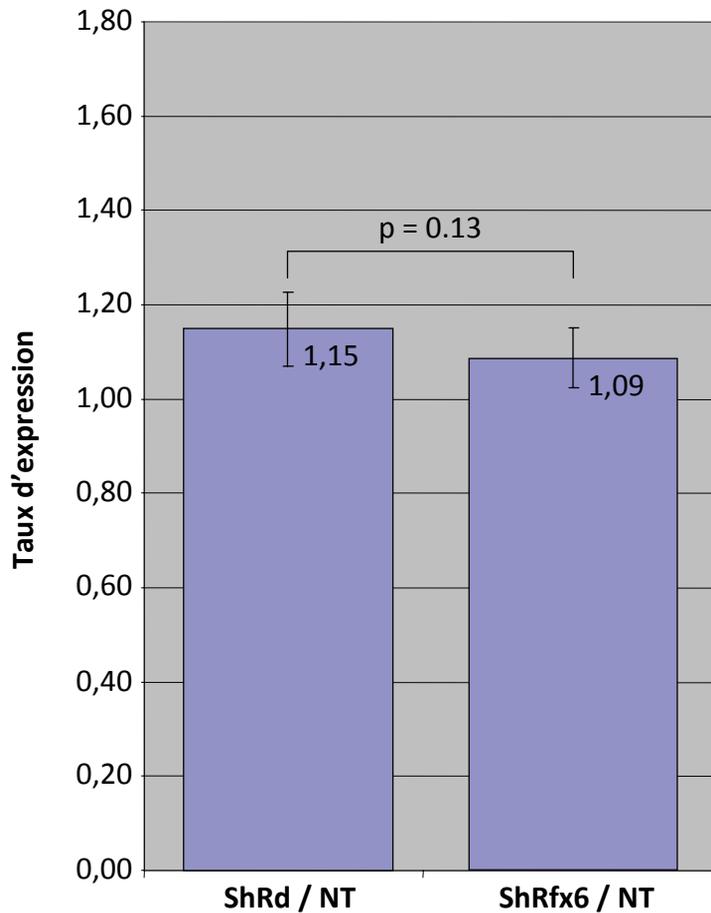
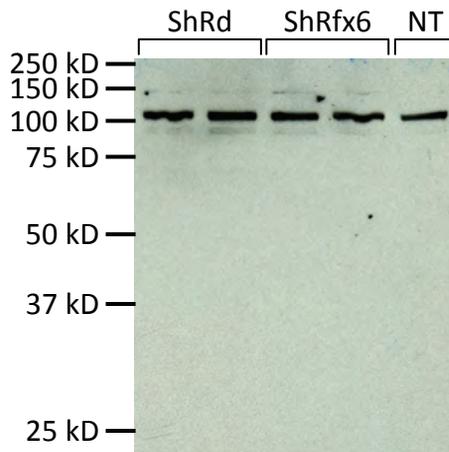
A**B**

Figure 31 : Le ShRfx6 n'induit pas de dégradation de la forme endogène de l'ARNm de *Rfx6*. (A) Quantification par RT-qPCR de l'ARNm de *Rfx6* dans des cellules β TC3 non traitées (NT), infectées avec le ShRd ou infectées avec le ShRfx6. (B) Western Blot pour RFX6 sur des extraits protéiques de cellules β TC3 non traitées (NT), infectées avec le ShRd (duplicata) ou infectées avec le ShRfx6 (duplicata). Le ShRfx6, comme le ShRd n'induit aucune diminution de la quantité d'ARNm ou de protéine RFX6 dans les cellules β TC3.

uniquement dans les cellules infectées (qui expriment la GFP), les cellules ont été triées par FACS selon leur expression de la GFP. Après extraction de l'ARN des cellules triées et des cellules non traitées, j'ai réalisé des RT-qPCR pour *Rfx6*. Cette analyse a été réalisée sur trois expériences différentes. Nous avons ainsi montré que le ShRfx6 n'induit pas de dégradation de la forme endogène du transcrite de *rfx6* par rapport aux cellules non traitées ou infectées avec le lentivirus exprimant le ShRd (Fig. 31A).

Dans le but de montrer que la non dégradation de la protéine RFX6 est liée à l'absence de dégradation de l'ARNm de *Rfx6* après l'infection avec le lentivirus exprimant le ShRfx6, nous avons confirmé ces résultats par western blot (WB). Pour cela, j'ai réalisé des extraits protéiques de cellules β TC3 non traitées ou GFP positives infectées avec le lentivirus exprimant le ShRd ou le ShRfx6, afin de faire des WB pour RFX6 (Fig. 31B). Nous avons ainsi confirmé que le ShRfx6 ne permet pas d'induire une dégradation de la protéine RFX6. Ces résultats montrent donc que le ShRfx6 ne permet pas d'induire une perte de fonction de *Rfx6*.

Ces résultats cependant contredisent la première caractérisation des ShRNA *Rfx6* #1 et ShRNA *Rfx6* #2. En effet, la première caractérisation des ShRNA *Rfx6* montre que le ShRNA *Rfx6* #2 induit une diminution de l'ARN bicistronique *Rfx6-IRES-GFP*. Toutefois, cette expérience n'était réalisée qu'en utilisant un fragment de l'ARNm de *Rfx6*. Ceci a pu modifier les structures secondaires de l'ARNm, rendant possible la dégradation de l'ARN bicistronique par le ShRNA *Rfx6* #2. De plus, il est à noter que le taux de transfection des cellules HEK293t au cours de ces tests était inférieur ou égal à 10 %. Ce faible taux de transfection a peut être mené à une interprétation erronée des résultats. Ces différentes hypothèses ne permettent cependant pas d'expliquer la diminution de la différenciation endocrine observée lors de la première analyse des épithélia (Fig. 29b). Toutefois, au vu des résultats obtenus lors des seconds tests sur l'impact du ShRfx6 sur la dégradation de l'ARNm de *Rfx6* (Fig. 31), nous pouvons supposer que ce résultat est un artefact d'expérimentation.

4. Caractérisation de la fonction de *Rfx6* chez le poisson zèbre

En parallèle à l'étude de *Rfx6* chez la souris, nous avons étudié en collaboration avec l'équipe du Dr Marianne Voz (Université de Liège, Liège), l'expression et la fonction de *zRfx6* chez le poisson zèbre.

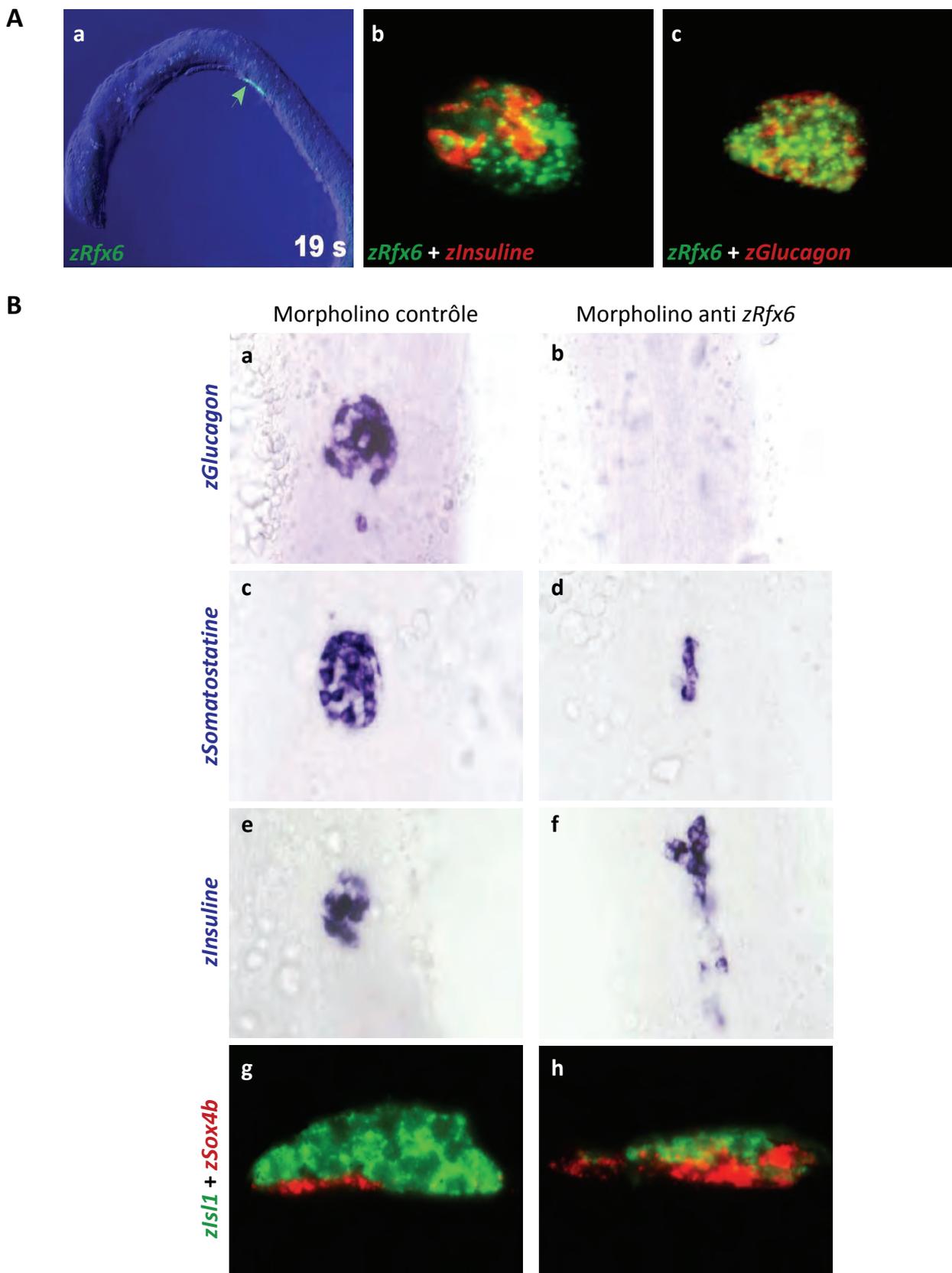


Figure 32 : La perte de fonction de *zRfx6* induit une réduction de la différenciation du pancréas endocrine et une accumulation des progéniteurs pancréatiques endocrines. (A) Hybridation *in situ* fluorescente pour *zRfx6* à 19 S (a). Double hybridation *in situ* fluorescente pour *zRfx6* (vert, b-c) et l'insuline (rouge, b) ou le glucagon (rouge, c). Rfx6 est exprimé uniquement dans le pancréas. (B) Hybridation *in situ* pour *zGlucagon* (a-b), *zSomatostatine* (c-d) et *zInsuline* (e-f) et double hybridation *in situ* fluorescente pour *zSox4b* (rouge, g-h) et *zIsl1* (vert) chez des embryons de poisson zèbre à 30 hpf, injectés avec un morpholino contrôle (a, c, e, g), et un morpholino contre *zRfx6* (b, d, f, h). 30 heures après la fécondation, le morpholino *zRfx6* induit une réduction de la différenciation endocrine pancréatique et une accumulation des progéniteurs endocriniens *zSox4b* positifs.

La caractérisation de l'expression de *zRfx6* chez le poisson zèbre a montré que *zRfx6* est exprimé uniquement dans le pancréas chez le poisson zèbre (Fig. 32Aa). L'expression chez le poisson zèbre commence dans le pancréas au stade 16 somites, et atteint son maximum à 24 hpf. Puis l'expression de *zRfx6* est maintenue au moins jusqu'à 72 hpf (Résultats non montrés).

L'étude de la perte de fonction de *zRfx6* a été réalisée par l'injection de morpholino contre *zRfx6*. Il a ainsi été montré que la perte de fonction de *zRfx6* induit une perte de la différenciation des cellules α et ε (Fig. 32Bb, et résultat non montré), une réduction de la différenciation des cellules δ (Fig. 32Bd) et une désorganisation des cellules β (Fig. 32Bf). De plus, cette réduction de la différenciation endocrine est liée à une accumulation des progéniteurs pancréatiques endocrines exprimant *zSox4b* (Mavropoulos et al., 2005 ; Fig. 32Bh).

Ces résultats suggèrent que *zRfx6* serait nécessaire à l'engagement des progéniteurs pancréatiques endocrines dans la différenciation endocrine. Cette conclusion est soutenue par d'autres résultats chez le poisson zèbre, présent dans l'article soumis et accepté dans *Development* (cf annexe).

DISCUSSION

Le développement du pancréas va s'initier par la mise en place de signaux externes à l'endoderme et par l'expression séquentielle de gènes dans l'endoderme, duquel le pancréas se développera. Le pancréas endocrine se différencie à partir des progéniteurs pancréatiques endocrines exprimant NEUROG3. Il a en effet été montré que l'expression du facteur de transcription NEUROG3 est nécessaire et suffisante pour induire la différenciation endocrine (Apelqvist et al., 1999 ; Gradwohl et al., 2000 ; Schwitzgebel et al., 2000 ; Grapin-Botton et al., 2001 ; Johansson et al., 2007). En dépit du rôle primordial de ce gène, un seul gène cible régulé directement par NEUROG3 était caractérisé au début de mon doctorat, *Neurod1* (Huang et al., 2000). L'objectif central de mon doctorat a donc été de caractériser le transcriptome pancréatique induit par NEUROG3 au cours de la différenciation pancréatique endocrine et d'étudier l'expression et la fonction des potentiels nouveaux gènes cibles de NEUROG3 mis en évidence.

A) Caractérisation du transcriptome pancréatique induit par NEUROG3 au cours de la différenciation du pancréas endocrine

Avant mon doctorat, différentes études transcriptomiques avaient été réalisées dans l'objectif de caractériser le transcriptome pancréatique contrôlé par NEUROG3. La première étude correspondant aux travaux de Gu avait utilisé une souris transgénique exprimant la GFP sous le contrôle du promoteur *Neurog3* (Gu et al., 2004). Récemment, une étude similaire a été réalisée en utilisant une souris hétérozygote « Knock In » de la GFP dans la séquence codant *Neurog3* (White et al., 2008). Ces deux études ont permis de caractériser les gènes exprimés dans les cellules exprimant NEUROG3, en analysant spécifiquement les cellules exprimant la GFP, mais ne permettaient pas de définir les gènes spécifiquement induits par NEUROG3. En outre, les différences de résultat entre ces deux analyses suggèrent que le promoteur utilisé par Gu pourrait ne pas contenir l'ensemble des éléments essentiels pour le contrôle de *Neurog3*. De plus, ces deux études ont utilisé des puces à ADNc ne couvrant que partiellement le génome et n'analysant que 12000 transcrits. Ainsi en comparant nos résultats aux leurs, on ne trouve pas de gènes tels que *RikenL19*, *Rfx6*, *Fev* ou *Riken cDNA 3100002J23*, que nous avons caractérisé comme spécifiquement induits par NEUROG3 lors de nos analyses transcriptomiques. Deux autres analyses ont caractérisé le

transcriptome des souris *Neurog3*^{-/-}, permettant de mettre en évidence les gènes nécessitant l'expression de NEUROG3 au cours du développement pancréatique (Petri et al., 2006 ; Juhl et al., 2008). Toutefois, comme pour les analyses de Gu et al. et de White et al., ces deux études ne permettaient pas de caractériser les gènes spécifiquement induits par NEUROG3. Par ailleurs, on peut noter que comme pour les études de Gu et al., et de White et al., la majorité de ces analyses n'a été réalisée que sur une partie du génome via l'utilisation de puce à ADNc ne couvrant que 12000 transcrits, et n'a pas permis de mettre en évidence des gènes tels que *Rfx6* et *Fev*. En outre, les stades analysés par ces différents groupes, à l'exception des travaux de Juhl et al., sont différents de E12,5, stade auquel les progéniteurs multipotents endocriniens sont compétents pour se différencier en cellules sécrétant l'insuline (Johansson et al., 2007). Ces études n'ont donc pas permis de mettre en évidence les gènes induits par NEUROG3 au cours de la différenciation des cellules β , cellules dont la caractérisation des programmes de différenciation et de maturation pourrait être nécessaire à la thérapie cellulaire du diabète de type 1.

Afin de déterminer les gènes spécifiquement induits par NEUROG3 au cours de la différenciation du pancréas endocrine et notamment des cellules β , le transcriptome pancréatique induit à E12,5 par la forme active de NEUROG3ERTM chez les souris *Neurog3*^{-/-} ; *Pdx1-Neurog3ER*TM a été caractérisé. Afin d'étudier les gènes dont l'expression est spécifiquement contrôlée par NEUROG3, nous avons comparé cette analyse transcriptomique avec le transcriptome des souris *Neurog3*^{-/-} caractérisé au laboratoire (non publié). Finalement, afin d'analyser des cibles de NEUROG3 induites dès le début de la différenciation endocrine, nous avons comparé ces analyses avec le transcriptome des progéniteurs endocrines (non publié) obtenu par l'utilisation de la souris « Knock Add On » *Neurog3-eYFP* générée dans notre laboratoire (Mellitzer et al., 2004). En outre, l'ensemble de nos analyses a été réalisé sur des puces à ADNc Affymetrix couvrant plus de 36000 transcrits et variants, à la différence des précédentes études qui n'analysaient que 12000 transcrits.

Ainsi, par la combinaison unique de modèles de souris et l'utilisation de la puce à ADNc Affymetrix couvrant plus de 36000 transcrits, nous avons établi une liste de 56 gènes dont certains étaient précédemment non caractérisés, tel que *Rfx6*. Nos analyses transcriptomiques sont en conséquent complémentaires des précédentes analyses transcriptomiques réalisées sur le pancréas en développement.

B) Le gène *RIKEN cDNA 1700086L19*

Parmi cette liste, nous avons étudié le gène *RIKEN cDNA 1700086L19* (*Riken L19*), précédemment non caractérisé. L'expression de *Riken L19* est en effet induite 46 fois par NEUROG3 chez les souris *Neurog3^{-/-}*; *Pdx1-Neurog3ERTM* injectées avec le 4-hydroxytamoxifène, perdue chez le mutant *Neurog3* et enrichie 122 fois dans les progéniteurs pancréatiques endocrines. Au cours du développement pancréatique, mes expériences montrent que son expression induite à E10,5 atteint son maximum à E15,5 pour être perdue à E18,5. A E15,5, l'expression de *Riken L19* est perdue chez les souris *Neurog3* mutantes. Ces résultats montrent que l'expression de *Riken L19* est spécifique du lignage endocrine. Toutefois, alors que NEUROG3 pourrait réguler l'expression de *Riken L19*, aucune cellule exprimant *Riken L19* n'a été caractérisée comme coexprimant NEUROG3. Ces résultats suggèrent que *Riken L19* ne serait pas une cible directe de NEUROG3. Mais son expression étant perdue chez les mutants *Neurog3*, l'expression de *Riken L19* pourrait être régulée par des gènes dont l'expression est régulée par NEUROG3. Par ailleurs, son expression est indépendante de celle d'*Arx* et *Pax4*, deux gènes impliqués respectivement dans la différenciation des cellules α et β (Collombat et al., 2003 ; Sosa-Pineda et al., 1997). Ces résultats suggèrent donc également que l'expression de *Riken L19* au cours du développement du pancréas endocrine n'est pas liée à la différenciation d'un lignage pancréatique endocrine particulier. Par ailleurs, bien que l'expression de *Riken L19* soit similaire à celle de *Neurod1* et de *Insm1*, deux gènes cibles directs de NEUROG3 impliqués dans la différenciation de l'ensemble des sous types cellulaires endocriniens (Naya et al., 1997 ; Mellitzer et al., 2006), ces résultats ne permettent pas d'affirmer que *Riken L19* est impliqué dans un processus de différenciation général de l'épithélium pancréatique endocrine au cours du développement pancréatique.

En outre, le patron d'expression de *Riken L19* dans les îlots de Langerhans est unique. *Riken L19* est exprimé dans les cellules sécrétant le glucagon, la somatostatine ou le polypeptide pancréatique dans les îlots de Langerhans. En effet, de nombreux gènes sont connus pour être exprimés à la fois dans les cellules α , β et PP tels que *Isl1* ou *Nkx2.2* (Ahlgren et al., 1997 ; Sussel et al., 1998), ou simplement dans les cellules β tels que *Pdx1* (Wu et al., 1997). Mais aucun gène n'est connu pour avoir ce patron d'expression chez

l'adulte. *Riken L19* pourrait ainsi avoir un rôle opposé à celui de PDX1 dans les îlots de Langerhans, où PDX1 est spécifiquement exprimé dans les cellules β (Ohlsson et al., 1993)

De plus, la séquence putative d'un domaine SERTA dans la séquence de *Riken L19* pourrait lui permettre dans le pancréas d'être impliqué dans la régulation du cycle cellulaire. En effet, les gènes contenant le domaine SERTA sont caractérisés pour leur fonction dans la régulation du cycle cellulaire. Les protéines SERTAD1 et SERTAD2 sont caractérisées par leur capacité à activer le complexe CDK4/Cyclin D nécessaire à la progression du cycle cellulaire (Sim et al., 2006), ou pour SERTAD1 à inhiber la croissance cellulaire de manière dépendante ou indépendante de P53 (Watanabe-Fukunaga et al., 2005). La présence d'un domaine SERTA dans la protéine RIKEN L19 pourrait ainsi permettre à RIKEN L19 de réguler le cycle cellulaire des cellules endocrines en différenciation, ainsi que des cellules sécrétant le glucagon, la somatostatine ou le polypeptide pancréatique dans les îlots de Langerhans.

C) Le gène *Rfx6*

En parallèle à la caractérisation du gène *Riken L19*, l'analyse transcriptomique a également révélé *Rfx6* (précédemment dénommé *Rfxdc1*), un autre gène non caractérisé avant cette étude, et membre de la famille des facteurs de transcription RFX à domaine « winged helix » (Gajiwala et al., 2000a; Gajiwala et al., 2000b). En effet, l'expression de *Rfx6* est induite par NEUROG3 chez les souris *Neurog3*^{-/-}; *Pdx1-Neurog3ER*TM après injection du 4-hydroxytamoxifène, perdue chez le mutant *Neurog3* et enrichie plus de 20 fois dans les progéniteurs pancréatiques endocrines. D'après notre étude transcriptomique, les autres gènes de la famille RFX, ne semblent pas être exprimés dans le pancréas (*Rfx2*, *Rfx4*) ou enrichis dans les progéniteurs pancréatiques endocrines (*Rfx1*, *RFX3*, *Rfx5*, *Rfx7*). Récemment, les travaux de Ait-Lounis et al. ont mis en évidence que *Rfx3* est nécessaire à la différenciation normale du pancréas endocrine (Ait-Lounis et al., 2007).

L'étude de l'expression de RFX6 montre que son expression varie au cours du développement embryonnaire, suggérant que RFX6 pourrait avoir plusieurs fonctions au cours du développement.

1. La fonction précoce de *Rfx6* dans l'endoderme intestinal primitif

Au cours du développement murin, nous avons montré que l'expression ubiquitaire de RFX6 va progressivement disparaître dans l'endoderme intestinal primitif au cours de sa différenciation à partir de E8,5. La perte de l'expression ubiquitaire de RFX6 dans l'épithélium des différents organes issus de l'endoderme intestinal primitif au cours de leur différenciation suggère que RFX6 serait nécessaire au maintien d'un état indifférencié de l'endoderme intestinal primitif, et que la perte de son expression serait nécessaire à sa différenciation. Plus précisément, dans les bourgeons pancréatiques, l'expression ubiquitaire de RFX6 à E8,5, se restreint progressivement pour être absente dans les progéniteurs pancréatiques exprimant PTF1a, qui formeront l'ensemble des sous types cellulaires pancréatiques (Kawaguchi et al., 2002). Ainsi, dans les bourgeons pancréatiques en différenciation, la perte d'expression de RFX6 pourrait être requise pour l'expression de *Ptf1a*, dont l'expression est nécessaire au développement pancréatique. La perte de l'expression ubiquitaire dans l'endoderme intestinal primitif qui formera l'intestin, remplacée par une expression éparse dans l'épithélium intestinal, correspond à la vague de différenciation antéropostérieure de l'épithélium intestinal (Simon-Assmann et al., 1993). Ces résultats suggèrent que comme dans le pancréas, la perte de l'expression ubiquitaire de RFX6 dans l'épithélium intestinal primitif formant l'épithélium intestinal pourrait être nécessaire à sa différenciation.

2. *Rfx6* dans la cascade génétique de différenciation du pancréas endocrine

Dans le pancréas en développement, la moitié des cellules exprimant NEUROG3 expriment RFX6. De plus, dans le pancréas des souris mutantes pour *Neurog3*, l'expression de *Rfx6* est perdue. Ceci démontre que l'expression de *Rfx6* est spécifique du lignage pancréatique endocrine bien que ces résultats ne permettent pas de conclure sur le fait que *Rfx6* serait une cible directe de NEUROG3. Cependant, l'observation d'un patron d'expression de *Rfx6* similaire à celui de *Insm1* et de *Neurod1* deux gènes cibles directs de NEUROG3 (Huang et al., 2000 ; Mellitzer et al., 2006) suggère que *Rfx6* pourrait être une cible directe de NEUROG3. Cette hypothèse est également étayée par la présence de site

consensus « E box », site de fixation des protéines à « bHLH » tel que NEUROG3, dans la séquence du promoteur de *Rfx6*.

Par ailleurs, l'expression de *Rfx6* est maintenue dans les souris mutantes pour *Arx* ou *Pax4*, deux gènes impliqués respectivement dans la différenciation des cellules α et β (Collombat et al., 2003 ; Sosa-Pineda et al., 1997) et pour *NeuroD1*, qui est nécessaire à la différenciation de l'ensemble des lignages insulaires (Naya et al., 1997). Ces résultats suggèrent que l'expression de *Rfx6* est indépendante de celle de *Arx*, *Pax4* et de *NeuroD1*, et que comme *Insm1*, *Rfx6* intervient soit en amont d'*Arx*, *Pax4* et *NeuroD1*, soit en parallèle à ces trois gènes chez la souris.

3. La famille des gènes *Rfx* dans la différenciation endocrine

Récemment, un autre gène de la famille *Rfx*, *Rfx3*, a été impliqué dans la différenciation du pancréas endocrine. En effet, Ait Lounis et al. ont montré que la perte de fonction de *Rfx3* induit une réduction de la différenciation des cellules α , β et ϵ , et une augmentation de la différenciation des cellules PP (Ait-Lounis et al., 2007). Suite à ce défaut du développement pancréatique, les souris *Rfx3* mutantes développent une intolérance au glucose. A E15,5, nous avons montré que RFX3 et RFX6 sont coexprimés dans les cellules pancréatiques endocrines en développement. Or, les protéines de la famille RFX sont connues pour former des homo- ou hétéro-dimères (Morotomi-Yano et al., 2002). Ces deux facteurs de transcription de la famille RFX pourraient donc se dimériser dans les progéniteurs pancréatiques endocrines en différenciation, formant ainsi un complexe qui pourrait être requis pour la différenciation des cellules pancréatiques endocrines. Par ailleurs, notre analyse transcriptomique montre que, bien que l'expression de *Rfx3* soit faiblement enrichie dans les progéniteurs pancréatiques endocrines (3,03 fois contre 21,14 fois pour *Rfx6*), son expression n'est pas perdue chez le mutant *Neurog3*. Ces résultats suggèrent donc que seule l'expression de RFX6 est spécifique du lignage pancréatique endocrine.

Par ailleurs, les travaux de Aughsteen ont montré que les îlots de Langerhans présentent à leur surface des cils primaires (Aughsteen, 2001). Toutefois, la perte de fonction de *Kif3a*, un gène impliqué dans la formation des cils, dans les progéniteurs pancréatiques n'induit aucun problème pendant le développement du pancréas endocrine,

aucune malformation des îlots de Langerhans, ni aucune anomalie dans la réponse au glucose chez la souris *Kif3a*^{-/-} adulte (Cano et al., 2006). Les travaux de Ait Lounis et al. ont eux montré que la perte de fonction de *Rfx3* induit une malformation des cils sur les cellules pancréatiques endocrines chez les nouveaux nés, ainsi qu'une malformation des îlots de Langerhans. Ces résultats suggèrent donc que la malformation des cils pouvait être à l'origine de la différenciation incorrecte des cellules pancréatiques endocrines (Ait-Lounis et al., 2007). Nos analyses suggèrent elles que les cils ne sont pas présents sur les cellules pancréatiques à E15,5. Ces résultats suggèrent donc que le phénotype embryonnaire observé chez les mutants *Rfx3* ne serait pas lié à une malformation des cils au cours du développement pancréatique. Toutefois, des cils seraient présents dans le pancréas à E15,5 (Durand et al., non publié). L'ensemble de ces résultats suggère donc que le phénotype pancréatique observé chez les mutants *Rfx3* ne serait pas lié à la malformation des cils puisque l'absence des cils d'après Cano et al. n'induit pas de phénotype pancréatique endocrine. La malformation des îlots de Langerhans chez le mutant *Rfx3* pourrait alors s'expliquer par la perte d'un hétérodimère putatif RFX3-RFX6, qui pourrait être nécessaire à la différenciation du pancréas endocrine.

4. La fonction de *Rfx6* dans le pancréas adulte

Chez la souris adulte, RFX6 est exprimé dans les cellules sécrétant l'insuline, le glucagon, la somatostatine et le polypeptide pancréatique. Mais alors qu'au stade embryonnaire l'expression de *Rfx6* est induite par NEUROG3, comment l'expression de RFX6 est elle maintenue ? En effet, bien que l'expression de NEUROG3 soit observée dans les îlots de Langerhans, son expression est très faible par rapport à son expression au cours du développement embryonnaire (Wang et al., 2009b). A ce stade, RFX6 pourrait autoréguler son expression, ou un autre facteur de transcription tel que PDX1 pourrait contrôler son expression. La totalité des cellules RFX6 semble également ciliée. Ainsi dans les îlots de Langerhans, RFX6 pourrait être impliqué dans la régulation de la formation des cils, comme RFX3 (Ait-Lounis et al., 2007). Toutefois, depuis les travaux de Cano et al. montrant que les cils ne sont pas requis dans l'homéostasie du glucose sanguin (Cano et al., 2006), la fonction des cils sur les cellules pancréatiques endocrines reste à déterminer.

5. La fonction de *Rfx6* dans la différenciation de l'intestin

En plus de son expression dans l'endoderme intestinal primitif et dans les cellules pancréatiques endocrines, RFX6 est également exprimé dans l'épithélium gastrique et intestinal en différenciation. A partir de E13,5, l'expression de RFX6 commence à se restreindre dans l'épithélium intestinal dans la région adjacente de l'estomac. Puis, à 15,5, l'expression de RFX6 est restreinte à quelques cellules épithéliales intestinales et, est similaire à celle de *NEUROG3* (Jenny et al., 2002). Comme dans le pancréas à E15,5, l'expression de *Rfx6* est perdue chez le mutant *Neurog3* dans l'intestin à E15,5. Jenny et al. ont montré que chez le mutant *Neurog3* la différenciation endocrine intestinale est perdue. Ainsi, la perte d'expression de *Rfx6* dans l'intestin chez le mutant *Neurog3* suggère que *Rfx6* pourrait être impliqué dans la différenciation des cellules endocrines intestinales.

6. La fonction de *Rfx6* dans la différenciation pancréatique du poisson zèbre

En collaboration avec l'équipe du Dr Marianne Voz, nous avons utilisé le modèle du poisson du zèbre pour déterminer la fonction de *Rfx6*. L'ensemble de l'analyse de la perte de fonction de *Rfx6* est présenté dans l'article accepté dans *Development* (cf annexe). L'orthologue *zRfx6*, mais pas *zRfx3* (non montré, cf annexe), est exprimé dans le lignage pancréatique endocrine. L'injection de deux morpholinos distincts montre que *zRFX6* est nécessaire à la différenciation correcte du pancréas endocrine chez le poisson zèbre. En effet, la plupart des cellules sécrétant le glucagon ou la ghréline sont perdues, et la différenciation des cellules sécrétant la somatostatine est fortement réduite. En revanche, la différenciation des cellules sécrétant l'insuline n'est que partiellement réduite par la perte de fonction de *zRfx6*. Cet impact réduit sur la différenciation des cellules β pourrait être dû au fait que la différenciation des cellules β commence dès 15 hpf (Biemar et al., 2001) alors que l'expression de *Rfx6* chez le poisson zèbre n'est pas détectée avant 17 hpf (16 somites). De même, la différenciation des cellules à somatostatine commençant à 17 hpf (Biemar et al., 2001) peut expliquer l'impact modéré de la perte de fonction de *zRFX6* sur la différenciation des cellules δ . Par ailleurs, la perte de fonction de *zRFX6* induit une accumulation des progéniteurs pancréatiques endocrines exprimant *zSox4b* (Mavropoulos et

al., 2005). Ces résultats suggèrent que zRFX6 serait nécessaire à la différenciation des progéniteurs endocrines en cellules α et ε chez le poisson zèbre.

7. Implication de RFX6 dans la thérapie cellulaire pancréatique

Les patients atteints de diabète de type 1 peuvent avoir recours à des greffes de pancréas ou d'îlots de Langerhans au cours de l'évolution de la maladie. Toutefois, le nombre de greffons pancréatiques ne permet pas de subvenir aux besoins. Afin de palier à ce manque, la thérapie cellulaire par différenciation de cellule souche en cellule insuline est aujourd'hui étudiée. D'Amour et al ont ainsi obtenu par différenciation de cellules souches des cellules sécrétant l'insuline (Fig. 33 ; D'Amour et al., 2006). Toutefois, dans ces cellules, la proinsuline n'est pas correctement métabolisée et la sécrétion du C-peptide après stimulation par le glucose est très faible par rapport à des cellules β humaines. Ces résultats suggèrent donc que les cellules β formées par les expériences de D'Amour et al. sont immatures. Récemment, ces cellules ayant atteint le stade 4 du protocole de différenciation (Fig. 33) ont été greffées chez des souris dans lesquelles les cellules β avaient été détruites par un traitement à la streptozotocine. Il a ainsi été mis en évidence que les cellules greffées en cours de différenciation se différencient en cellules β matures, pouvant réguler leur sécrétion d'insuline afin de maintenir le taux de glucose à un niveau physiologique (Kroon et al., 2008). Ces travaux suggèrent donc que la différenciation terminale correcte de ces cellules requiert leur implantation dans des conditions physiologiques.

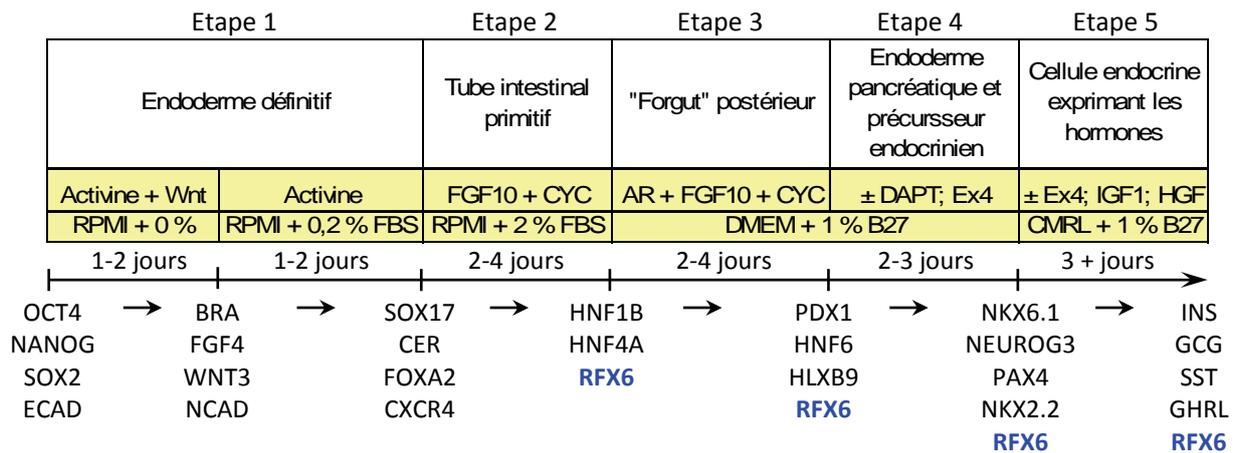


Figure 33 : Schéma des procédures de différenciation des cellules souches en cellules pancréatiques endocrines et expression protéique au cours de la différenciation pancréatique. Le protocole de différenciation des cellules souches en cellules pancréatiques endocrines est divisé en cinq étapes. Dans un premier temps, sous le contrôle des facteurs régulant les voies du TGF β et Wnt, les cellules souches vont se différencier en endoderme, exprimant SOX17, CERBERUS ou FOXA4. Ces cellules vont ensuite se différencier en cellules caractéristiques du tube intestinal primitif et du « Forgut » postérieur, en présence de FGF10 et cyclopamine (un inhibiteur de la signalisation Hedgehog), puis de FGF10, cyclopamine et d'acide rétinoïque. L'inactivation de Notch par le DAPT et l'ajout d'exendin-4 vont conduire ensuite ces cellules à se différencier en cellules exprimant NEUR OG3, PAX4, NKX2.2, facteurs de transcription caractéristiques des progéniteurs pancréatiques endocriniens ou des cellules endocrines en différenciation. Finalement, suite à l'ajout IGF1 et de HGF, ces cellules se différencieront en cellules sécrétant l'insuline, le glucagon, la somatostatine ou la ghréline. De par son expression au cours du développement embryonnaire chez la souris, RFX6 pourrait être utilisé comme un nouveau marqueur afin de vérifier la différenciation correcte des cellules souches en cellules sécrétant l'insuline. RFX6 pourrait ainsi être utilisé comme marqueur précoce dès le stade « forgut » postérieur afin de vérifier l'engagement correct de la différenciation des cellules souches. Le maintien de l'expression de RFX6 au cours de la différenciation des cellules pourrait également être utilisé comme marqueur à des stades plus tardifs tel que celui des « cellules endocrines exprimant les hormones ». AR : Acide Rétinoïque, CYC : Cyclopamine, DAPT : inhibiteur de la γ -sécrétase, Ex4 : Exendin-4, HGF : Hepatic Growth Factor, IGF1 : nsulin Growth Factor 1 (Adaptée de D'Amour et al., 2006)

On peut également noter qu'alors que de nombreux marqueurs sont connus pour les phases précoces de différenciation des cellules souches tels que SOX17 ou Cerberus pour la différenciation en endoderme, ou PDX1 et HLXB9 pour l'engagement dans la différenciation pancréatique (Fig. 33), peu de marqueurs sont connus pour les étapes de maturation des cellules endocrines à l'exception des hormones elles-mêmes. Au vu de l'expression précoce de RFX6 dans l'endoderme qui formera l'épithélium pancréatique, RFX6 pourrait être utilisé comme marqueur pour vérifier que les cellules souches s'engagent dans une différenciation correcte en endoderme pancréatique. RFX6 est également réexprimé dans les cellules pancréatiques endocrines en différenciation, puis est maintenue dans les îlots de Langerhans. De par cette expression et puisque que les facteurs de transcription tels que NKX2.2 ne permettent pas de caractériser précisément l'état de différenciation des cellules endocrines, RFX6 pourrait être utilisé comme marqueur d'une différenciation normale des

cellules souches en cellules endocrines, ainsi que comme marqueur des cellules endocrines matures, afin d'améliorer la caractérisation de ces cellules. Par ailleurs, le protocole de différenciation *in vitro* des cellules souches en cellules sécrétant l'insuline ne permet pas de former des cellules β matures (D'Amour et al., 2006). Ces résultats pourraient suggérer que d'autres facteurs que l'IGF1 et le HGF pourraient être nécessaires à la différenciation terminale des cellules sécrétant l'insuline. La caractérisation des gènes et des voies de signalisation induits par RFX6 lors de la différenciation et la maturation des cellules β pourrait permettre de définir de nouvelles molécules nécessaires à la différenciation des cellules β matures. En outre, le taux de cellules à insuline formé par le protocole de D'Amour est inférieur à 10% (D'Amour et al., 2006). La caractérisation des gènes induits par RFX6 dans les cellules β des îlots de Langerhans pourrait ainsi permettre de définir de nouveaux marqueurs de surfaces afin de purifier les cellules β matures dans l'objectif d'une greffe de ces cellules.

D) Conclusion générale et perspectives

Mon doctorat a eu pour objectif d'identifier de nouveaux gènes cibles de NEUROG3 impliqués dans la différenciation du pancréas endocrine. Pour cela, au cours de mon doctorat, j'ai caractérisé le transcriptome induit par NEUROG3 au cours de la différenciation pancréatique, à E12,5, stade auquel les progéniteurs pancréatiques endocrines ont la compétence de se différencier en cellules α et β (Johansson et al., 2007). Cette étude a permis d'établir une liste de 56 gènes potentiellement impliqués dans la différenciation du pancréas endocrine sous le contrôle de NEUROG3. J'ai ainsi poursuivi mon doctorat par l'étude de deux gènes de cette liste dont la séquence et la fonction étaient inconnues au début de cette étude : le gène *RIKEN cDNA 1700086L19* et le gène *Rfx6*.

L'étude du gène *RIKEN cDNA 1700086L19* a montré que le gène *Riken L19* est spécifiquement exprimé dans le lignage endocrine au cours du développement pancréatique endocrine. Son expression est également conservée dans les îlots de Langerhans, et semble restreinte aux cellules α , δ et PP. Je n'ai toutefois pas pu caractériser de séquence codante pour la protéine RIKEN L19 au cours de mon doctorat. En effet, alors que le dernier exon

putatif de la séquence nucléique de *Riken L19* contenait la séquence codant pour un domaine conservé SERTA, aucune séquence protéique contenant ce domaine SERTA ou un autre domaine conservé n'a pu être déterminée. Toutefois, une séquence a été caractérisée chez le macaque (XP_001096508), le gorille (XP_001153624) et le cheval (XP_001493324), et permet théoriquement la formation d'une protéine contenant un domaine SERTA. En comparant les séquences nucléiques du domaine SERTA chez la souris, on observe une homologie de 96 % entre la souris, le gorille et le macaque et de 94 % entre la souris et le cheval (Fig. 34). Ainsi, en comparant les séquences nucléiques de *Riken L19* du gorille, du macaque et du cheval avec la séquence génomique murine, une nouvelle séquence murine de *Riken L19* pourrait être caractérisée permettant la synthèse d'une protéine contenant le domaine SERTA. La caractérisation de cette séquence pourrait permettre de définir sa fonction par l'élaboration d'une stratégie de perte de fonction telle que le développement d'une souris mutante pour *Riken L19* ou de gain de fonction. Le développement d'un anticorps sera également nécessaire pour déterminer précisément dans quels types cellulaires RIKEN L19 est exprimé au cours du développement du pancréas endocrine et dans les îlots de Langerhans.

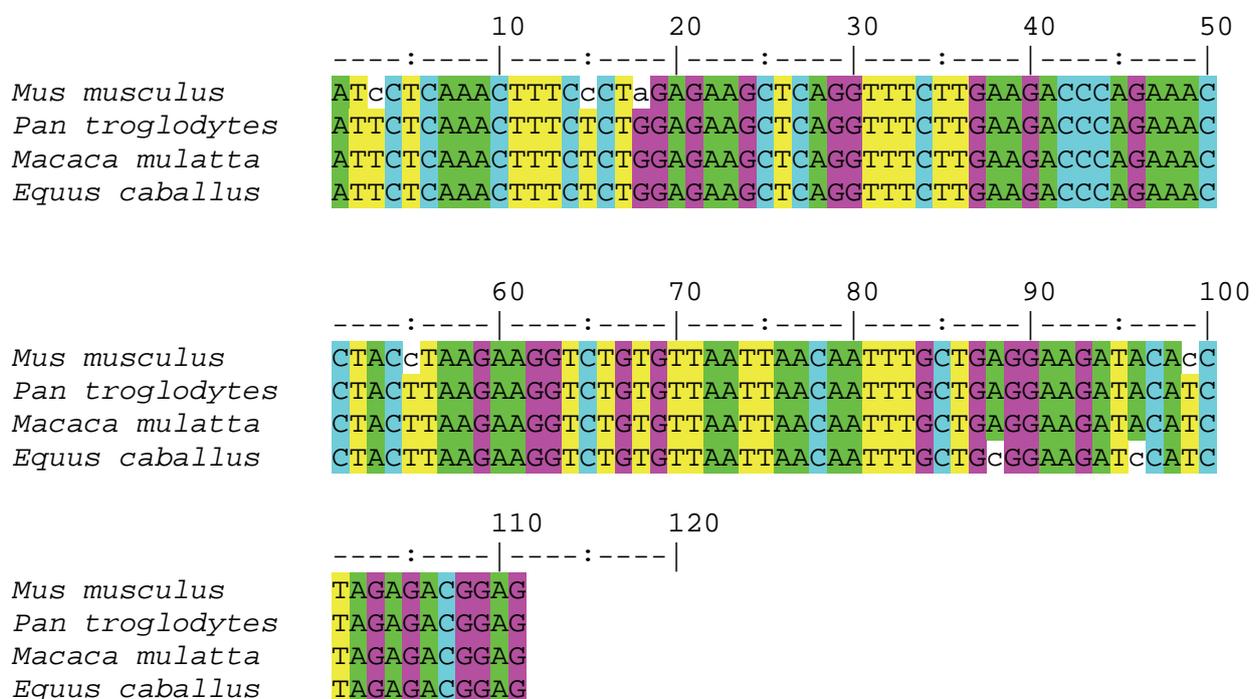


Figure 34 : Alignement des séquences nucléiques correspondant au domaine SERTA de *Riken L19* chez la souris, le gorille, le macaque et le cheval. La séquence murine du domaine SERTA présente une homologie de 96 % avec le gorille et le macaque, et de 94 % avec le cheval.

L'étude du gène *Rfx6*, dont le patron d'expression et la fonction n'étaient pas connus avant cette étude a montré que son expression évolue au cours du développement embryonnaire. Dans un premier temps, RFX6 est exprimé dans l'endoderme intestinal primitif. Puis, son expression va se restreindre aux lignages pancréatiques endocrines et intestinaux en différenciation. Chez l'adulte, son expression est maintenue dans les îlots de Langerhans. Bien que l'étude de sa fonction n'ait pu être adressée chez la souris par une approche de culture organotypique, la perte de fonction de *Rfx6* chez le poisson zèbre a montré que *Rfx6* est nécessaire à la différenciation du lignage pancréatique endocrine. Il sera donc intéressant de définir si la fonction caractérisée chez le poisson zèbre est conservée chez les mammifères.

Pour cela, la réalisation d'une souris mutante conditionnelle pour *Rfx6* est en cours, dans laquelle deux séquences *LoxP* seront placées de part et d'autre de l'exon trois, principal contributeur du domaine de liaison à l'ADN RFX. Le croisement d'une telle souris avec différentes souris transgéniques exprimant la *Cre* sous le contrôle de différents promoteurs nous permettra d'étudier les différentes fonctions de *Rfx6* au cours du développement pancréatique et intestinal. Le croisement de la souris *Rfx6^{LoxP/LoxP}* avec une souris *Pdx1-Cre* nous permettra d'étudier la fonction précoce de RFX6. En effet, la souris *Pdx1-Cre* permet une expression de la *Cre* dans les bourgeons pancréatiques (Gannon et al., 2000). En croisant la souris *Rfx6^{LoxP/LoxP}* avec une souris *HesX1-Cre* permettant une expression de la *Cre* dans l'endoderme intestinal primitif antérieur, nous pourrions observer la fonction de RFX6 dans l'endoderme intestinal primitif qui formera le pancréas (Andoniadou et al., 2007). Cette souris *Rfx6^{LoxP/LoxP}* ; *HesX1-Cre* devrait nous permettre d'étudier si l'expression de RFX6 est nécessaire à la perte de l'expression de *shh* dans les bourgeons pancréatiques. En effet, il a été montré que la perte d'expression de *Shh* est nécessaire à l'engagement de la différenciation pancréatique du bourgeon dorsal (Hebrok et al., 1998). La famille des gènes *Hnf* a également été impliquée dans la différenciation pancréatique. Dans cette famille, les gènes *Hnf1β* (*Tcf2* ; Haumaitre et al., 2005), *Hnf3β* (*Foxa2* ; Weinstein et al., 1994 ; Ang et al., 1994) et *Hnf6* (*Oc-1* ; Rausa et al., 1997) sont exprimés tôt au cours de la différenciation pancréatique. *Tcf2* et *Foxa2* ont notamment été caractérisés comme étant exprimés dans l'endoderme intestinal primitif. La perte de fonction pour *Tcf2* a en outre montré que l'expression de *Tcf2* est nécessaire au développement pancréatique précoce (Haumaitre et

al., 2005). Il serait donc intéressant de déterminer si *Rfx6* est régulé ou régule l'expression de ces gènes au cours du développement précoce du pancréas.

L'utilisation du mutant conditionnel pour *Rfx6* nous permettra également d'étudier la fonction de RFX6 au cours de la différenciation des cellules endocrines. Ceci pourra être réalisé en croisant la souris *Rfx6*^{LoxP/LoxP} avec une souris *Neurog3-Cre*, exprimant la Cre dans les progéniteurs pancréatiques endocrines multipotents (Gu et al., 2002). Cette souris *Rfx6*^{LoxP/LoxP} ; *Neurog3-Cre* pourrait en outre nous permettre de déterminer la place de *Rfx6* par rapport à *Insm1* dans la cascade génétique de différenciation des cellules pancréatiques endocrines en étudiant l'expression de *Insm1* dans les progéniteurs endocrines n'exprimant plus RFX6. Cette souris nous permettra également d'étudier l'impact de la perte de fonction de *Rfx6* sur les îlots de Langerhans au cours du développement pancréatique. La souris mutante pour *Rfx6* nous permettra également de déterminer spécifiquement quelle est la fonction de RFX6 chez l'adulte dans les îlots de Langerhans, et si RFX6 est impliqué dans l'homéostasie du glucose sanguin. Pour cela, la souris *Rfx6*^{LoxP/LoxP} pourra être croisée avec une souris exprimant une forme inducible de la Cre sous le contrôle du promoteur de *Pdx1*, permettant d'étudier la fonction de RFX6 dans l'ensemble des îlots de Langerhans (Herrera et al., 2000). Le mutant conditionnel de *Rfx6* pourra également être croisé avec une souris exprimant la Cre sous le contrôle du promoteur de l'insuline afin d'étudier spécifiquement le rôle de RFX6 dans les cellules β des îlots de Langerhans (Herrera, 2000).

Une caractérisation plus précise de l'expression et de la fonction de RFX6 dans l'estomac et l'intestin sera également nécessaire pour y déterminer si RFX6 est impliqué dans la différenciation de l'ensemble des sous types endocriniens gastriques et intestinaux comme le suggère la perte d'expression de *Rfx6* dans l'intestin des mutants *Neurog3*. Car bien que l'expression de *Neurog3* soit requise pour la différenciation de l'ensemble des cellules endocrines intestinales, seule une partie des cellules endocrines gastriques se différencie sous le contrôle de NEUROG3 (Jenny et al., 2002). Pour cela, la génération d'une souris exprimant la Cre sous le contrôle du promoteur de RFX6 sera nécessaire afin de permettre des expériences de traçage cellulaire. Par ailleurs, la souris *Rfx6*^{LoxP/LoxP} pourrait également être utilisée pour étudier la fonction de RFX6 dans l'épithélium intestinal. Pour cela, la souris *Rfx6*^{LoxP/LoxP} pourra être croisée avec une souris exprimant la Cre sous le contrôle du promoteur de la villin, permettant d'exprimer la Cre dans l'épithélium intestinal (El marjou et al., 2000).

Au-delà de ces études fonctionnelles, il sera également intéressant de définir comment RFX6 agit sur ces gènes cibles. Pour cela, la caractérisation des partenaires protéiques de RFX6 sera requise. En effet, les protéines de la famille RFX sont caractérisées pour agir en homo- ou hétéro-dimère (Katan-Khaykovich et al., 1998). Il sera ainsi intéressant de déterminer par co-immunoprécipitation l'éventuelle interaction dans le pancréas entre RFX6 et RFX3, ainsi qu'avec d'autres membres de la famille RFX (Ait-Lounis et al., 2007). La caractérisation des gènes cibles de RFX6 nous permettra en outre de déterminer les cascades génétiques induites par RFX6 au cours de la différenciation du pancréas endocrine. Nous pourrions également déterminer si RFX6 est impliqué dans l'expression des gènes régulant la formation de cils sur les cellules pancréatiques endocrines. Pour cela, nous pourrions nous baser sur la comparaison des transcriptomes des souris précédemment décrites.

Finalement, au-delà de l'étude de ces deux gènes, la caractérisation des autres gènes mis en évidence dans cette étude transcriptomique, permettra de mieux définir les cascades génétiques nécessaires au développement des cellules pancréatiques endocrines et à la purification de cellules β issues de la différenciation des cellules souches. On peut notamment citer le gène *expressed sequence AI987662* (ou *Hepacam 2*) qui après analyse bioinformatique a été caractérisé comme pouvant être une protéine membranaire contenant un domaine immunoglobuline. L'expression de ce gène est induite 6 fois chez les souris *Neurog3*^{-/-} ; *Pdx1-Neurog3ER*TM injectées avec le 4-hydroxytamoxifène, enrichi plus de 29 fois dans les progéniteurs pancréatiques endocrines et perdue chez les souris *Neurog3*^{-/-}. Ce gène pourrait ainsi être utilisé comme marqueur des cellules β issues de la différenciation des cellules souches afin de les purifier.

Par ailleurs, afin de définir comment est régulée l'expression de ces gènes et si NEUROG3 régule directement ces gènes, il serait intéressant de définir si ces gènes sont des cibles directes de NEUROG3. Pour cela, face à l'absence d'anticorps contre NEUROG3 fonctionnant en chromatine immunoprécipitation, il pourrait être intéressant d'utiliser la technique utilisée par Mellitzer et al. qui avait permis de montrer que *Insm1* est une cible directe de NEUROG3 (Mellitzer et al., 2006).

Une étude détaillée du promoteur de *Rfx6* pourrait également nous permettre de caractériser les facteurs de transcription nécessaires à la régulation de son expression au cours du développement embryonnaire et dans les îlots de Langerhans.

En conclusion, ce travail a conduit à la caractérisation du répertoire des gènes activés en aval de *NEUROG3* lors de la différenciation du pancréas endocrine. Parmi ce répertoire, nous avons montré que *RFX6*, un facteur de transcription de la famille *RFX*, dont la fonction n'était pas connue avant cette étude, requiert *NEUROG3* pour son expression pancréatique ; et chez le poisson zèbre, *zRFX6* est nécessaire à la différenciation normale des cellules pancréatique endocrines. Par ailleurs, aujourd'hui, bien que la formation de cellules sécrétant l'insuline à partir de cellules souches soit possible *in vitro*, ces cellules ne peuvent réguler leur sécrétion en fonction du taux de glucose (D'Amour et al., 2006). La caractérisation des cascades génétiques induites par *NEUROG3*, ainsi que par ces gènes cibles tel que *RFX6* devraient nous permettre de préciser les gènes et les voies de signalisation impliqués dans la différenciation et la maturation des cellules β .

MATERIEL ET METHODES

A) Matériel biologique

1. Les lignées murines

La lignée de souris transgénique Pdx1-Neurog3ERTM-IRES-EGFP nous a été donnée par le Dr Anne Grapin-Botton (EPFL, Lausanne ; Johansson et al., 2007). La lignée de souris *Neurod1^{+/-LacZ}* nous a été donnée par le Dr Jacqueline E. Lee (Université du Colorado, Boulder ; Naya et al., 1997). La lignée de souris *Neurog3^{+/-}* (Gradwohl et al., 2000) a été établie à l'IGBMC. Ces lignées sont entretenues par croisement avec des souris CD1. Les souris CD1 constituent également la source de tissus sauvages. Les souris gestantes sont sacrifiées par dislocation cervicale. L'utérus de la mère est prélevé puis les embryons extraits et disséqués dans du PBS 1X froid. Pour déterminer le stade embryonnaire, les souris sont mises en croisement et le matin du bouchon vaginal est désigné comme le stade embryonnaire E0,5.

Les marquages réalisés sur les tissus *Arx^{-/-}* ou *Pax4^{-/-}* ont été faits sur des tractus digestifs embryonnaires fixés et congelés par l'équipe de Ahmed Mansouri (Max-Planck Institute, Göttingen ; Collombat et al., 2003 ; Collombat et al., 2003 ; Sosa-Pineda et al., 1997).

2. Les lignées cellulaires

2.A. culture de la lignée HEK293t

Ces cellules ont été établies à partir de cellules rénales embryonnaires humaines, exprimant l'antigène SV40 T. Ces cellules sont cultivées à 37°C, 5% CO₂ dans le milieu suivant : DMEM, sérum de veau fœtal 10%, gentamycine 0,04 mg/ml. Elles sont passées au trente cinquième une fois par semaine et le milieu est changé tous les trois jours.

2.B. culture de la lignée β TC3

Ces cellules ont été établies à partir d'insulinome pancréatique murin (Efrat et al., 1988). Ces cellules sont cultivées à 37°C, 5% CO₂ dans du milieu : DMEM, sérum de veau

foetal 10%, pénicilline-streptomycine 1%. Elles sont passées au quart une fois par semaine et le milieu est changé tous les trois jours.

B) Culture organotypique

1. Dissection et culture des épithélia pancréatiques

Les tractus digestifs sont disséqués au stade E11,5 dans du HBSS. Après une incubation des tractus dans du RPMI additionné de Pénicilline-Streptomycine 1 % et de collagénase (0,5 mg/ml) pendant cinq minutes à 37°C, 5 % CO₂, les épithélia pancréatiques dorsaux sont ensuite isolés par dissociation mécanique du mésenchyme. Ils sont mis en culture dans des boîtes quatre puits plastique de 2,4 cm² (NUNC, NuncleonTMΔ) sur une membrane FHLC 0,45 μm (Millipore). Chaque puits contient 600 μl de milieu de culture DMEM/HAM-F12 avec une concentration finale de glucose à 3,15g/l additionné de 10 % de sérum de veau foetal inactivé à la chaleur, de gentamycine 0,04 mg/ml. Les épithélia pancréatiques sont cultivés, regroupés jusqu'à 9 par puits, et placés dans un incubateur à 37°C, 5 % CO₂. Le milieu de culture est renouvelé tous les trois jours.

2. Infection des épithélia

Avant l'infection, un milieu de « pré-incubation » contenant 20 μg/ml de dextran DEAE et x μl de virus (350 ng de P24 (Protéine de la capsid) de lentivirus par épithélium) dilué dans 45 μl de milieu de culture d'épithélia pancréatiques est incubé quinze minutes dans un incubateur à 37°C, 5 % CO₂.

Dans un couvercle de plaque 96 puits, ce milieu de « pré-incubation » est ensuite déposé sur les épithélia pancréatiques disposés dans une goutte de γ μl de HBSS, afin de former une goutte finale de 90 μl. Cette goutte « épithélia-virus » est incubée pendant deux heures dans un incubateur à 37°C, 5 % CO₂. Les épithélia infectés sont ensuite mis en culture comme décrit précédemment (cf § B1).

C) Techniques d’histologie

Les tissus sont fixés dans une solution de PBS 1X contenant 4 % de paraformaldéhyde (PFA) à 4°C sous agitation pour les embryons et les tractus et, sur glace pour les épithélia. Après trois rinçages au PBS 1X de 5 min, les organes ou embryons sont incubés dans du sucrose 20% à 4°C sous agitation ou sur glace pour les épithélia, puis inclus dans du milieu de cryopréservation, le Tissu-Tek (Sakura) après avoir éliminé un maximum de sucrose. Les différents temps d’incubation dans la PFA et le sucrose en fonction de l’échantillon sont indiqués à la table 2. Les échantillons sont orientés dans les moules et mis à figer dans de la carboglace avant d’être stockés à -80°C. Les échantillons sont ensuite coupés à une épaisseur de 6 ou 10 µm puis recueillies sur des lames superfrost-plus (Kindler GmbH & Co).

Tissu ou Stade de développement		Temps d'incubation dans la paraformaldéhyde 4%	Temps d'incubation dans le sucrose
Epithélium pancréatique		10 minutes	15 minutes
Embryon	E8,5 (8S)	1 heure	1 à 2 heures
	E9,5	1 à 2 heures	2 à 4 heures
	E10,5	1 à 2 heures	2 à 4 heures
	E12,5	2 à 4 heures	2 à 4 heures
	E15,5	1 jour	1 jour
	E18,5	1 jour	1 jour
Adulte (organe)		1 jour	1 jour

Table 2 : Temps d’incubations pour l’inclusion des tissus.

D) Techniques d'immuno-histochimie

Après décongélation, les lames portant les sections de tissus sont rincées dans du PBS 1X.

1. Démasquage

Pour les expériences d'immuno-histochimie nécessitant un démasquage (pour les anticorps de lapin ou de rat anti-RFX3 et anti-RFX6), les lames sont immergées dans 200 ml de tampon citrate 10 mM pH 6 puis incubés dans un four à micro-ondes à 650 Watts pendant 4 minutes, puis 160 Watts pendant 15 minutes sous atmosphère humide. Les lames sont ensuite laissées à refroidir à température ambiante pendant au moins 30 minutes, puis rincées 3 fois 5 minutes dans du PBS 1X.

2. Blocage des sites aspécifiques et incubation de l'anticorps primaire

Les sites aspécifiques sont bloqués en présence d'une solution de PBST (PBS 1X, Triton X100 0,1%) contenant 5 % de sérum de chèvre (inactivé à la chaleur) pendant 1h à température ambiante. Après 3 lavages de 5 min au PBST pour éliminer l'excès d'anticorps non fixé, les sections sont incubées avec l'anticorps primaire comme décrit ci-dessous. L'ensemble des anticorps utilisés sont décrits à la table 3.

3. Anticorps secondaires et révélation

3.A. Immuno-histochimie à la peroxydase

Les coupes sont incubées en présence de l'anticorps secondaire biotinylé, dilué dans du PBST, pendant 1h à température ambiante. Après 3 lavages dans du PBST, les peroxydases endogènes sont inactivées pendant 30 min dans de l'eau oxygénée 0,5 % (diluée dans du méthanol) à température ambiante. Après 3 lavages dans du PBST, le complexe AB (avidine-biotine, ABC kit *Elite Vectastain*) est appliqué aux lames durant 1h à température ambiante. La révélation est faite après 3 lavages de 5 min dans du PBST, en déposant la solution contenant le chromogène DAB (3,3'-diamino-benzidine-tetra-hydrochloride ; DAB

0,3 mg/ml, H₂O₂ 0,03 %) dans du PBS 1X. Après 3 à 10 min de révélation en fonction du développement du signal, la réaction est stoppée par 2 brefs passages dans de l'eau. Les coupes sont séchées et montées avec de l'Eukitt (Labonord).

Anticorps Primaires	Production	Dilution	Origine
Anti-NEUROG3	Cochon d'Inde	1/1000	M. Sander (U. C. Irvine)
Anti-PDX1	Lapin	1/1000	C. Wright (Vanderbilt University)
Anti-PDX1	Cochon d'Inde	1/1000	BCBC
Anti-PTF1a	Lapin	1/1000	B. Breant (Hopital Robert Debré, Paris)
Anti-RFX3	Lapin	1/100	W. Reith (Université de Genève)
Anti-RFX6	Lapin	1/500	G. Gradwohl (IGBMC, Illkirch)
Anti-RFX6	Rat	1/200	G. Gradwohl (IGBMC, Illkirch)
Anti-insuline	Cochon d'Inde	1/1000	Linco
Anti-glucagon	Cochon d'Inde	1/2000	Linco
Anti-polypeptide pancréatique	Cochon d'Inde	1/1000	Linco
Anti-somatostatine	Lapin	1/200	Dako
Anti-Ki67	Souris	1/100	Novo Castra
Anti-GFP	Lapin	1/500	Molecular Probes

Anticorps Secondaires	Production	Dilution	Origine
Anti-Ig de Lapin-Alexa 488	Chèvre	1/1000	Molecular Probes
Anti-Ig de cochon d'Inde-Cy3	Chèvre	1/1000	Jackson ImmunoResearch laboratories
Anti-Ig de lapin-Cy3	Chèvre	1/1000	Jackson ImmunoResearch laboratories
Anti-Ig de souris-Cy3	Chèvre	1/1000	Jackson ImmunoResearch laboratories
Anti-Ig de lapin biotinylé	Chèvre	1/200	Vector Laboratories
Anti-Ig de rat biotinylé	Chèvre	1/200	Vector Laboratories

Table 3 : Anticorps primaires et secondaires.

3.B. Immunofluorescence

L'anticorps secondaire, porteur du fluorochrome est dilué dans du PBST, et incubé 1h à température ambiante sur les lames. Cette étape et toutes les suivantes se font à l'abri de la lumière. L'ensemble des anticorps utilisés sont décrits à la table 3. Pour l'anticorps de lapin anti-RFX3 et l'anticorps de rat anti-RFX6, le signal est amplifié en révélant avec un anticorps couplé à la biotine dirigé contre les immunoglobulines de lapin ou de rat, puis avec un

conjugué Streptavidin-Alexa 488 (1:500, Molecular Probes). Après trois lavages de 5 min au PBST, les noyaux sont marqués 5 min à température ambiante avec du Dapi (Sigma), dilué au 1/50000^e dans du PBS 1X. Après 2 lavages dans du PBS 1X, les lames sont montées avec une solution d’Aqua Polymount (Polyscience).

4. Acquisition des images

Les images ont été prises avec un microscope à fluorescence Leica DMIRE avec un filtre GFP, Cy3 ou Dapi, grâce à une caméra CoolSnap HQ et le logiciel MetaVue V 6.1r3, ou sur un microscope DMIRE2 équipé d'une tête confocale Leica TCS SP2 AOBS, ou un microscope Zeiss axiophot équipé de la caméra Olympus Colour View et du logiciel DP-Soft 5.0.

E) Analyse quantitative

1. Dans le pancréas embryonnaire

Pour déterminer le nombre de cellules α ou β par pancréas chez les embryons, des coupes congelées sériées (séries de deux lames) de 6 μm sont réalisées sur la totalité du pancréas. Une analyse immuno-histochimique est ensuite réalisée pour chaque antigène sur des sections adjacentes et les cellules positives comptées.

2. Dans les épithélia pancréatiques embryonnaires après la culture

Pour déterminer le nombre de cellules GFP et sécrétant l’insuline ou le glucagon par épithélium, des coupes congelées de 6 μm sont réalisées sur la totalité de l’épithélium. Une analyse immuno-histochimique est ensuite réalisée sur chaque section. Après acquisition des images, la surface de marquage pour chaque antigène est déterminée avec le logiciel Metamorph.

F) Technique d'hybridation *in situ*

1. Synthèse de la sonde ARN

La synthèse de la sonde ARN est réalisée par transcription *in vitro* à partir de la matrice d'ADNc : *Neurog3* (Gradwohl et al., 2000); *NeuroD1* (Gradwohl et al., 2000), *Insm1* (Mellitzer et al., 2006), *Rfx3* (transcrit d'un fragment d'ADNc de 2.7 kb, Clone image 4483833, IRAV35-E2), *Rfx6* (transcrit d'un fragment d'ADNc de 0.9 kb, cloné à partir d'ARN pancréatique à E13.5 avec les oligonucléotides 5' c ggaattcgccacgtggagacatcctat et 3' ggactagtaatctgggttgcaagttgg) et *RIKEN cDNA 1700086L19* (transcrit d'un fragment d'ADNc de 0.65 kb, cloné à partir d'ARN pancréatique à E13.5 avec les oligonucléotides 5' CGGAATTCAGGCGCTTCAGAGAAGG et 3' GGACTAGTACTGGGAACAGCACTCC).

La synthèse de la sonde se fait dans un volume total de 20 µl à partir de 0,5 µg de plasmide linéarisé en présence de tampon de transcription 1X, des 4 ribonucléotides (ATP, CTP, GTP 1mM chacun; UTP 0,6mM), dUTP-dig 0,4mM (Boehringer), de RNasine 1 U/µl (Promega), de DTT 12,5mM (Boehringer) et de 2 U/µl de Sp6 ou T7 RNA Polymérase (NEB BioLabs). La transcription se fait à 37°C durant 2h et est arrêtée par ajout d'EDTA 20mM final. La sonde est précipitée 30 min à -80°C, en présence de LiCl 0,4M et de 2 volumes d'éthanol. Après une centrifugation de 15 min à 20800 g, le surnageant est éliminé et le culot repris dans 22µl de H₂O. 1µl est analysé sur gel d'agarose afin de déterminer la concentration de la sonde. Une seconde précipitation est faite en ajoutant du LiCl 0,4 M et 2 volumes d'éthanol. Le culot est repris dans 20 µl d'eau DEPC et 180 µl de tampon d'hybridation.

Tampon d'hybridation: tampon A 1X, formamide 50%, sulfate de dextran 10%, tRNA 1mg/ml, Denhardt's 1X. Tampon A 10X: NaCl 1,95 M, Tris HCl 90mM, Tris Base 11mM, Na₂HPO₄ 2H₂O 50 mM, NaH₂PO₄ 50 mM, EDTA 50 mM ; pH7,5. Solution de Denhardt's 100X: BSA 2 %, Ficoll 2 %, PVP 2 % (polyvinylpyrrolidone) stocké à -20°C.

2. Hybridation *in situ*

L'étape d'hybridation se fait en ajoutant la sonde, diluée au 1/100^e, dans du tampon d'hybridation préchauffé à 65°C, et dénaturée 10 min à 70°C, sur les coupes de tissu.

L'hybridation de la sonde fait pendant deux jours à 65 °C entre lame et lamelle dans une boîte saturée en 5X SSC / 50% formamide. Trois passages (1x15 min, 2x30 min) dans la solution de lavage (1X SSC, 50% formamide, 0,1 % Tween) à 65°C permettent d'éliminer l'excès de sonde non hybridée puis les coupes sont incubées 2x30 min dans du MABT 1X (Acide Maléique 100 mM, NaCl 150 mM, Tween 0,1 %; pH 7,5) à température ambiante. Les sites aspécifiques pour l'anticorps anti-dig sont bloqués par incubation dans une solution de blocage à température ambiante (réactif de blocage 2 % (Boehringer), sérum de chèvre 20 %, MABT 1X). L'anticorps anti-dig dilué dans cette même solution (1/2500^e, Roche Diagnostics) est appliqué sur les lames ; la réaction se fait pendant 12h à température ambiante dans une chambre humide. On procède ensuite à 4 lavages de 20 min dans du MABT à température ambiante puis 2 incubations dans du tampon de révélation de phosphatase alcaline en absence de substrat (NaCl 100 mM, MgCl₂ 50 mM, Tris pH9,5 100 mM, Tween 20 0.1 %, levamisole 0,5 mg/ml;). La révélation se fait durant 1 à 7 jours en fonction du développement du signal à l'obscurité, par une réaction colorée utilisant le bleu de tétrazolium (NBT) 262,5 µg/ml et le 5-Bromo,4-Chloro,3-Indolyphosphate (BCIP) 175 µg/ml dilués dans le tampon de révélation. La révélation est arrêtée par 2 passages rapides dans de l'eau. Les coupes sont séchées et montées avec de l'Eukitt (Labonord).

Solution de SSC 20X : NaCl 3 M, citrate de sodium 0,3 M.

G) Techniques de biologie moléculaire

1. Génotypage

1.A Préparation de l'ADN génomique

Pour déterminer les génotypes, l'ADN génomique est préparé à partir d'une biopsie de queue d'embryon ou de souris âgées de 2 semaines.

Les queues embryonnaires sont digérées par la protéinase K (100 µg/ml) dans 300 µl de tampon de digestion (KCl 50 mM, Tris 10mM pH 8,3, MgCl₂ 2 mM, gélatine 0,1mg/ml, Nonidet-P40 0,45 %, Tween 20 0,45 %) pendant 12H à 55°C. Ce tampon sans SDS permet d'analyser directement l'échantillon par PCR sans purifier l'ADN. Après la digestion, la protéinase K est inactivée par chauffage à 95°C durant 10 min.

Les queues de souriceaux sont digérées par la protéinase K (100 µg/ml) dans 500 µl de tampon de digestion (Tris 100mM pH 8.5, NaC 200mM, EDTA 1, 5mM, SDS 0.2%) pendant 12H à 55°C. Après une centrifugation de 10 min à 20800 g, 500 µl d'isopropanol sont ajoutés au surnageant. Après une centrifugation de 10 min à 20800 g et un lavage du culot dans l'éthanol à 70%, l'ADN est repris dans 200 µl de tampon Tris-EDTA (10/1 ; pH=8).

1.B Réaction de polymérisation en chaîne (PCR)

Pour les allèles *Neurog3* sauvage et mutant, 2µl d'ADN génomique sont mélangés à 48 µl d'une solution contenant : les dNTP à 0.2 mM chacun, le tampon PCR 1X (10 mM Tris-HCl pH 8,8, 1,5 mM MgCl₂, 50 mM KCl et 0,1 % Triton X-100) chacun des 2 oligonucléotides à 0,5 µM, et l'ADN Polymérase DyNazyme™ II (FINNZYMES) à 0,02 U/µl. Pour le transgène *Pdx1-Neurog3ERTM* et pour les allèles *Neurod1* sauvage et mutant, 2 µl d'ADN génomique sont mélangés à 12,5 µl d'un tampon commercial (Red'y'Starmix, EUROGENTEC). Chacun des 2 oligonucléotides est ajouté à une concentration finale de 1µM et le volume réactionnel complété à 25 µl par de l'H₂O.

Les réactions de PCR sont réalisées selon les conditions indiquées dans la table 4. Les analyses des produits de PCR se font par électrophorèse sur un gel d'agarose 2 % contenant 1 ng/ml de Bromure d'Ethidium dans du tampon TAE.

Gène	Oligonucléotide 5'/3'	Polymérase	cycle
<i>Ngn3</i> sauvage	Ngn3 5' : TCTCGCCTCTTCTGGCTTTC Ngn3 3' : CGGCAGATTTGAATGAGGGC	DyNazyme™II	94°C 3 min, (94°C 30 sec, 55°C 30 sec, 72°C 1 min) 35 cycles, 72°C 7 min.
<i>Ngn3</i> mutant	Neo 5' : GCAGCGCATCGCCTTCTATC Ngn3 3' : CGGCAGATTTGAATGAGGGC	DyNazyme™II	94°C 3 min, (94°C 30 sec, 55°C 30 sec, 72°C 1 min) 38 cycles, 72°C 7 min.
<i>Gfp</i>	Gfp 5' : CCTGAAGTTCATCTGCACCAC Gfp 3' : TTGTAGTTGTA CTCCAGCTTGTGC	HotGoldStar	94°C 10 min, (94°C 1 min, 55°C 1 min, 72°C 1 min) 35 cycles, 72°C 7 min.
<i>NeuroD1</i> sauvage	NeuroD1 5' : AAAGCCCCCTAACTGACTG NeuroD1 3' : TGATCCCCTGTTTCTTCC	DyNazyme™II	94°C 3 min, (94°C 30 sec, 55°C 30 sec, 72°C 1 min) 35 cycles, 72°C 7 min.
<i>NeuroD1</i> mutant	LacZ 5' : GCACATCCCCCTTTCGCCAGCTGGCGTAAT LacZ 3' : CGCGTCTGGCCTTCCTGTAGCCAGCTTCA	HotGoldStar	94°C 10 min, (94°C 1 min, 60°C 1 min, 72°C 1 min) 30 cycles, 72°C 7 min.

Table 4 : Oligonucléotides et conditions de RT-PCR pour le géotypage.

2. Extraction d'ARN

Après culture, les explants pancréatiques sont prélevés, rincés avec du PBS 1X et mis dans 300 µl de solution d'extraction (Tri-Reagent, Molecular Research Center). Le tissu est ensuite homogénéisé et les cellules lysées par trituration (pipetages et refoulements successifs). Pour les ARN issus de tissus embryonnaires, le lysat cellulaire est préparé comme pour les ARN issus d'explants en culture. Pour l'extraction d'ARN de pancréas adulte, les tissus sont placés pendant la dissection dans une solution de PBS 1X complétement en RNasin Ribonucléase Inhibitor (Promega, 4,8U/ml) ; puis rincé deux fois après la dissection du pancréas dans une solution de PBS 1X complétement en RNasin Ribonucléase Inhibitor (Promega, 4,8U/ml). Le pancréas est ensuite placé dans 8 ml de solution d'extraction, et broyé.

Le lysat cellulaire subit ensuite deux extractions-précipitations successives afin de purifier les ARN des protéines et de l'ADN génomique. La première extraction se fait par ajout de chloroforme (un cinquième du volume de Tri-Reagent) au lysat cellulaire. Après homogénéisation, repos pendant 2 min et centrifugation (20800 g, 15 min), le surnageant est prélevé et l'ARN précipité par ajout d'isopropanol (un demi volume de Tri-Reagent) et incubé à -20 °C pendant 12h. 0,25 µl de glycogène 20 µg/µl et 0,25 µl de Pellet Paint (Novagen) sont ajoutés pour les explants pancréatiques. Après centrifugation (20800 g, 30 min) puis un lavage dans 500 µl d'éthanol 75%, le culot est resuspendu dans de l'eau DEPC (AMBION). La seconde extraction est réalisée par l'ajout d'un volume de Phénol:Chlorophorme (5 :1) pH 4,5 puis centrifugation. Finalement, l'ARN est précipité par ajout de NaAc 0,6 M et 2,5 volumes d'éthanol 100% à -20°C pendant 12h. Après centrifugation, le culot d'ARN obtenu est lavé dans l'éthanol 75% et finalement dissout dans de l'eau DEPC.

3. Reverse-transcription (RT)/qPCR

3.A Transcription Reverse

La réaction commence par une incubation à 65 °C pendant 5 min dans un mélange de 13 µl contenant 250 ng d'ARN totaux, dNTP 0,77 mM et d'oligodT 1,93 mM ; suivie d'une

incubation sur la glace d'une minute au moins. Un mélange de 7 μ l composé du tampon de transcription 1X, DTT 14 mM, RNasin Ribonuclease inhibitor (Promega) 40 unités et de la Reverse Transcriptase SuperScript III (Invitrogen) 200 unités est ensuite ajouté. Le volume réactionnel de 20 μ l obtenu est incubé pendant 1h à 50°C, puis 15 min à 70°C. Le témoin négatif de transcription reverse est réalisé dans les mêmes conditions mais sans la reverse transcriptase.

3.B qPCR

Les réactions de qPCR sont réalisées avec le kit « LightCycler 480 SYBR Green I Master » (Roche) sur un LightCycler 480 (Roche). Pour chaque réaction de qPCR, 1 μ l d'ADNc issu de la réaction de RT est mélangé à 9 μ l d'une solution contenant : chacun des 2 oligonucléotides à 0,5 μ M, et 5 μ l du master mix 2X et 3 μ l d'eau. Les oligonucléotides utilisés pour chaque gène sont décrits dans la table 5. Le protocole de PCR est celui décrit par le kit, et la température d'hybridation des oligonucléotides est de 60°C pour l'ensemble des couples.

Gène	Oligonucléotide 5'/3'
<i>Dll1</i>	qDll1 5' : TTGAGGACAGCTTCTCTTGC
	qDll1 3' : AGGCACAGTCATCCACATTG
<i>GAPDH</i>	qGAPDH 5' : CTGAGGACCAGGTTGTCTCC
	qGAPDH 3' : TGTGAGGGAGATGCTCAGTG
<i>Insm1</i>	qInsm1 5' : AAGTGCCACCCGTCTGAGAA
	qInsm1 3' : AGCTTTGCCCTCAAAGGA
<i>NeuroD1</i>	qNeurod1 5' : CTTGGCCAAGAACTACATCTGG
	qNeurod1 3' : GGAGTAGGGATGCACCGGGAA
<i>Nkx2.2</i>	qNkx2.2 5' : AAAGGTATGGAGGTGACGCCT
	qNkx2.2 3' : AGCTGTACTGGGCGTTGTA
<i>pax4</i>	qPax4 5' : ACCTCATCCCAGGCCTATCT
	qPax4 3' : AGGCCTCTTATGGCCAGTTT
<i>Rfx6</i>	qRfx6 5' : TCTGCACCCATTCTCAACTG
	qRfx6 3' : GAGCAATGTGTTGGAGCTGA

Table 5 : Oligonucléotides utilisés pour les qPCR.

4. Puce à ADN

L'ARN utilisé pour l'analyse par puce à ADN a été extrait comme décrit précédemment (cf § G)2). 100 ng d'ARN ont été utilisés pour une amplification linéaire (RiboAmp® OA RNA Amplification Kit, Arcturus). Puis les sondes d'ARNc ont été réalisées (Bioarray High yield RNA transcription labeling Kit, Enzo) et hybridées sur les puces à ADN Affymetrix murine 430 2.0. A l'exception de l'extraction de l'ARN, l'ensemble de ces étapes a été réalisé par la plateforme « biopuces » de l'IGBMC. La normalisation des données a été réalisée avec l'algorithme Mass5. Les sondes perdues et exprimés correspondent respectivement au critère « present call » et « absent call » de l'analyse des puces Affymetrix. Ce critère est défini par un test de wilcoxon sur la différence de signal pour une sonde donnée entre les oligonucléotides normaux et les oligonucléotides mutés composant cette sonde.

Les chiffres des analyses de puce à ADNc présentées dans ma thèse sont en valeur absolue.

L'analyse statistique des données a été réalisée à l'aide du logiciel Zoe (Dembele, 2007).

5. RACE

Afin de déterminer la partie 5' d'un transcrit, des expériences de RACE ont été réalisées avec le kit « 5' RACE System, Version 2.0 » (Invitrogen). Les oligonucléotides utilisés sont : GSP1, TTCTTCTTCTGCATATT ; GSP2, AATGCTGACTCAGACCGAAGA ; GSP nested, CGATCTAGGACAGCCTCGGAAGA.

6. Clonage

Le clonage des séquences utilisées pour les ISH a été réalisé à partir d'ARN issu de pancréas dorsaux à E13,5 avec la polymérase « Phusion » (Finnzyme) selon le protocole du kit, avec les oligonucléotides précédemment décrits (cf § F.1), avec une température d'hybridation des oligonucléotides de 60°C. Après purification sur gel de l'ADNc d'intérêt (avec le kit Qiaquick Gel Extraction Kit, Qiagen), un ajout de dATP est réalisé avec l'ADN Polymérase DyNazyme™ II (FINNZYMES). Après purification de l'ADNc (avec le kit Qiaquick

PCR purification Kit, Qiagen), l'ADNc est cloné dans le vecteur pCRII-TOPO avec le kit « TOPO TA cloning kit Dual Promoter » (Invitrogen).

Le clonage des séquences de *RIKEN cDNA 1700086L19* et de *Rfx6* a été réalisé comme pour le clonage des séquences utilisées pour les sondes d'ISH, avec les oligonucléotides décrits à la table 6

Gène	Oligonucléotide 5'/3'
<i>Rfx6</i>	Rfx6 5' : GACACAGAGCTCCTTCAAGC Rfx6 3' : CCATACATTCTTTTTTATTTTCCTC
<i>Riken L19</i>	Riken L19 5' : CGGAATTCAGGCGCTTCAGAGAAGG Riken L19 3' : GGACTAGTACTGGGAACAGCACTCC

Table 6 : Oligonucléotides utilisés pour les clonages.

7. Northern Blot

Pour le Northern Blot, 10 µg d'ARN sont incubés 10 min à 65°C en présence de bleu de dépôt (2 volumes du volume d'ARN), puis placés sur glace jusqu'au dépôt sur le gel. Ils migrent ensuite dans un gel agarose 1%, MOPS 1X, formaldéhyde 2M. Après migration, le gel est incubé dans du SSC 20X pendant 15 min. Le gel est ensuite transféré sur une membrane de nylon (Roche) par diffusion par capillarité de SSC20X pendant 14h dans un sandwich papier Watmann-gel-membrane de nylon-papier Watmann-papier filtre, déposé sur une feuille de papier filtre dont les rebords trempent dans le SSC20X. La membrane est ensuite exposée aux UV avec une puissance de 0,12 Joules.cm⁻². La présence d'ARN sur la membrane est ensuite vérifiée par coloration de la membrane en l'incubant 15 min dans l'acide acétique 5%, puis 10 min dans une solution de NaAc 0,5 M pH5, bleu de méthylène 0,004%, Acide acétique glacial 1% ; puis dans l'EUP.

Dans un tube à hybridation en rotation dans un four à hybridation chauffé à 68°C, la membrane est incubée 1h à 65°C. En parallèle, 50 ng de sonde /ml de solution d'hybridation sont dénaturés pendant 10 min à 65°C dans 100µl d'EUP. Après l'heure d'incubation de la

membrane dans la solution d'hybridation, la sonde dénaturée y est ajoutée et l'ensemble est incubé 14h à 68 °C.

Dans un plat en pyrex en agitation dans un four à hybridation chauffé à 68°C, la membrane est lavée 2 fois 15 min dans 500ml d'une solution SSC 2X, SDS 0,1% préchauffée à 65°C ; puis 2 fois 30 min dans 500ml d'une solution SSC 0,1X, SDS 0,1% préchauffée à 65°C. La membrane est ensuite rincée 1 min dans la solution de lavage à température ambiante. Dans un sac, la membrane est ensuite agitée 1 heure dans la solution de blocage ; puis 30 min dans la solution de blocage avec l'anticorps alcaline phosphatase-anti-dig dilué au 1/10000^e (Roche). La membrane est ensuite lavée 2 fois 15 min dans la solution de lavage ; puis incubée 2 min dans la solution de détection.

Dans un sac, la membrane est révélée par une incubation de 10 min à 37 °C dans la solution de détection avec le CSPD dilué au centième (Roche), puis par détection par un film radiographique.

La taille des bandes est déterminée à l'aide du logiciel Chemidoc (Biorad).

Bleu de dépôt : deux volumes d'ARN ; formamide déionisée 66%, formaldéhyde 7,5%, MOPS 1,3X, bleu de bromophénol, Xylène Xianol ; Solution d'hybridation : Roche Dig Easy Hyb Granules ; Solution de lavage : acide maléique 100 mM, NaCl 150 mM, pH 7.5, 0,3% Tween[®] 20 ; Tampon MOPS : MOPS 0,25 M, NaAc 0,05 M, EDTA 0,01 M, pH 7 ; Solution de SSC 20X : NaCl 3 M, citrate de sodium 0,3 M, pH 7 ; Solution de blocage : MAB 1X, Blocking Reagent 1 % (Roche) ; MAB 10X : acide maléique 100 mM, NaCl 150 mM, pH 7.5 ; Solution de détection : Tris-HCl 100 mM, NaCl 100 mM, pH 9.5

8. Extraction de protéines

L'extraction des protéines des cellules en culture se fait par l'incubation des cellules dans une solution de NaCl 150 mM, NP40 1%, Tris pH8 50 mM, inhibiteur de protéase 1X (Roche) pendant 30 minutes sur glace. Les cellules sont ensuite grattées et l'extrait cellulaire est centrifugé 15 min à 20800 g. Les protéines localisées dans le surnageant sont ensuite dosées par la technique de Bradford par un dosage de la DO à 280 nm d'une solution protéine/Protein Assay (Biorad) 1 : 1000.

9. Western Blot

Pour le western blot, 30 µg de protéines sont incubés dans le tampon Laemli pendant 5 min à 100°C. La solution Laemli-protéines est ensuite migrée dans un gel acrylamide 10 %, Lower tris 25 %, ammonium persulfate 0,1%, TEMED 0,0006 %. Les protéines sont ensuite transférées sur une membrane de nitrocellulose 0,45 µm (Wathman®). Après un rinçage dans le PBS, la membrane est incubée 12 h dans une solution de PBS 1X, Régilait 5% dans laquelle l'anticorps de lapin anti Rfx6 est dilué au millième. Après quatre rinçages dans du PBS 1X, 0,05% Tween 20 et un rinçage dans le PBS, la membrane est incubée 20 min dans une de solution PBS 1X, Régilait 0,5%. Après un rinçage dans le PBS, la membrane est incubée 20 min dans du PBS 1X dans lequel l'anticorps de chèvre anti Ig de lapin couplé à la peroxydase est dilué au vingt millième (Jackson ImmunoResearch Laboratories). La membrane est ensuite révélée avec un substrat de la peroxydase (Immobilon Western, Millipore) en présence d'un film radiographique (Kodak).

Tampon Laemli : Upper tris 25%, SDS 4% Glycérol 20%, β-mercaptoéthanol 10%, H₂O 41% ; Upper Tris : Tris-base 30,3 g, SDS 10% 20 ml, pH 6,8, QSP H₂O 500 ml ; Lower Tris : Tris-base 90,85 g, SDS 10% 20 ml, pH 8,8, QSP H₂O 500 ml.

10. Clonage et test des ShRNA

10.A Clonage des ShRNA et de leur séquence cible

Le Clonage des ShRNA a été réalisé par la dimérisation d'oligonucléotides, qui ont ensuite été clonés dans le vecteur pSuper (Fig. 35A). Les oligonucléotides utilisés pour les ShRNA Rfx6 #1 et ShRNA Rfx6 #2 sont décrits dans la table 7. La dimérisation des oligonucléotides est réalisée en présence de 125 pmoles de chaque oligonucléotide dans 50 µl de NaCl 100mM. L'ensemble est placé dans de l'eau à ébullition pendant 10 min, puis laissé refroidir doucement à température ambiante en les conservant dans l'eau. En parallèle, le vecteur pSuper (Dr. Philippe Ravassard, Hôpital de la Pitié Salpêtrière, Paris ; Fig. 35A) est digéré par BglIII et HindIII, puis déphosphorylé par deux incubations successives de 1h avec l'Antartic Phosphatase (New England Biolabs). Les dimères d'oligonucléotides sont ensuite insérés dans le vecteur pSuper linéarisé par une incubation de 16h en présence de

T4 DNA ligase (Invitrogen). Les clones d'intérêt sont ensuite sélectionnés par transformation bactérienne en présence du produit de ligation, suivi d'un crible de colonies.

ShRNA	Oligonucléotide 5'/3'
Rfx6 #1	ShRfx6 #1 5' : p-GATCTCCCCAAAGAGCAGGAACACTACAGAATTTTCAAGAGAAATTCTGTAGT TCCTGCTCTTTTTTTTA ShRfx6 #1 3' : p-AGCTTAAAAAAGAGCAGGAACACTACAGAATTTCTCTTGAAAATTCTGTAG TTCTGCTCTTTGGGGA
Rfx6 #2	ShRfx6 #2 5' : p-GATCTCCCCAACTCTAACTATCAGACTATGTTTCAAGAGAACATAGTCTGAT AGTTAGAGTTTTTTTA ShRfx6 #2 3' : p-AGCTTAAAAAAGACTCTAACTATCAGACTATGTTCTCTTGAAACATAGTCTGA TAGTTAGAGTTGGGGA

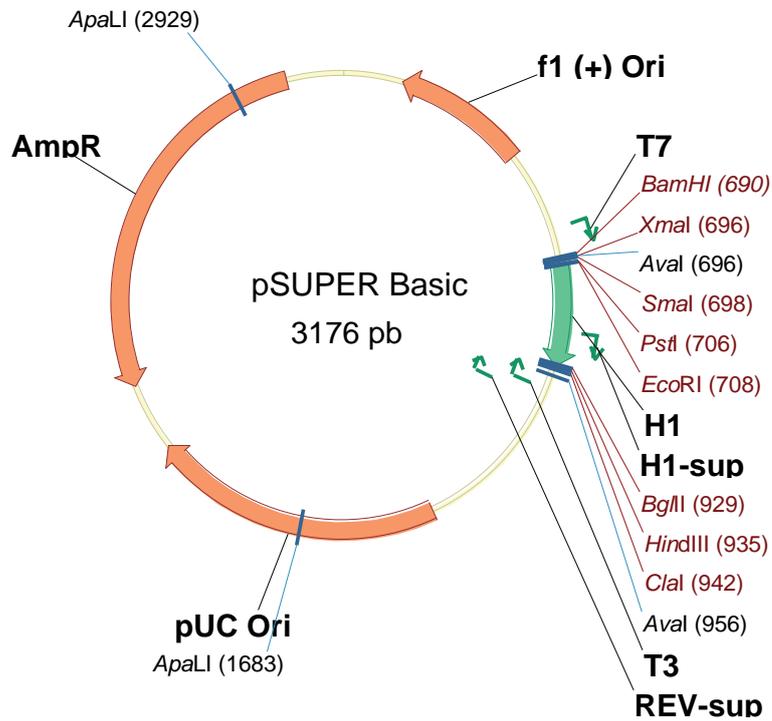
Table 7 : Oligonucléotides utilisés pour la synthèse des ShRNA.

Séquence cible ShRNA	Oligonucléotide 5'/3'
Rfx6 #1	Rfx6 #1 5' : CACCCCATTTTGC GAAGATTTGT Rfx6 #1 3' : GGACTAGTGCAATGTGTTGGAGCTGAGA
Rfx6 #2	Rfx6 #2 5' : CACCCCATTTTGC GAAGATTTGT Rfx6 #2 3' : GGACTAGTAATCTGGGTTTGCAAGTTGG

Table 8 : Oligonucléotides utilisés pour le clonage des séquences cibles des ShRNA.

Pour le clonage des séquences cibles des ShRNA, les fragments d'ADNc Rfx6#1 et Rfx6#2 ont été synthétisés par PCR avec la polymérase « Phusion » (Finnzymes) à partir des vecteurs pCRII-Topo-Rfx6#73 et pCRII-Topo-Rfx6#75 avec les oligonucléotides décrits dans la table 8. Les produits de PCR ont ensuite été insérés dans le vecteur pENTR/D-TOPO (Invitrogen). Les clones d'intérêt sont ensuite sélectionnés par transformation bactérienne en présence du produit de clonage, suivi d'un crible de colonies. Les clones obtenus sont ensuite sous cloné dans le vecteur pADRSV-IRES-GFP (Dr. Philippe Ravassard, Hôpital de la Pitié Salpêtrière, Paris ; Fig. 35B) par une recombinaison avec l'enzyme

A



B

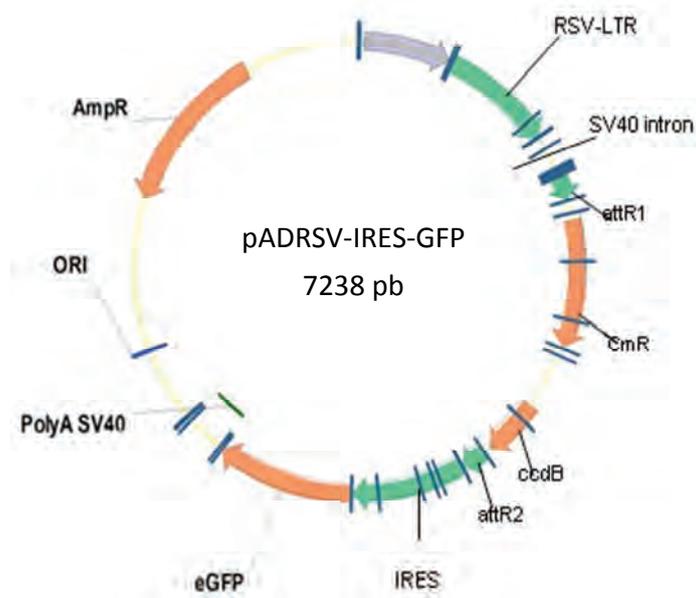


Figure 35 : Vecteurs utilisés pour le clonage et le test d'efficacité *in vitro* des ShRNA. (A) pSuper. (B) pADRSV-IRES-eGFP

LR clonase (Invitrogen). Les clones d'intérêt sont ensuite sélectionnés par transformation bactérienne en présence du produit de recombinaison, suivi d'un crible de colonies. Les séquences cibles des ShRNA Rfx6#1 et #2 sont ainsi clonés dans le vecteur pADRSV-IRES-GFP : pADRSV-Rfx6#1-IRES-GFP et pADRSV-Rfx6#2-IRES-GFP.

10.B Test de l'efficacité des ShRNA

Pour déterminer l'efficacité des ShRNA, 390000 cellules HEK293t sontensemencées dans des boîtes de 6 cm de diamètre (Falcon). Le lendemain, 250 ng de pADRSV-Rfx6-IRES-GFP avec 2,3 µg de pBluescript, de pSuper, de pSuper-ShRNA-Rfx6#1 ou de pSuper-ShRNA-Rfx6#2 sont transfectés dans les cellules HEK293t avec du JetPEI (Polyplus transfection). 1, 2 et 3 jours après transfection, les cellules sont trypsinisées, puis reprises dans une solution de PBS 1X, PFA 1%. Elles sont ensuite triées sur FACS calibre pour déterminer le taux de cellules GFP positives.

10.C Production des ShRNA

Après validation des ShRNA, la séquence H1-ShRNA est sous clonée par digestion du vecteur pSuper-ShRNA avec les enzymes de restriction BamHI et ClaI. En parallèle, le vecteur pTrip-CMV-eGFP-deltaU3 (Dr. Philippe Ravassard, Hôpital de la Pitié Salpêtrière, Paris ; Fig. 36) est digéré par BamHI et ClaI. La séquence H1-ShRNA est ensuite insérée dans le vecteur pTrip-CMV-eGFP-deltaU3 linéarisé par une incubation de 16h en présence de T4 DNA ligase (Invitrogen). Les clones d'intérêt sont ensuite sélectionnés par transformation bactérienne en présence du produit de ligation, suivi d'un crible de colonies.

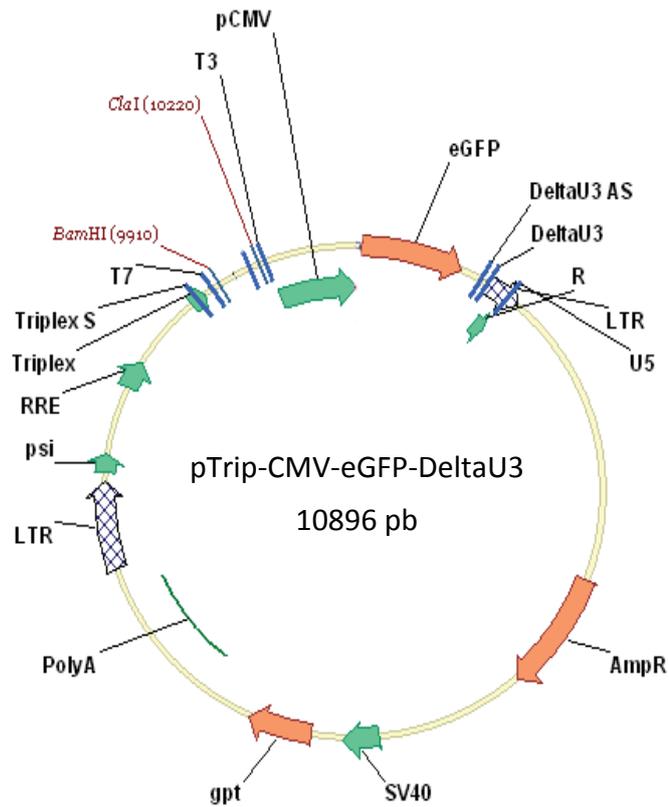


Figure 36 : Vecteur utilisé pour la production des ShRNA, pTrip-CMV-eGFP-DeltaU3.

H) Test Statistique

L'évidence statistique est déterminée par un test T de Student, à l'exception de l'analyse des études de puce à ADN.

I) Analyse bioinformatique

Les alignements de séquence et les arbres phylogénétiques ont été faits avec la suite de logiciel MyBioPack XP 3.6 (http://mybiopack.ifrance.com/mybiopack_fr.html).

REFERENCES

- Abdullah, J. M., Jing, X., Spassov, D. S., Nachtman, R. G., and Jurecic, R. (2001). Cloning and characterization of Hepp, a novel gene expressed preferentially in hematopoietic progenitors and mature blood cells. *Blood Cells Mol Dis* 27, 667-676.
- Aftab, S., Semenc, L., Chu, J. S., and Chen, N. (2008). Identification and characterization of novel human tissue-specific RFX transcription factors. *BMC Evol Biol* 8, 226.
- Ahlgren, U., Jonsson, J., and Edlund, H. (1996). The morphogenesis of the pancreatic mesenchyme is uncoupled from that of the pancreatic epithelium in IPF1/PDX1-deficient mice. *Development* 122, 1409-1416.
- Ahlgren, U., Jonsson, J., Jonsson, L., Simu, K., and Edlund, H. (1998). beta-cell-specific inactivation of the mouse *Ip1/Pdx1* gene results in loss of the beta-cell phenotype and maturity onset diabetes. *Genes Dev* 12, 1763-1768.
- Ahlgren, U., Pfaff, S. L., Jessell, T. M., Edlund, T., and Edlund, H. (1997). Independent requirement for ISL1 in formation of pancreatic mesenchyme and islet cells. *Nature* 385, 257-260.
- Ait-Lounis, A., Baas, D., Barras, E., Benadiba, C., Charollais, A., Nlend Nlend, R., Liegeois, D., Meda, P., Durand, B., and Reith, W. (2007). Novel function of the ciliogenic transcription factor RFX3 in development of the endocrine pancreas. *Diabetes* 56, 950-959.
- Akiyama, H., Kim, J. E., Nakashima, K., Balmes, G., Iwai, N., Deng, J. M., Zhang, Z., Martin, J. F., Behringer, R. R., Nakamura, T., and de Crombrughe, B. (2005). Osteo-chondroprogenitor cells are derived from Sox9 expressing precursors. *Proc Natl Acad Sci U S A* 102, 14665-14670.
- Andoniadou, C. L., Signore, M., Sajedi, E., Gaston-Massuet, C., Kelberman, D., Burns, A. J., Itasaki, N., Dattani, M., and Martinez-Barbera, J. P. (2007). Lack of the murine homeobox gene *Hesx1* leads to a posterior transformation of the anterior forebrain. *Development* 134, 1499-1508.
- Andralojc, K. M., Mercuri, A., Nowak, K. W., Albarello, L., Calcagno, R., Luzi, L., Bonifacio, E., Doglioni, C., and Piemonti, L. (2009). Ghrelin-producing epsilon cells in the developing and adult human pancreas. *Diabetologia* 52, 486-493.
- Ang, S. L., and Rossant, J. (1994). HNF-3 beta is essential for node and notochord formation in mouse development. *Cell* 78, 561-574.
- Apelqvist, A., Ahlgren, U., and Edlund, H. (1997). Sonic hedgehog directs specialised mesoderm differentiation in the intestine and pancreas. *Curr Biol* 7, 801-804.
- Apelqvist, A., Li, H., Sommer, L., Beatus, P., Anderson, D. J., Honjo, T., Hrabe de Angelis, M., Lendahl, U., and Edlund, H. (1999). Notch signalling controls pancreatic cell differentiation. *Nature* 400, 877-881.

- Araki, R., Takahashi, H., Fukumura, R., Sun, F., Umeda, N., Sujino, M., Inouye, S. T., Saito, T., and Abe, M. (2004). Restricted expression and photic induction of a novel mouse regulatory factor X4 transcript in the suprachiasmatic nucleus. *J Biol Chem* *279*, 10237-10242.
- Artner, I., Blanchi, B., Raum, J. C., Guo, M., Kaneko, T., Cordes, S., Sieweke, M., and Stein, R. (2007). MafB is required for islet beta cell maturation. *Proc Natl Acad Sci U S A* *104*, 3853-3858.
- Artner, I., Le Lay, J., Hang, Y., Elghazi, L., Schisler, J. C., Henderson, E., Sosa-Pineda, B., and Stein, R. (2006). MafB: an activator of the glucagon gene expressed in developing islet alpha- and beta-cells. *Diabetes* *55*, 297-304.
- Aughsteeen, A. A. (2001). The ultrastructure of primary cilia in the endocrine and excretory duct cells of the pancreas of mice and rats. *Eur J Morphol* *39*, 277-283.
- Baas, D., Meiniel, A., Benadiba, C., Bonnafe, E., Meiniel, O., Reith, W., and Durand, B. (2006). A deficiency in RFX3 causes hydrocephalus associated with abnormal differentiation of ependymal cells. *Eur J Neurosci* *24*, 1020-1030.
- Balemans, W., and Van Hul, W. (2002). Extracellular regulation of BMP signaling in vertebrates: a cocktail of modulators. *Dev Biol* *250*, 231-250.
- Bastien, J., and Rochette-Egly, C. (2004). Nuclear retinoid receptors and the transcription of retinoid-target genes. *Gene* *328*, 1-16.
- Bayha, E., Jorgensen, 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., Erler, T., Russell, A., and James, R. (1995). Expression of Cdx-2 in the mouse embryo and placenta: possible role in patterning of the extra-embryonic membranes. *Dev Dyn* *204*, 219-227.
- Beckers, A., Alten, L., Viebahn, C., Andre, P., and Gossler, A. (2007). The mouse homeobox gene *Noto* regulates node morphogenesis, notochordal ciliogenesis, and left right patterning. *Proc Natl Acad Sci U S A* *104*, 15765-15770.
- Beglinger, C., Taylor, I. L., Grossman, M. I., and Solomon, T. E. (1984). Pancreatic polypeptide inhibits exocrine pancreatic responses to six stimulants. *Am J Physiol* *246*, G286-291.
- Bennetts, J. S., Fowles, L. F., Berkman, J. L., van Bueren, K. L., Richman, J. M., Simpson, F., and Wicking, C. (2006). Evolutionary conservation and murine embryonic expression of the gene encoding the SERTA domain-containing protein CDCA4 (HEPP). *Gene* *374*, 153-165.
- Bergman, R. N., and Ader, M. (2000). Free fatty acids and pathogenesis of type 2 diabetes mellitus. *Trends Endocrinol Metab* *11*, 351-356.

- Bhushan, A., Itoh, N., Kato, S., Thiery, J. P., Czernichow, P., Bellusci, S., and Scharfmann, R. (2001). Fgf10 is essential for maintaining the proliferative capacity of epithelial progenitor cells during early pancreatic organogenesis. *Development* *128*, 5109-5117.
- Biemar, F., Argenton, F., Schmidtke, R., Epperlein, S., Peers, B., and Driever, W. (2001). Pancreas development in zebrafish: early dispersed appearance of endocrine hormone expressing cells and their convergence to form the definitive islet. *Dev Biol* *230*, 189-203.
- Blackshear, P. J., Graves, J. P., Stumpo, D. J., Cobos, I., Rubenstein, J. L., and Zeldin, D. C. (2003). Graded phenotypic response to partial and complete deficiency of a brain-specific transcript variant of the winged helix transcription factor RFX4. *Development* *130*, 4539-4552.
- Bonnafe, E., Touka, M., AitLounis, A., Baas, D., Barras, E., Ucla, C., Moreau, A., Flamant, F., Dubruille, R., Couble, P., Collignon, J., Durand, B., and Reith, W. (2004). The transcription factor RFX3 directs nodal cilium development and left-right asymmetry specification. *Mol Cell Biol* *24*, 4417-4427.
- Bort, R., Martinez-Barbera, J. P., Beddington, R. S., and Zaret, K. S. (2004). Hex homeobox gene-dependent tissue positioning is required for organogenesis of the ventral pancreas. *Development* *131*, 797-806.
- Bottcher, R. T., and Niehrs, C. (2005). Fibroblast growth factor signaling during early vertebrate development. *Endocr Rev* *26*, 63-77.
- Bouwmeester, T., Kim, S., Sasai, Y., Lu, B., and De Robertis, E. M. (1996). Cerberus is a head-inducing secreted factor expressed in the anterior endoderm of Spemann's organizer. *Nature* *382*, 595-601.
- Buse, J. B., Rosenstock, J., Sesti, G., Schmidt, W. E., Montanya, E., Brett, J. H., Zychma, M., and Blonde, L. (2009). Liraglutide once a day versus exenatide twice a day for type 2 diabetes: a 26-week randomised, parallel-group, multinational, open-label trial (LEAD-6). *Lancet* *374*, 39-47.
- Calgaro, S., Boube, M., Cribbs, D. L., and Bourbon, H. M. (2002). The *Drosophila* gene *taranis* encodes a novel trithorax group member potentially linked to the cell cycle regulatory apparatus. *Genetics* *160*, 547-560.
- Cano, D. A., Sekine, S., and Hebrok, M. (2006). Primary cilia deletion in pancreatic epithelial cells results in cyst formation and pancreatitis. *Gastroenterology* *131*, 1856-1869.
- Chao, C. S., Loomis, Z. L., Lee, J. E., and Sussel, L. (2007). Genetic identification of a novel NeuroD1 function in the early differentiation of islet alpha, PP and epsilon cells. *Dev Biol* *312*, 523-532.

- Chen, Y., Pan, F. C., Brandes, N., Afelik, S., Solter, M., and Pieler, T. (2004). Retinoic acid signaling is essential for pancreas development and promotes endocrine at the expense of exocrine cell differentiation in *Xenopus*. *Dev Biol* 271, 144-160.
- Cheong, J. K., Gunaratnam, L., Zang, Z. J., Yang, C. M., Sun, X., Nasr, S. L., Sim, K. G., Peh, B. K., Rashid, S. B., Bonventre, J. V., Salto-Tellez, M., and Hsu, S. I. (2009). TRIP-Br2 promotes oncogenesis in nude mice and is frequently overexpressed in multiple human tumors. *J Transl Med* 7, 8.
- Chey, W. Y., and Escoffery, R. (1976). Secretion cells in the gastrointestinal tract. *Endocrinology* 98, 1390-1395.
- Cho, J. M., Song, D. J., Bergeron, J., Benlimame, N., Wold, M. S., and Alaoui-Jamali, M. A. (2000). RBT1, a novel transcriptional co-activator, binds the second subunit of replication protein A. *Nucleic Acids Res* 28, 3478-3485.
- Cissell, M. A., Zhao, L., Sussel, L., Henderson, E., and Stein, R. (2003). Transcription factor occupancy of the insulin gene in vivo. Evidence for direct regulation by Nkx2.2. *J Biol Chem* 278, 751-756.
- Clemente, F., de Caro, A., and Figarella, C. (1972). [Composition of human pancreatic juice. Immunoenzymological studies]. *Eur J Biochem* 31, 186-193.
- Collombat, P., Hecksher-Sorensen, J., Broccoli, V., Krull, J., Ponte, I., Mundiger, T., Smith, J., Gruss, P., Serup, P., and Mansouri, A. (2005). The simultaneous loss of Arx and Pax4 genes promotes a somatostatin-producing cell fate specification at the expense of the alpha- and beta-cell lineages in the mouse endocrine pancreas. *Development* 132, 2969-2980.
- Collombat, P., Hecksher-Sorensen, J., Krull, J., Berger, J., Riedel, D., Herrera, P. L., Serup, P., and Mansouri, A. (2007). Embryonic endocrine pancreas and mature beta cells acquire alpha and PP cell phenotypes upon Arx misexpression. *J Clin Invest* 117, 961-970.
- Collombat, P., Hecksher-Sorensen, J., Serup, P. and Mansouri, A. (2006). Specifying pancreatic endocrine cell fates. *Mech Dev* 123, 501-512.
- Collombat, P., Mansouri, A., Hecksher-Sorensen, J., Serup, P., Krull, J., Gradwohl, G., and Gruss, P. (2003). Opposing actions of Arx and Pax4 in endocrine pancreas development. *Genes Dev* 17, 2591-2603.
- Collombat, P., Xu, X., Ravassard, P., Sosa-Pineda, B., Dussaud, S., Billestrup, N., Madsen, O.D., Serup, P., Heimberg, H., and Mansouri, A. (2009). The Ectopic Expression of Pax4 in the Mouse Pancreas Converts Progenitor Cells into alpha and Subsequently beta Cells. *Cell* 138, 449-462.
- D'Amour, K. A., Bang, A. G., Eliazar, S., Kelly, O. G., Agulnick, A. D., Smart, N. G., Moorman, M. A., Kroon, E., Carpenter, M. K., and Baetge, E. E. (2006). Production of pancreatic

hormone-expressing endocrine cells from human embryonic stem cells. *Nat Biotechnol* 24, 1392-1401.

Davies, K., and Conlon, K. C. (2009). Neuroendocrine tumors of the pancreas. *Curr Gastroenterol Rep* 11, 119-127.

Dembele, D, Jost, B, Thibault-Carpentier, C and Kastner, P (2007). Zoe : a user friendly tool for filtering microarray data. *Proceedings JOBIM, Marseille 10-12 Juillet 2007*, 377-378.

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. *Mech Dev* 123, 42-55.

Deutsch, G., Jung, J., Zheng, M., Lora, J., and Zaret, K. S. (2001). A bipotential precursor population for pancreas and liver within the embryonic endoderm. *Development* 128, 871-881.

Dichmann, D. S., Miller, C. P., Jensen, J., Scott Heller, R., and Serup, P. (2003). Expression and misexpression of members of the FGF and TGFbeta families of growth factors in the developing mouse pancreas. *Dev Dyn* 226, 663-674.

Dimagno, M. J., and Dimagno, E. P. (2009). Chronic pancreatitis. *Curr Opin Gastroenterol*.

Dominguez-Bendala, J., Klein, D., Ribeiro, M., Ricordi, C., Inverardi, L., Pastori, R., and Edlund, H. (2005). TAT-mediated neurogenin 3 protein transduction stimulates pancreatic endocrine differentiation in vitro. *Diabetes* 54, 720-726.

Doria, A., Patti, M. E., and Kahn, C. R. (2008). The emerging genetic architecture of type 2 diabetes. *Cell Metab* 8, 186-200.

Dorn, A., Durand, B., Marfing, C., Le Meur, M., Benoist, C., and Mathis, D. (1987). Conserved major histocompatibility complex class II boxes--X and Y--are transcriptional control elements and specifically bind nuclear proteins. *Proc Natl Acad Sci U S A* 84, 6249-6253.

Doyle, M. J., Loomis, Z. L., and Sussel, L. (2007a). Nkx2.2-repressor activity is sufficient to specify alpha-cells and a small number of beta-cells in the pancreatic islet. *Development* 134, 515-523.

Doyle, M. J., and Sussel, L. (2007b). Nkx2.2 regulates beta-cell function in the mature islet. *Diabetes* 56, 1999-2007.

Du, A., Hunter, C. S., Murray, J., Noble, D., Cai, C. L., Evans, S. M., Stein, R., and May, C. L. (2009). Islet-1 is required for the maturation, proliferation and survival of the endocrine pancreas. *Diabetes*.

- Dubruille, R., Laurencon, A., Vandaele, C., Shishido, E., Coulon-Bublex, M., Swoboda, P., Couble, P., Kernan, M., and Durand, B. (2002). *Drosophila* regulatory factor X is necessary for ciliated sensory neuron differentiation. *Development* *129*, 5487-5498.
- Edlund, H. (2002). Pancreatic organogenesis--developmental mechanisms and implications for therapy. *Nat Rev Genet* *3*, 524-532.
- Efrat, S., Linde, S., Kofod, H., Spector, D., Delannoy, M., Grant, S., Hanahan, D., and Baekkeskov, S. (1988). Beta-cell lines derived from transgenic mice expressing a hybrid insulin gene-oncogene. *Proc Natl Acad Sci U S A* *85*, 9037-9041.
- Eizirik, D. L., Colli, M. L., and Ortis, F. (2009). The role of inflammation in insulinitis and beta-cell loss in type 1 diabetes. *Nat Rev Endocrinol* *5*, 219-226.
- Elghazi, L., Cras-Meneur, C., Czernichow, P., and Scharfmann, R. (2002). Role for FGFR2IIIb-mediated signals in controlling pancreatic endocrine progenitor cell proliferation. *Proc Natl Acad Sci U S A* *99*, 3884-3889.
- Emery, P., Durand, B., Mach, B., and Reith, W. (1996). RFX proteins, a novel family of DNA binding proteins conserved in the eukaryotic kingdom. *Nucleic Acids Res* *24*, 803-807.
- Esni, F., Ghosh, B., Biankin, A. V., Lin, J. W., Albert, M. A., Yu, X., MacDonald, R. J., Civin, C. I., Real, F. X., Pack, M. A., Ball, D. W., and Leach, S. D. (2004). Notch inhibits Ptf1 function and acinar cell differentiation in developing mouse and zebrafish pancreas. *Development* *131*, 4213-4224.
- Feng, C., Xu, W., and Zuo, Z. (2009). Knockout of the regulatory factor X1 gene leads to early embryonic lethality. *Biochem Biophys Res Commun* *386*, 715-717.
- Field, H. A., Dong, P. D., Beis, D., and Stainier, D. Y. (2003). Formation of the digestive system in zebrafish. II. Pancreas morphogenesis. *Dev Biol* *261*, 197-208.
- Figarella, C. (1973). Les protéines du suc pancréatique humain. *Arch Fr Mal App Dig* *62*, 337-353.
- Freedman, S. D., and Scheele, G. A. (1994). Acid-base interactions during exocrine pancreatic secretion. Primary role for ductal bicarbonate in acinar lumen function. *Ann N Y Acad Sci* *713*, 199-206.
- Gajiwala, K. S., and Burley, S. K. (2000a). Winged helix proteins. *Curr Opin Struct Biol* *10*, 110-116.
- Gajiwala, K. S., Chen, H., Cornille, F., Roques, B. P., Reith, W., Mach, B., and Burley, S. K. (2000b). Structure of the winged-helix protein hRFX1 reveals a new mode of DNA binding. *Nature* *403*, 916-921.

- Gannon, M., Ables, E. T., Crawford, L., Lowe, D., Offield, M. F., Magnuson, M. A., and Wright, C. V. (2008). *pdx-1* function is specifically required in embryonic beta cells to generate appropriate numbers of endocrine cell types and maintain glucose homeostasis. *Dev Biol* *314*, 406-417.
- Gannon, M., Gamer, L. W., and Wright, C. V. (2001). Regulatory regions driving developmental and tissue-specific expression of the essential pancreatic gene *pdx1*. *Dev Biol* *238*, 185-201.
- Gannon, M., Herrera, P. L., and Wright, C. V. (2000). Mosaic Cre-mediated recombination in pancreas using the *pdx-1* enhancer/promoter. *Genesis* *26*, 143-144.
- Garcia, M., Hernandez-Lorenzo, P., San Roman, J. I., and Calvo, J. J. (2008). Pancreatic duct secretion: experimental methods, ion transport mechanisms and regulation. *J Physiol Biochem* *64*, 243-257.
- Gierl, M. S., Karoulias, N., Wende, H., Strehle, M., and Birchmeier, C. (2006). The zinc-finger factor *Insm1* (IA-1) is essential for the development of pancreatic beta cells and intestinal endocrine cells. *Genes Dev* *20*, 2465-2478.
- Githens, S. (1988). The pancreatic duct cell: proliferative capabilities, specific characteristics, metaplasia, isolation, and culture. *J Pediatr Gastroenterol Nutr* *7*, 486-506.
- 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. *Proc Natl Acad Sci U S A* *97*, 1607-1611.
- Grapin-Botton, A., Majithia, A. R., and Melton, D. A. (2001). Key events of pancreas formation are triggered in gut endoderm by ectopic expression of pancreatic regulatory genes. *Genes Dev* *15*, 444-454.
- Green, R. P., Cohn, S. M., Sacchettini, J. C., Jackson, K. E., and Gordon, J. I. (1992). The mouse intestinal fatty acid binding protein gene: nucleotide sequence, pattern of developmental and regional expression, and proposed structure of its protein product. *DNA Cell Biol* *11*, 31-41.
- Greenwood, A. L., Li, S., Jones, K., and Melton, D. A. (2007). Notch signaling reveals developmental plasticity of Pax4(+) pancreatic endocrine progenitors and shunts them to a duct fate. *Mech Dev* *124*, 97-107.
- Gu, G., Dubauskaite, J., and Melton, D. A. (2002). Direct evidence for the pancreatic lineage: NGN3+ cells are islet progenitors and are distinct from duct progenitors. *Development* *129*, 2447-2457.
- Gu, G., Wells, J. M., Dombkowski, D., Preffer, F., Aronow, B., and Melton, D. A. (2004). Global expression analysis of gene regulatory pathways during endocrine pancreatic development. *Development* *131*, 165-179.

- Guz, Y., Montminy, M. R., Stein, R., Leonard, J., Gamer, L. W., Wright, C. V., and Teitelman, G. (1995). 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. *Development* *121*, 11-18.
- Hald, J., Sprinkel, A. E., Ray, M., Serup, P., Wright, C., and Madsen, O. D. (2008). Generation and characterization of Ptf1a antiserum and localization of Ptf1a in relation to Nkx6.1 and Pdx1 during the earliest stages of mouse pancreas development. *J Histochem Cytochem* *56*, 587-595.
- Harrison, K. A., Thaler, J., Pfaff, S. L., Gu, H., and Kehrl, J. H. (1999). Pancreas dorsal lobe agenesis and abnormal islets of Langerhans in Hlxb9-deficient mice. *Nat Genet* *23*, 71-75.
- Haumaitre, C., Barbacci, E., Jenny, M., Ott, M. O., Gradwohl, G., and Cereghini, S. (2005). Lack of TCF2/vHNF1 in mice leads to pancreas agenesis. *Proc Natl Acad Sci U S A* *102*, 1490-1495.
- Hayashi, S., and McMahon, A. P. (2002). Efficient recombination in diverse tissues by a tamoxifen-inducible form of Cre: a tool for temporally regulated gene activation/inactivation in the mouse. *Dev Biol* *244*, 305-318.
- Hebrok, M., Kim, S. K., and Melton, D. A. (1998). Notochord repression of endodermal Sonic hedgehog permits pancreas development. *Genes Dev* *12*, 1705-1713.
- Hebrok, M., Kim, S. K., St Jacques, B., McMahon, A. P., and Melton, D. A. (2000). Regulation of pancreas development by hedgehog signaling. *Development* *127*, 4905-4913.
- Heiser, P. W., Lau, J., Taketo, M. M., Herrera, P. L., and Hebrok, M. (2006). Stabilization of beta-catenin impacts pancreas growth. *Development* *133*, 2023-2032.
- Heller, R. S., Dichmann, D. S., Jensen, J., Miller, C., Wong, G., Madsen, O. D., and Serup, P. (2002). Expression patterns of Wnts, Frizzleds, sFRPs, and misexpression in transgenic mice suggesting a role for Wnts in pancreas and foregut pattern formation. *Dev Dyn* *225*, 260-270.
- Heller, R. S., Jenny, M., Collombat, P., Mansouri, A., Tomasetto, C., Madsen, O. D., Mellitzer, G., Gradwohl, G., and Serup, P. (2005). Genetic determinants of pancreatic epsilon-cell development. *Dev Biol* *286*, 217-224.
- Heller, R. S., Stoffers, D. A., Liu, A., Schedl, A., Crenshaw, E. B., 3rd, Madsen, O. D., and Serup, P. (2004). The role of Brn4/Pou3f4 and Pax6 in forming the pancreatic glucagon cell identity. *Dev Biol* *268*, 123-134.
- Henning, S.J. (1987). Functional development of the gastrointestinal tract, in *Physiology of the Gastrointestinal Tract*, 2nd ed (Johnson LR, ed). Raven Press, New York, 285-300.

- Henseleit, K. D., Nelson, S. B., Kuhlbrodt, K., Hennings, J. C., Ericson, J., and Sander, M. (2005). NKX6 transcription factor activity is required for alpha- and beta-cell development in the pancreas. *Development* *132*, 3139-3149.
- Heremans, Y., Van De Casteele, M., in't Veld, P., Gradwohl, G., Serup, P., Madsen, O., Pipeleers, D., and Heimberg, H. (2002). Recapitulation of embryonic neuroendocrine differentiation in adult human pancreatic duct cells expressing neurogenin 3. *J Cell Biol* *159*, 303-312.
- Herrera, P. L. (2000). Adult insulin- and glucagon-producing cells differentiate from two independent cell lineages. *Development* *127*, 2317-2322.
- Holland, A. M., Hale, M. A., Kagami, H., Hammer, R. E., and MacDonald, R. J. (2002). Experimental control of pancreatic development and maintenance. *Proc Natl Acad Sci U S A* *99*, 12236-12241.
- Horvath, G. C., Kistler, W. S., and Kistler, M. K. (2004). RFX2 is a potential transcriptional regulatory factor for histone H1t and other genes expressed during the meiotic phase of spermatogenesis. *Biol Reprod* *71*, 1551-1559.
- Hsu, S. I., Yang, C. M., Sim, K. G., Hentschel, D. M., O'Leary, E., and Bonventre, J. V. (2001). TRIP-Br: a novel family of PHD zinc finger- and bromodomain-interacting proteins that regulate the transcriptional activity of E2F-1/DP-1. *Embo J* *20*, 2273-2285.
- 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. *Mol Cell Biol* *20*, 3292-3307.
- Huang, M., Zhou, Z., and Elledge, S. J. (1998). The DNA replication and damage checkpoint pathways induce transcription by inhibition of the Crt1 repressor. *Cell* *94*, 595-605.
- Jacquemin, P., Lemaigre, F. P., and Rousseau, G. G. (2003). The Onecut transcription factor HNF-6 (OC-1) is required for timely specification of the pancreas and acts upstream of Pdx-1 in the specification cascade. *Dev Biol* *258*, 105-116.
- Jarikji, Z. H., Vanamala, S., Beck, C. W., Wright, C. V., Leach, S. D., and Horb, M. E. (2007). Differential ability of Ptf1a and Ptf1a-VP16 to convert stomach, duodenum and liver to pancreas. *Dev Biol* *304*, 786-799.
- Jenny, M., Uhl, C., Roche, C., Duluc, I., Guillermin, V., Guillemot, F., Jensen, J., Kedinger, M., and Gradwohl, G. (2002). Neurogenin3 is differentially required for endocrine cell fate specification in the intestinal and gastric epithelium. *Embo J* *21*, 6338-6347.
- Jensen, J., Heller, R. S., Funder-Nielsen, T., Pedersen, E. E., Lindsell, C., Weinmaster, G., Madsen, O. D., and Serup, P. (2000a). Independent development of pancreatic alpha- and beta-cells from neurogenin3-expressing precursors: a role for the notch pathway in repression of premature differentiation. *Diabetes* *49*, 163-176.

- Jensen, J., Pedersen, E. E., Galante, P., Hald, J., Heller, R. S., Ishibashi, M., Kageyama, R., Guillemot, F., Serup, P., and Madsen, O. D. (2000b). Control of endodermal endocrine development by Hes-1. *Nat Genet* 24, 36-44.
- Johansson, K. A., Dursun, U., Jordan, N., Gu, G., Beermann, F., Gradwohl, G., and Grapin-Botton, A. (2007). Temporal control of neurogenin3 activity in pancreas progenitors reveals competence windows for the generation of different endocrine cell types. *Dev Cell* 12, 457-465.
- Jonsson, J., Carlsson, L., Edlund, T., and Edlund, H. (1994). Insulin-promoter-factor 1 is required for pancreas development in mice. *Nature* 371, 606-609.
- Jorgensen, M. C., Ahnfelt-Ronne, J., Hald, J., Madsen, O. D., Serup, P., and Hecksher-Sorensen, J. (2007). An illustrated review of early pancreas development in the mouse. *Endocr Rev* 28, 685-705.
- Juhl, K., Sarkar, S. A., Wong, R., Jensen, J., and Hutton, J. C. (2008). Mouse pancreatic endocrine cell transcriptome defined in the embryonic Ngn3-null mouse. *Diabetes* 57, 2755-2761.
- Kankel, M. W., Hurlbut, G. D., Upadhyay, G., Yajnik, V., Yedvobnick, B., and Artavanis-Tsakonas, S. (2007). Investigating the genetic circuitry of mastermind in *Drosophila*, a notch signal effector. *Genetics* 177, 2493-2505.
- Karlsson, O., Thor, S., Norberg, T., Ohlsson, H., and Edlund, T. (1990). Insulin gene enhancer binding protein Isl-1 is a member of a novel class of proteins containing both a homeo- and a Cys-His domain. *Nature* 344, 879-882.
- Katan-Khaykovich, Y., and Shaul, Y. (1998). RFX1, a single DNA-binding protein with a split dimerization domain, generates alternative complexes. *J Biol Chem* 273, 24504-24512.
- Kawaguchi, Y., Cooper, B., Gannon, M., Ray, M., MacDonald, R. J., and Wright, C. V. (2002). The role of the transcriptional regulator Ptf1a in converting intestinal to pancreatic progenitors. *Nat Genet* 32, 128-134.
- Kim, S. K., Hebrok, M., Li, E., Oh, S. P., Schrewe, H., Harmon, E. B., Lee, J. S., and Melton, D. A. (2000). Activin receptor patterning of foregut organogenesis. *Genes Dev* 14, 1866-1871.
- Kim, S. K., Hebrok, M., and Melton, D. A. (1997). Notochord to endoderm signaling is required for pancreas development. *Development* 124, 4243-4252.
- Kim, S. K., and MacDonald, R. J. (2002). Signaling and transcriptional control of pancreatic organogenesis. *Curr Opin Genet Dev* 12, 540-547
- Kojima, M., Hosoda, H., Date, Y., Nakazato, M., Matsuo, H., and Kangawa, K. (1999). Ghrelin is a growth-hormone-releasing acylated peptide from stomach. *Nature* 402, 656-660.

- Krapp, A., Knofler, M., Frutiger, S., Hughes, G. J., Hagenbuchle, O., and Wellauer, P. K. (1996). The p48 DNA-binding subunit of transcription factor PTF1 is a new exocrine pancreas-specific basic helix-loop-helix protein. *Embo J* *15*, 4317-4329.
- Krapp, A., Knofler, M., Ledermann, B., Burki, K., Berney, C., Zoerkler, N., Hagenbuchle, O., and Wellauer, P. K. (1998). The bHLH protein PTF1-p48 is essential for the formation of the exocrine and the correct spatial organization of the endocrine pancreas. *Genes Dev* *12*, 3752-3763.
- Kreivi, J. P., and Lamond, A. I. (1996). RNA splicing: unexpected spliceosome diversity. *Curr Biol* *6*, 802-805.
- Kroon, E., Martinson, L. A., Kadoya, K., Bang, A. G., Kelly, O. G., Eliazar, S., Young, H., Richardson, M., Smart, N. G., Cunningham, J., Agulnick, A. D., D'Amour, K. A., Carpenter, M. K., and Baetge, E. E. (2008). Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells in vivo. *Nat Biotechnol* *26*, 443-452.
- Lai, E. C. (2004). Notch signaling: control of cell communication and cell fate. *Development* *131*, 965-973.
- Lammert, E., Brown, J., and Melton, D. A. (2000). Notch gene expression during pancreatic organogenesis. *Mech Dev* *94*, 199-203.
- Lammert, E., Cleaver, O., and Melton, D. (2001). Induction of pancreatic differentiation by signals from blood vessels. *Science* *294*, 564-567.
- Lau, J., Kawahira, H., and Hebrok, M. (2006). Hedgehog signaling in pancreas development and disease. *Cell Mol Life Sci* *63*, 642-652.
- Ledermann, H. M. (1995). Is maturity onset diabetes at young age (MODY) more common in Europe than previously assumed? *Lancet* *345*, 648.
- Li, H., Arber, S., Jessell, T. M., and Edlund, H. (1999). Selective agenesis of the dorsal pancreas in mice lacking homeobox gene Hlxb9. *Nat Genet* *23*, 67-70.
- Lioubinski, O., Muller, M., Wegner, M., and Sander, M. (2003). Expression of Sox transcription factors in the developing mouse pancreas. *Dev Dyn* *227*, 402-408.
- Logan, C. Y., and Nusse, R. (2004). The Wnt signaling pathway in development and disease. *Annu Rev Cell Dev Biol* *20*, 781-810.
- Lorent, K., Yeo, S. Y., Oda, T., Chandrasekharappa, S., Chitnis, A., Matthews, R. P., and Pack, M. (2004). Inhibition of Jagged-mediated Notch signaling disrupts zebrafish biliary development and generates multi-organ defects compatible with an Alagille syndrome phenocopy. *Development* *131*, 5753-5766.

- Lottmann, H., Vanselow, J., Hessabi, B., and Walther, R. (2001). The Tet-On system in transgenic mice: inhibition of the mouse *pdx-1* gene activity by antisense RNA expression in pancreatic beta-cells. *J Mol Med* 79, 321-328.
- Lubelsky, Y., Reuven, N., and Shaul, Y. (2005). Autorepression of *rfx1* gene expression: functional conservation from yeast to humans in response to DNA replication arrest. *Mol Cell Biol* 25, 10665-10673.
- Lum, L., and Beachy, P. A. (2004). The Hedgehog response network: sensors, switches, and routers. *Science* 304, 1755-1759.
- Lynn, F. C., Smith, S. B., Wilson, M. E., Yang, K. Y., Nekrep, N., and German, M. S. (2007). Sox9 coordinates a transcriptional network in pancreatic progenitor cells. *Proc Natl Acad Sci U S A* 104, 10500-10505.
- MacFarlane, W. M., Read, M. L., Gilligan, M., Bujalska, I., and Docherty, K. (1994). Glucose modulates the binding activity of the beta-cell transcription factor IUF1 in a phosphorylation-dependent manner. *Biochem J* 303 (Pt 2), 625-631.
- Marshak, S., Totary, H., Cerasi, E., and Melloul, D. (1996). Purification of the beta-cell glucose-sensitive factor that transactivates the insulin gene differentially in normal and transformed islet cells. *Proc Natl Acad Sci U S A* 93, 15057-15062.
- Marsich, E., Vetere, A., Di Piazza, M., Tell, G., and Paoletti, S. (2003). The PAX6 gene is activated by the basic helix-loop-helix transcription factor NeuroD/BETA2. *Biochem J* 376, 707-715.
- Martin, M., Gallego-Llamas, J., Ribes, V., Kedingler, M., Niederreither, K., Chambon, P., Dolle, P., and Gradwohl, G. (2005). Dorsal pancreas agenesis in retinoic acid-deficient *Raldh2* mutant mice. *Dev Biol* 284, 399-411.
- Matsuoka, T. A., Zhao, L., Artner, I., Jarrett, H. W., Friedman, D., Means, A., and Stein, R. (2003). Members of the large Maf transcription family regulate insulin gene transcription in islet beta cells. *Mol Cell Biol* 23, 6049-6062.
- Matsushita, H., Uenaka, A., Ono, T., Hasegawa, K., Sato, S., Koizumi, F., Nakagawa, K., Toda, M., Shingo, T., Ichikawa, T., Noguchi, Y., Tamiya, T., Furuta, T., Kawase, T., Date, I., and Nakayama, E. (2005). Identification of glioma-specific RFX4-E and -F isoforms and humoral immune response in patients. *Cancer Sci* 96, 801-809.
- Mavropoulos, A., Devos, N., Biemar, F., Zecchin, E., Argenton, F., Edlund, H., Motte, P., Martial, J. A., and Peers, B. (2005). *sox4b* is a key player of pancreatic alpha cell differentiation in zebrafish. *Dev Biol* 285, 211-223.
- McClintock, T. S., Glasser, C. E., Bose, S. C., and Bergman, D. A. (2008). Tissue expression patterns identify mouse cilia genes. *Physiol Genomics* 32, 198-206.

- Mellitzer, G., Bonne, S., Luco, R. F., Van De Casteele, M., Lenne-Samuel, N., Collombat, P., Mansouri, A., Lee, J., Lan, M., Pipeleers, D., Nielsen, F. C., Ferrer, J., Gradwohl, G., and Heimberg, H. (2006). IA1 is NGN3-dependent and essential for differentiation of the endocrine pancreas. *Embo J* 25, 1344-1352.
- Mellitzer, G., Martin, M., Sidhoum-Jenny, M., Orvain, C., Barths, J., Seymour, P. A., Sander, M., and Gradwohl, G. (2004). Pancreatic islet progenitor cells in neurogenin 3-yellow fluorescent protein knock-add-on mice. *Mol Endocrinol* 18, 2765-2776.
- Meyer, P. (1977). *Physiologie humaine*. Flammarion Medecine-Science, 248-279.
- Michael, M. D., Kulkarni, R. N., Postic, C., Previs, S. F., Shulman, G. I., Magnuson, M. A., and Kahn, C. R. (2000). Loss of insulin signaling in hepatocytes leads to severe insulin resistance and progressive hepatic dysfunction. *Mol Cell* 6, 87-97.
- Miralles, F., Czernichow, P., Ozaki, K., Itoh, N., and Scharfmann, R. (1999). Signaling through fibroblast growth factor receptor 2b plays a key role in the development of the exocrine pancreas. *Proc Natl Acad Sci U S A* 96, 6267-6272.
- Miralles, F., Czernichow, P., and Scharfmann, R. (1998). Follistatin regulates the relative proportions of endocrine versus exocrine tissue during pancreatic development. *Development* 125, 1017-1024.
- Molotkov, A., Molotkova, N., and Duyster, G. (2005). Retinoic acid generated by Raldh2 in mesoderm is required for mouse dorsal endodermal pancreas development. *Dev Dyn* 232, 950-957.
- Morotomi-Yano, K., Yano, K., Saito, H., Sun, Z., Iwama, A., and Miki, Y. (2002). Human regulatory factor X 4 (RFX4) is a testis-specific dimeric DNA-binding protein that cooperates with other human RFX members. *J Biol Chem* 277, 836-842.
- Murphy, K. G., and Bloom, S. R. (2004). Gut hormones in the control of appetite. *Exp Physiol* 89, 507-516.
- Murtaugh, L. C., Law, A. C., Dor, Y., and Melton, D. A. (2005). Beta-catenin is essential for pancreatic acinar but not islet development. *Development* 132, 4663-4674.
- Murtaugh, L. C., Stanger, B. Z., Kwan, K. M., and Melton, D. A. (2003). Notch signaling controls multiple steps of pancreatic differentiation. *Proc Natl Acad Sci U S A* 100, 14920-14925.
- 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 Dev* 11, 2323-2334.

- Nelson, S. B., Schaffer, A. E., and Sander, M. (2007). The transcription factors Nkx6.1 and Nkx6.2 possess equivalent activities in promoting beta-cell fate specification in Pdx1+ pancreatic progenitor cells. *Development* 134, 2491-2500.
- Nishimura, W., Kondo, T., Salameh, T., El Khattabi, I., Dodge, R., Bonner-Weir, S., and Sharma, A. (2006). A switch from MafB to MafA expression accompanies differentiation to pancreatic beta-cells. *Dev Biol* 293, 526-539.
- Offield, M. F., Jetton, T. L., Labosky, P. A., Ray, M., Stein, R. W., Magnuson, M. A., Hogan, B. L., and Wright, C. V. (1996). PDX-1 is required for pancreatic outgrowth and differentiation of the rostral duodenum. *Development* 122, 983-995.
- Ohlsson, H., Karlsson, K., and Edlund, T. (1993). IPF1, a homeodomain-containing transactivator of the insulin gene. *Embo J* 12, 4251-4259.
- Olbrot, M., Rud, J., Moss, L. G., and Sharma, A. (2002). Identification of beta-cell-specific insulin gene transcription factor RIPE3b1 as mammalian MafA. *Proc Natl Acad Sci U S A* 99, 6737-6742.
- Oster, A., Jensen, J., Edlund, H., and Larsson, L. I. (1998). Homeobox gene product Nkx 6.1 immunoreactivity in nuclei of endocrine cells of rat and mouse stomach. *J Histochem Cytochem* 46, 717-721.
- Ostrom, M., Loffler, K. A., Edfalk, S., Selander, L., Dahl, U., Ricordi, C., Jeon, J., Correa-Medina, M., Diez, J., and Edlund, H. (2008). Retinoic acid promotes the generation of pancreatic endocrine progenitor cells and their further differentiation into beta-cells. *PLoS One* 3, e2841.
- Papadopoulou, S., and Edlund, H. (2005). Attenuated Wnt signaling perturbs pancreatic growth but not pancreatic function. *Diabetes* 54, 2844-2851.
- Pedersen, J. K., Nelson, S. B., Jorgensen, M. C., Henseleit, K. D., Fujitani, Y., Wright, C. V., Sander, M., and Serup, P. (2005). Endodermal expression of Nkx6 genes depends differentially on Pdx1. *Dev Biol* 288, 487-501.
- Petri, A., Ahnfelt-Ronne, J., Frederiksen, K. S., Edwards, D. G., Madsen, D., Serup, P., Fleckner, J., and Heller, R. S. (2006). The effect of neurogenin3 deficiency on pancreatic gene expression in embryonic mice. *J Mol Endocrinol* 37, 301-316.
- Picard, D. (1994). Regulation of protein function through expression of chimaeric proteins. *Curr Opin Biotechnol* 5, 511-515.
- Polak, J. M., Bloom, S. R., Rayford, P. L., Pearse, A. G., Buchan, A. M., and Thompson, J. C. (1975). Identification of cholecystokinin-secreting cells. *Lancet* 2, 1016-1018.

- Prado, C. L., Pugh-Bernard, A. E., Elghazi, L., Sosa-Pineda, B., and Sussel, L. (2004). Ghrelin cells replace insulin-producing beta cells in two mouse models of pancreas development. *Proc Natl Acad Sci U S A* *101*, 2924-2929.
- Pulkkinen, M. A., Spencer-Dene, B., Dickson, C., and Otonkoski, T. (2003). The IIIb isoform of fibroblast growth factor receptor 2 is required for proper growth and branching of pancreatic ductal epithelium but not for differentiation of exocrine or endocrine cells. *Mech Dev* *120*, 167-175.
- Raum, J. C., Gerrish, K., Artner, I., Henderson, E., Guo, M., Sussel, L., Schisler, J. C., Newgard, C. B., and Stein, R. (2006). FoxA2, Nkx2.2, and PDX-1 regulate islet beta-cell-specific mafA expression through conserved sequences located between base pairs -8118 and -7750 upstream from the transcription start site. *Mol Cell Biol* *26*, 5735-5743.
- Rausa, F., Samadani, U., Ye, H., Lim, L., Fletcher, C. F., Jenkins, N. A., Copeland, N. G., and Costa, R. H. (1997). The cut-homeodomain transcriptional activator HNF-6 is coexpressed with its target gene HNF-3 beta in the developing murine liver and pancreas. *Dev Biol* *192*, 228-246.
- Reith, W., Herrero-Sanchez, C., Kobr, M., Silacci, P., Berte, C., Barras, E., Fey, S., and Mach, B. (1990). MHC class II regulatory factor RFX has a novel DNA-binding domain and a functionally independent dimerization domain. *Genes Dev* *4*, 1528-1540.
- Reith, W., and Mach, B. (2001). The bare lymphocyte syndrome and the regulation of MHC expression. *Annu Rev Immunol* *19*, 331-373.
- Reith, W., Satola, S., Sanchez, C. H., Amaldi, I., Lisowska-Groszpiere, B., Griscelli, C., Hadam, M. R., and Mach, B. (1988). Congenital immunodeficiency with a regulatory defect in MHC class II gene expression lacks a specific HLA-DR promoter binding protein, RF-X. *Cell* *53*, 897-906.
- Reith, W., Ucla, C., Barras, E., Gaud, A., Durand, B., Herrero-Sanchez, C., Kobr, M., and Mach, B. (1994). RFX1, a transactivator of hepatitis B virus enhancer I, belongs to a novel family of homodimeric and heterodimeric DNA-binding proteins. *Mol Cell Biol* *14*, 1230-1244.
- Rhinn, M., Dierich, A., Shawlot, W., Behringer, R. R., Le Meur, M., and Ang, S. L. (1998). Sequential roles for Otx2 in visceral endoderm and neuroectoderm for forebrain and midbrain induction and specification. *Development* *125*, 845-856.
- Ross, S. A., McCaffery, P. J., Drager, U. C., and De Luca, L. M. (2000). Retinoids in embryonal development. *Physiol Rev* *80*, 1021-1054.
- Rudnick, A., Ling, T. Y., Odagiri, H., Rutter, W. J., and German, M. S. (1994). Pancreatic beta cells express a diverse set of homeobox genes. *Proc Natl Acad Sci U S A* *91*, 12203-12207.

- Rulifson, I. C., Karnik, S. K., Heiser, P. W., ten Berge, D., Chen, H., Gu, X., Taketo, M. M., Nusse, R., Hebrok, M., and Kim, S. K. (2007). Wnt signaling regulates pancreatic beta cell proliferation. *Proc Natl Acad Sci U S A* *104*, 6247-6252.
- Samaras, S. E., Zhao, L., Means, A., Henderson, E., Matsuoka, T. A., and Stein, R. (2003). The islet beta cell-enriched RIPE3b1/Maf transcription factor regulates pdx-1 expression. *J Biol Chem* *278*, 12263-12270.
- Sander, M., Neubuser, A., Kalamaras, J., Ee, H. C., Martin, G. R., and German, M. S. (1997). Genetic analysis reveals that PAX6 is required for normal transcription of pancreatic hormone genes and islet development. *Genes Dev* *11*, 1662-1673.
- Sander, M., Sussel, L., Connors, J., Scheel, D., Kalamaras, J., Dela Cruz, F., Schwitzgebel, V., Hayes-Jordan, A., and German, M. (2000). Homeobox gene Nkx6.1 lies downstream of Nkx2.2 in the major pathway of beta-cell formation in the pancreas. *Development* *127*, 5533-5540.
- Sanvito, F., Herrera, P. L., Huarte, J., Nichols, A., Montesano, R., Orci, L., and Vassalli, J. D. (1994). TGF-beta 1 influences the relative development of the exocrine and endocrine pancreas in vitro. *Development* *120*, 3451-3462.
- Schubert, M. L. (2003). Gastric secretion. *Curr Opin Gastroenterol* *19*, 519-525.
- Schweisguth, F. (2004). Regulation of notch signaling activity. *Curr Biol* *14*, R129-138.
- Schwitzgebel, V. M., Scheel, D. W., Connors, J. R., Kalamaras, J., Lee, J. E., Anderson, D. J., Sussel, L., Johnson, J. D., and German, M. S. (2000). Expression of neurogenin3 reveals an islet cell precursor population in the pancreas. *Development* *127*, 3533-3542.
- Sengupta, P. K., Fargo, J., and Smith, B. D. (2002). The RFX family interacts at the collagen (COL1A2) start site and represses transcription. *J Biol Chem* *277*, 24926-24937.
- Seymour, P. A., Freude, K. K., Dubois, C. L., Shih, H. P., Patel, N. A., and Sander, M. (2008). A dosage-dependent requirement for Sox9 in pancreatic endocrine cell formation. *Dev Biol* *323*, 19-30.
- Seymour, P. A., Freude, K. K., Tran, M. N., Mayes, E. E., Jensen, J., Kist, R., Scherer, G., and Sander, M. (2007). SOX9 is required for maintenance of the pancreatic progenitor cell pool. *Proc Natl Acad Sci U S A* *104*, 1865-1870.
- Siegrist, C. A., Durand, B., Emery, P., David, E., Hearing, P., Mach, B., and Reith, W. (1993). RFX1 is identical to enhancer factor C and functions as a transactivator of the hepatitis B virus enhancer. *Mol Cell Biol* *13*, 6375-6384.
- Sim, K. G., Cheong, J. K., and Hsu, S. I. (2006). The TRIP-Br family of transcriptional regulators is essential for the execution of cyclin E-mediated cell cycle progression. *Cell Cycle* *5*, 1111-1115.

- Simon-Assmann, P., and Kedinger, M. (1993). Heterotypic cellular cooperation in gut morphogenesis and differentiation. *Semin Cell Biol* 4, 221-230.
- Skipper, M., and Lewis, J. (2000). Getting to the guts of enteroendocrine differentiation. *Nat Genet*, 24, 3-4
- Slack, J. M. (1995). Developmental biology of the pancreas. *Development* 121, 1569-1580.
- Soll, A. H., Yamada, T., Park, J., and Thomas, L. P. (1984). Release of somatostatinlike immunoreactivity from canine fundic mucosal cells in primary culture. *Am J Physiol* 247, G558-566.
- 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.
- St-Onge, L., Sosa-Pineda, B., Chowdhury, K., Mansouri, A., and Gruss, P. (1997). Pax6 is required for differentiation of glucagon-producing alpha-cells in mouse pancreas. *Nature* 387, 406-409.
- Stafford, D., and Prince, V. E. (2002). Retinoic acid signaling is required for a critical early step in zebrafish pancreatic development. *Curr Biol* 12, 1215-1220.
- Stafford, D., White, R. J., Kinkel, M. D., Linville, A., Schilling, T. F., and Prince, V. E. (2006). Retinoids signal directly to zebrafish endoderm to specify insulin-expressing beta-cells. *Development* 133, 949-956.
- Stanger, B. Z., and Dor, Y. (2006). Dissecting the cellular origins of pancreatic cancer. *Cell Cycle* 5, 43-46.
- Strowski, M. Z., Kohler, M., Chen, H. Y., Trumbauer, M. E., Li, Z., Szalkowski, D., Gopal-Truter, S., Fisher, J. K., Schaeffer, J. M., Blake, A. D., Zhang, B. B., and Wilkinson, H. A. (2003). Somatostatin receptor subtype 5 regulates insulin secretion and glucose homeostasis. *Mol Endocrinol* 17, 93-106.
- Sugimoto, M., Nakamura, T., Ohtani, N., Hampson, L., Hampson, I. N., Shimamoto, A., Furuichi, Y., Okumura, K., Niwa, S., Taya, Y., and Hara, E. (1999). Regulation of CDK4 activity by a novel CDK4-binding protein, p34(SEI-1). *Genes Dev* 13, 3027-3033.
- Sussel, L., Kalamaras, J., Hartigan-O'Connor, D. J., Meneses, J. J., Pedersen, R. A., Rubenstein, J. L., and German, M. S. (1998). Mice lacking the homeodomain transcription factor Nkx2.2 have diabetes due to arrested differentiation of pancreatic beta cells. *Development* 125, 2213-2221.
- Swoboda, P., Adler, H. T., and Thomas, J. H. (2000). The RFX-type transcription factor DAF-19 regulates sensory neuron cilium formation in *C. elegans*. *Mol Cell* 5, 411-421.

- Thomas, M. K., Devon, O. N., Lee, J. H., Peter, A., Schlosser, D. A., Tenser, M. S., and Habener, J. F. (2001). Development of diabetes mellitus in aging transgenic mice following suppression of pancreatic homeoprotein IDX-1. *J Clin Invest* *108*, 319-329.
- Thomas, M. K., Rastalsky, N., Lee, J. H., and Habener, J. F. (2000). Hedgehog signaling regulation of insulin production by pancreatic beta-cells. *Diabetes* *49*, 2039-2047.
- Thomas, P., and Beddington, R. (1996). Anterior primitive endoderm may be responsible for patterning the anterior neural plate in the mouse embryo. *Curr Biol* *6*, 1487-1496.
- von Bubnoff, A., and Cho, K. W. (2001). Intracellular BMP signaling regulation in vertebrates: pathway or network? *Dev Biol* *239*, 1-14.
- Wang, G. J., Gao, C. F., Wei, D., Wang, C., and Ding, S. Q. (2009a). Acute pancreatitis: etiology and common pathogenesis. *World J Gastroenterol* *15*, 1427-1430.
- Wang, Q., Elghazi, L., Martin, S., Martins, I., Srinivasan, R. S., Geng, X., Sleeman, M., Collombat, P., Houghton, J., and Sosa-Pineda, B. (2008). Ghrelin is a novel target of Pax4 in endocrine progenitors of the pancreas and duodenum. *Dev Dyn* *237*, 51-61.
- Wang, S., Jensen, J. N., Seymour, P. A., Hsu, W., Dor, Y., Sander, M., Magnuson, M. A., Serup, P., and Gu, G. (2009b). Sustained Neurog3 expression in hormone-expressing islet cells is required for endocrine maturation and function. *Proc Natl Acad Sci U S A* *106*, 9715-9720.
- Watanabe-Fukunaga, R., Iida, S., Shimizu, Y., Nagata, S., and Fukunaga, R. (2005). SEI family of nuclear factors regulates p53-dependent transcriptional activation. *Genes Cells* *10*, 851-860.
- Weinstein, D. C., Ruiz i Altaba, A., Chen, W. S., Hoodless, P., Prezioso, V. R., Jessell, T. M., and Darnell, J. E., Jr. (1994). The winged-helix transcription factor HNF-3 beta is required for notochord development in the mouse embryo. *Cell* *78*, 575-588.
- Wells, J. M., Esni, F., Boivin, G. P., Aronow, B. J., Stuart, W., Combs, C., Sklenka, A., Leach, S. D., and Lowy, A. M. (2007). Wnt/beta-catenin signaling is required for development of the exocrine pancreas. *BMC Dev Biol* *7*, 4.
- Wells, J. M., and Melton, D. A. (1999). Vertebrate endoderm development. *Annu Rev Cell Dev Biol* *15*, 393-410.
- Wells, J. M., and Melton, D. A. (2000). Early mouse endoderm is patterned by soluble factors from adjacent germ layers. *Development* *127*, 1563-1572.
- White, P., May, C. L., Lamounier, R. N., Brestelli, J. E., and Kaestner, K. H. (2008). Defining pancreatic endocrine precursors and their descendants. *Diabetes* *57*, 654-668.

- White, S. A., Shaw, J. A., and Sutherland, D. E. (2009). Pancreas transplantation. *Lancet* *373*, 1808-1817.
- Wierup, N., Svensson, H., Mulder, H., and Sundler, F. (2002). The ghrelin cell: a novel developmentally regulated islet cell in the human pancreas. *Regul Pept* *107*, 63-69.
- Wild, S., Roglic, G., Green, A., Sicree, R., and King, H. (2004). Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care* *27*, 1047-1053.
- Wolfe, S. A., van Wert, J., and Grimes, S. R. (2006). Transcription factor RFX2 is abundant in rat testis and enriched in nuclei of primary spermatocytes where it appears to be required for transcription of the testis-specific histone H1t gene. *J Cell Biochem* *99*, 735-746.
- Wolfe, S. A., Vanwert, J. M., and Grimes, S. R. (2008). Transcription factor RFX4 binding to the testis-specific histone H1t promoter in spermatocytes may be important for regulation of H1t gene transcription during spermatogenesis. *J Cell Biochem* *105*, 61-69.
- Wu, K. L., Gannon, M., Peshavaria, M., Offield, M. F., Henderson, E., Ray, M., Marks, A., Gamer, L. W., Wright, C. V., and Stein, R. (1997). Hepatocyte nuclear factor 3beta is involved in pancreatic beta-cell-specific transcription of the pdx-1 gene. *Mol Cell Biol* *17*, 6002-6013.
- Xi, Y., Formentini, A., Nakajima, G., Kornmann, M., and Ju, J. (2008). Validation of biomarkers associated with 5-fluorouracil and thymidylate synthase in colorectal cancer. *Oncol Rep* *19*, 257-262.
- Xi, Y., Nakajima, G., Schmitz, J. C., Chu, E., and Ju, J. (2006). Multi-level gene expression profiles affected by thymidylate synthase and 5-fluorouracil in colon cancer. *BMC Genomics* *7*, 68.
- Xu, X., D'Hoker, J., Stange, G., Bonne, S., De Leu, N., Xiao, X., Van de Casteele, M., Mellitzer, G., Ling, Z., Pipeleers, D., Bouwens, L., Scharfmann, R., Gradwohl, G., and Heimberg, H. (2008). Beta cells can be generated from endogenous progenitors in injured adult mouse pancreas. *Cell* *132*, 197-207.
- Xu, Y., Sengupta, P. K., Seto, E., and Smith, B. D. (2006). Regulatory factor for X-box family proteins differentially interact with histone deacetylases to repress collagen alpha2(I) gene (COL1A2) expression. *J Biol Chem* *281*, 9260-9270.
- Yee, N. S., Lorent, K., and Pack, M. (2005). Exocrine pancreas development in zebrafish. *Dev Biol* *284*, 84-101.
- Zarbalis, K., May, S. R., Shen, Y., Ekker, M., Rubenstein, J. L., and Peterson, A. S. (2004). A focused and efficient genetic screening strategy in the mouse: identification of mutations that disrupt cortical development. *PLoS Biol* *2*, E219.

Zhang, C., Moriguchi, T., Kajihara, M., Esaki, R., Harada, A., Shimohata, H., Oishi, H., Hamada, M., Morito, N., Hasegawa, K., Kudo, T., Engel, J. D., Yamamoto, M., and Takahashi, S. (2005). MafA is a key regulator of glucose-stimulated insulin secretion. *Mol Cell Biol* 25, 4969-4976.

Zhang, D., Zeldin, D. C., and Blackshear, P. J. (2007). Regulatory factor X4 variant 3: a transcription factor involved in brain development and disease. *J Neurosci Res* 85, 3515-3522.

Zhang, T., Liu, W. D., Saunee, N. A., Breslin, M. B., and Lan, M. S. (2009). Zinc finger transcription factor INSM1 interrupts cyclin D1 and CDK4 binding and induces cell cycle arrest. *J Biol Chem* 284, 5574-5581.

ABREVIATIONS

AA	Acides Aminés	DMEM	Dulbecco's Modified Eagle Medium
ADN	Acide DésoxyriboNucléique		
ADNc	Acide DésoxyriboNucléique complémentaire	DMSO	DiMéthyl SulfOxyde
AR	Acide Rétinoïque	DTT	DiThioThréitol
ARN	Acide RiboNucléique	E	Stade de développement Embryonnaire
ARNm	Acide RiboNucléique messenger	EGFP	Enhanced Green Fluorescent Protein
ARNt	Acide RiboNucléique de transfert	ERTM	Estrogen Receptor ^{TaMoxifen}
Arx	Aristaless related homeobox	EST	Expressed Sequence Tag
ATP	Adénosine TriPhosphate	eYFP	Enhanced Yellow Fluorescent Protein
bHLH	basic Helix-Loop-Helix	FACS	Fluorescence-Activated Cell Sorting
BSA	Albumine Sérique de Bœuf	FGF	Fibroblast Growth Factor
Cdca4	Cell division cycle associated 4	Foxa2	Forkhead box A2 (=HNF3β)
Cdk	Cyclin Dependant Kinase	Frz	Frizzled
Cdx2	Caudal type homeo box 2	GAPDH	GlycérAldehyde-3-Phosphate DésHydrogenase
cf	<i>Confer</i>	Gcg	Glucagon
CMH	complexe majeur d'histocompatibilité	GTP	Guanosine TriPhosphate
CYC	Cyclopamine	h	heure
CTP	Cytidine TriPhosphate	Hes1	Hairy and enhancer of split 1
DAPT	N-[N-(3,5-difluorophenacetyl)-1-alanyl]-S-phenylglycine t-butyl ester	Hesx1	homeo box gene expressed in ES cells 1
DEAE	Diethylamino Ethanol	Hlxb9	Homeobox gene HB9
DEPC	Diethyl Pyrocarbonate	HMG	High Mobility Group
Dll1	Delta-like 1	HNF	Hepatic Nuclear Factor
		hpf	hours post fertilization
		HSP90	Heat-Shock Protein 90

I-FABP	Intestinal Fatty Acid Binding Protein	PCR	Polymerase Chain Reaction
Insm1	Insulinoma-associated 1	Pdx1	Pancreas duodenum homeobox 1
IRES	Internal Ribosomal Entry Site	PFA	Paraformaldéhyde
ISH	<i>In Situ</i> Hybridisation	PHD	Plant HomeoDomain
Isl1	ISL1 transcription factor, LIM/homeodomain	PolyA	PolyAdenylation
jpc	jour post coitum	PP	Polypeptide Pancréatique
kb	kilo base	Ptc	Patched
LIM	LIn-11 and Mec-3	Ptf1a	Pancreas specific Transcription Factor, 1a
LRP	Low density lipoprotein-Related Protein	RACE	Rapid Amplification of cDNA Ends
MABT	Maleic Acid Buffer Triton	Raldh2	aldehyde dehydrogenase family 1, subfamily A2
min	minute	RBPJκ	Recombination signal Binding Protein for immunoglobulin kappa J region
MODY	Maturity Onset Diabetes of the Young	RFX6	Regulatory Factor X 6
NCBI	National Center for Biotechnology Information	Rfxdc1	Regulatory Factor X Domain Containing 1
NeuroD1	Neurogenic differentiation 1	Riken L19	RIKEN cDNA 1700086L19
Neurog3	Neurogenin3	RPMI	Roswell Park Memorial Institute
Nkx	NK homeobox	S	Somite
Oc-1	one cut domain, family member 1 (=HNF6)	SDS	Dodécyl Sulfate de Sodium
Otx2	Orthodenticle homolog 2	sec	seconde
Pax 4	Paired box gene 4	SERTA	SEI-1, RBT1 and TARA motif
Pax 6	Paired box gene 6	Sertad	SERTA domain containing
pb	paire de bases	SFRP	Secreted Frizzled-Related Protein
PBS	Phosphate Buffer Salin	Shh	Sonic HedgeHog
PBST	Phosphate Buffer Salin Triton		

SSC	Sodium Salt Citrate	Tris	Tris-hydroxyméthyl-aminométhane
Tam	4-Hydroxytamoxifen		
Taq	Thermus Aquaticus	TSA	Tyramide System Amplification
Tara	Taranis		
TCF1	HNF1 homeobox B	UTP	Uridine riPhosphate
TEMED	N,N,N',N'-Tetramethyl-1-,2-diaminomethane	UTP-dig	Uridine TriPhosphate couplée à la digoxygénine
TGFβ	Transforming Growth Factor, beta	UTR	UnTranslated Region
		YFP	Yellow Fluorescent Protein

ANNEXE

Development 137, 203-212 (2010) doi:10.1242/dev.041673

Rfx6 is an Ngn3-dependent winged helix transcription factor required for pancreatic islet cell development

Josselin Soyer^{1,*}, Lydie Flasse^{2,*}, Wolfgang Raffelsberger¹, Anthony Beucher¹, Christophe Orvain¹, Bernard Peers², Philippe Ravassard³, Julien Vermot¹, Marianne L. Voz^{2,†}, Georg Mellitzer^{1,†} and Gérard Gradwohl^{1,†}

SUMMARY

The transcription factor neurogenin 3 (Neurog3 or Ngn3) controls islet cell fate specification in multipotent pancreatic progenitor cells in the mouse embryo. However, our knowledge of the genetic programs implemented by *Ngn3*, which control generic and islet subtype-specific properties, is still fragmentary. Gene expression profiling in isolated Ngn3-positive progenitor cells resulted in the identification of the uncharacterized winged helix transcription factor Rfx6. Rfx6 is initially expressed broadly in the gut endoderm, notably in Pdx1-positive cells in the developing pancreatic buds, and then becomes progressively restricted to the endocrine lineage, suggesting a dual function in both endoderm development and islet cell differentiation. Rfx6 is found in postmitotic islet progenitor cells in the embryo and is maintained in all developing and adult islet cell types. *Rfx6* is dependent on Ngn3 and acts upstream of or in parallel with *NeuroD*, *Pax4* and *Arx* transcription factors during islet cell differentiation. In zebrafish, the *Rfx6* ortholog is similarly found in progenitors and hormone expressing cells of the islet lineage. Loss-of-function studies in zebrafish revealed that *rfx6* is required for the differentiation of *glucagon*-, *ghrelin*- and *somatostatin*-expressing cells, which, in the absence of *rfx6*, are blocked at the progenitor stage. By contrast, beta cells, whose number is only slightly reduced, were no longer clustered in a compact islet. These data unveil Rfx6 as a novel regulator of islet cell development.

KEY WORDS: Neurogenin 3, Pancreas, Rfx, Cell differentiation, Endocrine, Transcription factor, Mouse, Zebrafish

INTRODUCTION

Deciphering the mechanisms controlling the progressive restriction of the fate of stem/progenitor cells and their differentiation into highly specialized cells is not only a major issue in stem cell biology but will also have an important impact on future cell-based or regenerative therapies in major disease such as type 1 diabetes, where insulin-producing beta cells are destroyed. In 2000, the Edmonton protocol of cadaveric islet transplantation from allogeneic donors reported successful restoration of insulin production and glycemic stability in patients with type 1 diabetes mellitus (Shapiro et al., 2000). These studies provided the proof-of-principle for a cell-based therapy in diabetes and launched a new area of islet cell transplantation. However, major limitations have still to be overcome such as the reoccurrence of the autoimmune destruction of the beta cells and the scarcity of transplantable islets due to the paucity of donors. In the last decade, significant knowledge has been acquired on the transcriptional regulation and signals controlling beta cell development during mouse embryogenesis (Murtaugh, 2007; Claiborn and Stoffers, 2008). Thanks to these findings, major progress has been achieved in the generation of insulin-producing cells from human embryonic stem cells (hESCs) by recapitulating embryonic differentiation programs (D'Amour et al., 2006; Madsen and Serup, 2006). However, the

cells generated are still immature and remain different from normal glucose-responsive single-hormone-positive human islet beta cells (Kroon et al., 2008). This limitation underlines the crucial importance to pursue basic research to gain a highly detailed knowledge of the developmental program leading to functional beta cells.

During mouse pancreas embryogenesis, the basic helix-loop-helix (bHLH) transcription factor neurogenin 3 (Neurog3 or Ngn3) is the master gene controlling endocrine cell fate decisions in uncommitted multipotent pancreatic endodermal progenitor cells. Ngn3 is transiently expressed in endocrine progenitor cells which do not yet express hormones (Gradwohl et al., 2000; Schwitzgebel et al., 2000). In the absence of Ngn3, all pancreatic endocrine cells, including alpha-, beta-, delta-, PP- and epsilon-cells, which produce glucagon, insulin, somatostatin, pancreatic polypeptide and ghrelin hormones, respectively, fail to develop (Gradwohl et al., 2000; Heller et al., 2005). Consequently, islets of Langerhans do not form and mice die from diabetes. Importantly, ectopic Ngn3 expression is also sufficient to generate all islet cell types in mouse either in vivo (Johansson et al., 2007) or ex vivo in pancreatic explant cultures (M. Martin and G.G., unpublished). Ngn3 thus controls a complex network of transcription factors, leading to mature islet cells. In agreement with these findings, lineage tracing studies demonstrated that all pancreatic endocrine cells derive from Ngn3-positive progenitor cells (Gu et al., 2002; Schonhoff et al., 2004; Heller et al., 2005). As in the pancreas, the differentiation of endocrine cells of the gastrointestinal tract relies on Ngn3 (Jenny et al., 2002; Lee et al., 2002). Important downstream target genes of Ngn3 in the pancreas include *Arx* and *Pax4*, the major regulators of the alpha and beta cell fate (Sosa-Pineda et al., 1997; Collombat et al., 2003). Although some evidence in the literature suggests that transcription factors such as *NeuroD* (*Neurod1* – Mouse Genome Informatics) (Huang et al., 2000), *Pax4* (Smith et al., 2003), *Insm1*

¹Institute of Genetics and Molecular and Cell Biology (IGBMC), Inserm U-964, CNRS UMR7104, University of Strasbourg, Illkirch, France. ²GIGA-R, University of Liege, Avenue Hospital, 1, 4000 Sart-Tilman, Belgium. ³Centre de Recherche de l'Institut du Cerveau et de la Moelle CNRS UMR7225, Inserm UMRS795, Université Pierre et Marie Curie, Paris, France.

*These authors contributed equally to this work

†Authors for correspondence (mellit@igbmc.fr; gradwohl@igbmc.fr; mvoz@ulg.ac.be)

(Mellitzer et al., 2006) or *Nkx2.2* (Watada et al., 2003) might be direct targets of Ngn3, it has not yet been proven that Ngn3 binds the promoter of these genes in vivo. Our knowledge of the Ngn3-regulated program is thus still incomplete, and we do not know much about how this transcription factor integrates the generic program of endocrine differentiation with the programs that specify the different islet cell types. Therefore, to determine the gene expression profile of islet progenitor cells and thus identify potential novel downstream effectors of Ngn3 function, we performed Affymetrix microarray analysis on purified Ngn3-positive progenitor cells. Here we show that this strategy led to the identification of Rfx6, a novel Ngn3-dependent winged helix transcription factor. We report that in both mouse and zebrafish, *Rfx6* is found in islet progenitor cells and maintained in developing and mature islet cells. Furthermore functional studies in zebrafish revealed that *rxf6* is an important regulator of endocrine cell differentiation. These findings suggest that the study of other genes identified in this study might reveal the full genetic program implementing Ngn3 endocrinogenic function. This information might in turn be relevant to promote and scrutinize the differentiation of hESCs to the beta lineage.

MATERIALS AND METHODS

Preparation of single-cell suspensions and RNA, probe synthesis and microarray hybridization and analysis

Single-cell suspensions were prepared from E15.5 *Ngn3^{eYFP/+}* pancreas as described previously (Mellitzer et al., 2004). eYFP-positive and -negative cells were sorted directly into Trizol reagent (Invitrogen) and immediately processed for total RNA isolation. On average, 1500–2000 eYFP-positive cells were obtained per embryonic pancreas. One hundred nanograms of RNA from each biological quadruplicate was then used for linear amplification (T7, RiboAmp OA Kit, Arcturus). cRNA probes were generated with the Enzo Bioarray High Yield RNA Transcription Labelling Kit and hybridized on the GeneChip Mouse Genome 430 2.0 Array following Affymetrix standard protocols. Affymetrix raw gene expression data were normalized using the GC Robust Multi-array Average (GCRMA) procedure. The data were filtered in order to remove probe sets with constant low-level expression. The filtered data sets were subsequently subjected to *t*-tests with multiple testing correction and control of the global and local false discovery rate (FDR and *fdr*, respectively) using the OCplus package (Ploner et al., 2006) available from Bioconductor (<http://bioconductor.org>). Only transcripts with at least a 2-fold change in expression with an FDR below 0.05 were retained. Using these criteria, 1445 genes (1831 probe sets) were found up- (550, FC 2–386) or down- (895, FC 2–143) regulated in islet progenitor cells. Transcription factors and transcriptional regulators were identified using Gene Ontology annotations. Microarray data are available at <http://genomics.betacell.org>, on the RAD database https://www.cbil.upenn.edu/RADQuerier/php/displayStudy.php?study_id=3100 and ArrayExpress (accession number E-CBIL-48).

In situ hybridization and immunohistochemistry on mouse tissues

Tissues were fixed in paraformaldehyde, embedded and frozen using standard methods. Detailed protocols for in situ hybridization immunofluorescence and immunohistochemistry on frozen sections are available on request. Mouse cRNA probes used included: *Ngn3* and *NeuroD* (Gradwohl et al., 2000), *Insm1/IAI* (Mellitzer et al., 2006), *Rfx3* (transcribed from a 2.7 kb cDNA fragment, image clone 4483833, IRVAV35-E2) and *Rfx6* (transcribed from a 0.9 kb cDNA fragment, cloned from E13.5 pancreatic RNA with oligo 5' CGGAATTCGCCACGTGGAGACATCCTAT and 3' GGACTAGTAATCTGGGTTTGAAGTTGG). The following primary antibodies were used: guinea pig or rabbit anti-Pdx1 at 1:1000 (provided by C. Wright, Vanderbilt University, Nashville, TN, USA), guinea pig anti-Ngn3 at 1:1000 (provided by M. Sander, California University, Irvine, CA, USA), anti-insulin at 1:1000 (Linco), anti-glucagon at 1:2000 (Linco), anti-peptide pancreatic (PP) at 1:1000 (Linco), rabbit anti-somatostatin at 1:200 (Dako), rabbit anti-Rfx6 at 1:1000, rat anti-Rfx6 at 1:200, mouse anti-

Ki-67 at 1:100 (Novocastra). Secondary antibodies used were: anti-rabbit Alexa 488 at 1:1000 (Molecular Probes), Cy3 anti-rabbit, anti-rat and anti-guinea pig at 1:1000 (Jackson Immunoresearch), biotin-coupled anti-rabbit at 1:200 (Vector Laboratories). For rat anti-Rfx6, signal amplification was performed using biotin anti-rat coupled antibody at 1:200 (Vector Laboratories) and streptavidin-Alexa 488 conjugate at 1:500 (Molecular Probes). Nuclei were stained with DAPI and slides mounted in Aqua-Poly/Mount (Polysciences).

Cloning of the *rxf6* ortholog from zebrafish

rxf6 partial cDNA was cloned by two rounds of PCR performed on cDNAs of embryos at 24 hours post-fertilization (hpf). The primers used for amplification were O146 (TGCCCTTTTGGACCAGATTGTAGTG) and O139 (GAACGACTGGAGCTGCTGATGGAT) for the first PCR, followed by a nested PCR with O146 and O147 (GCTACGCTTCTCTGGAC-ATCACCT), giving rise to a 972 bp fragment in the coding region. This fragment was cloned into a pCRII-TOPO vector (Invitrogen) and used as template for preparing labelled antisense RNA probes.

Morpholino sequences and injections

The *rxf6* morpholinos (MOs) were designed by Gene Tools and are complementary to either the exon 2 splice donor site (MO1: GTCCTCAAGCCTAATGAAACAAAAC) or the exon 2 splice acceptor site (MO2: AATAAAAACGCCTCTTACCTTTCCG). A standard control MO, having the sequence 5'-CCTCTTACCTCAGTTACAATTATA-3', has also been designed by Gene Tools in a way that it should have no target and no significant biological activity. The MOs were dissolved at a concentration of 3 µg/µl in 1× Danieau buffer containing 0.5% rhodamine dextran and microinjected at the 1- to 2-cell stage at a dose of 3 ng. Injected embryos were then grown in the presence of 0.003% 1-phenyl-2-thiourea until the desired stage, fixed overnight in 4% paraformaldehyde and stored in 100% methanol before analysis.

Riboprobes, wholemount in situ hybridizations (WISH) and imaging

Antisense riboprobes were made by transcribing linearized cDNA clones with SP6, T7 or T3 polymerase using digoxigenin or DNP labelling mix (Roche) according to manufacturer's instructions. They were subsequently purified on NucAway spin columns (Ambion) and ethanol-precipitated. The zebrafish *sox4b* (Mavropoulos et al., 2005), *isl1* (Korz et al., 1993), *neurod* (Korz et al., 1998), *insulin* (Milewski et al., 1998), *somatostatin 2* (Devos et al., 2002), *ghrelin* (NCBI: AL918922) and *glucagon* (Argenton et al., 1999) probes have been described elsewhere. Single-wholemount, double-fluorescent in situ hybridizations and fluorescent imaging were carried out as described (Mavropoulos et al., 2005).

Cilia imaging

Immunostaining was performed on 24 hpf embryos where primary cilia were labelled with anti-acetylated tubulin (Sigma, T6793) and GFP-expressing cells were labelled with anti-GFP (Millipore, AB3080). Cilia imaging was performed using the Leica sp2 confocal microscope to image tg(*pax6:GFP*) and acetylated-tubulin-labelled embryos in order to visualize cilia concomitantly with pancreatic cells. Embryos were dissected prior to imaging and mounted on a slide. Images were taken using a 63X/1.2 HCX PL APOchromat water immersion lens. Stacks were reconstructed and processed using Imap (Bitplane).

RESULTS

Differential expression of transcription factors in the islet lineage

To identify the complete panel of genes activated specifically in islet progenitor cells, we determined the genes differentially expressed in Ngn3-positive versus Ngn3-negative cells at E15.5, a stage when the proportion of Ngn3 cells culminates in the embryonic pancreas. Ngn3-positive cells were FACS-purified from *Ngn3^{eYFP/+}* mice, where the enhanced yellow fluorescent protein (eYFP) has been introduced in the 3' UTR region, leaving the coding sequence intact and thus wild-type levels of Ngn3 protein as supported by normal

islet cell development and glucose homeostasis in *Ngn3^{eYFP/eYFP}* mice (Mellitzer et al., 2004). As transcription factor levels are crucial to transactivate the appropriate targets, this mouse model is thus ideally designed to reveal the full *Ngn3*-dependent transcriptome. Importantly, due to the stability of the eYFP protein, both *Ngn3*-positive cells and their progeny are isolated using this strategy. We used Affymetrix GeneChip 430 2.0 Arrays containing 39,000 transcripts, a chip that has not previously been used to characterize the genes enriched in purified islet progenitor cells. The present report focuses on transcriptional regulators that have been identified in this study. We found 47 transcriptional regulators upregulated in sorted eYFP/*Ngn3*-positive cells (see Table S1 in the supplementary material). As expected, transcripts for genes encoding transcription factors known to be expressed downstream of *Ngn3* such as *NeuroD*, *Arx4*, *Pax4*, *Pax6*, *Insm1* and *Mafa* were strongly enriched in purified eYFP/*Ngn3*-cells. Importantly, several transcription factors for which a function in islet development has not yet been reported were also identified. These include *Fev*, *Mlxipl*, *Rfxdc1*, *Vdr*, *Dach1* and *Lhx1*. Below, we characterize the expression and function of *Rfxdc1*, one of these novel islet specific transcription factors, also called *Rfx6*, and enriched 65 times in eYFP/*Ngn3*-positive cells.

Rfx6 is a novel islet-specific winged helix transcription factor

Rfx6 is a member of the Rfx transcription factor family that bind X-boxes of DNA sequences with a DNA binding domain containing a winged helix motif. Among the Rfx family, so far only *Rfx3* has been reported to be expressed in the embryonic pancreas and crucial for islet cell development (Ait-Lounis et al., 2007). The *Rfx6* gene is located on position qB3 on mouse chromosome 10 and contains 19 exons in 52.62 kb. Two transcripts are predicted in Ensembl, ENSMUST0000002054 (19 exons, 3419 bps) and ENSMUST00000050455 (16 exons, 3,088 bps), of which only the first encodes a protein containing the predicted Rfx DNA binding domain (PF02257). Using RT-PCR strategies and northern blots (data not shown), only one transcript could be identified from embryonic pancreas RNA (E13.5, E15.5), corresponding to the longest predicted transcript and encoding a protein of 927 amino acids (UNiProtKb Q8C7R7) sharing 87% identity with the human ortholog. Multiple alignment of the DNA binding domains of Rfx proteins revealed that *Rfx6* is most similar to *Rfx4* (see Fig. S1 in the supplementary material).

Rfx6 is expressed in the gut endoderm and becomes progressively restricted to the islet lineage in the embryonic pancreas

To determine *Rfx6* expression during mouse embryogenesis and pancreas development, we performed a series of in situ hybridization and immunohistochemistry experiments. At E9.0 *Rfx6* transcripts were broadly found in the gut endoderm (Fig. 1A,B). To better determine *Rfx6* expression we generated polyclonal antibodies in rabbits and rats against a glutathione-S-transferase (GST) fusion protein containing the N-terminus amino-acid 2-65, a peptide sharing no homology with other members of the Rfx family. Using this tool, we found *Rfx6* expression in the nascent pancreatic buds of E9.5 in *Pdx1*⁺ pancreatic progenitor cells (Fig. 1C-E). At E10.5, the ubiquitous endodermal expression of *Rfx6* persisted caudally in the prospective intestinal epithelium from the duodenum to the colon and rostrally in the proximal stomach and more anterior gut endoderm (Fig. 1F,G; data not shown). At this stage, *Rfx6* was also present in the lung epithelium, another tissue of endodermal origin (data not shown).

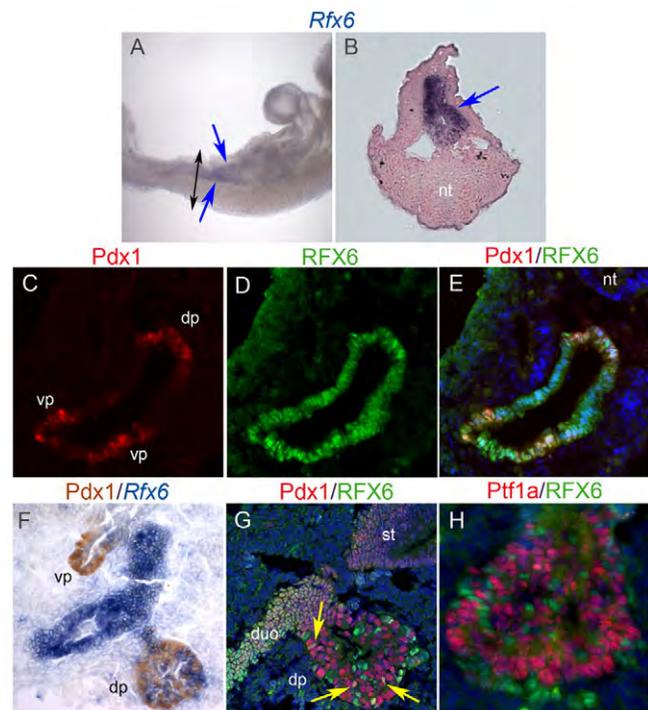


Fig. 1. Expression of *Rfx6* in the gut endoderm and in the developing pancreatic buds in the mouse embryo.

(A) Wholemount in situ hybridization (ISH) on an E9.0 mouse embryo (lateral view) with an *Rfx6* antisense probe (purple). (B) Transversal section of embryo shown in A at the level of the black arrows. (C-E, G, H) Double immunofluorescence for *Rfx6* (green) and *Pdx1* (red) on cryosections of E9.5 (C-E), and E10.5 (G) embryos and *Rfx6* and *Ptf1a* (red) at E10.5 (H). (F) ISH for *Rfx6* followed by immunohistochemistry for *Pdx1*. Nuclei are stained with DAPI (blue) in E, G and H. *Rfx6* is found in the gut endoderm (blue arrows in A,B) and in the developing pancreatic regions in *Pdx1*-positive progenitor cells at E9.5 (C-E). At E10.5, *Rfx6* is excluded from *Pdx1*- (F,G) and *Ptf1a*- (H) positive pancreatic endodermal progenitor cells and restricted to differentiating endocrine cells. All sections are sagittal, B is transverse. In E and G arrows point to *Pdx1/Rfx6* double-positive cells. dp, dorsal pancreas; duo, duodenum; nt, neural tube; st, stomach; vp, ventral pancreas.

By sharp contrast, after E9.5, *Rfx6* transcripts and *Rfx6* protein were progressively excluded from multipotent pancreatic endodermal progenitor cells. Indeed, at E10.5, *Rfx6* was found essentially in cells which do not express the pancreatic progenitor markers *Pdx1* or *Ptf1a* (Fig. 1F-H). At this stage, only rare *Rfx6*-high and *Pdx1*-low cells were found, which were likely to be either islet progenitors or early *Pdx1*-positive *insulin*-expressing cells. Instead, *Rfx6* marked *Ngn3*-positive cells (Fig. 2A-C) and alpha cells at E10.5 (Fig. 5A). These data suggest that in the embryonic pancreas from E10.5, *Rfx6* expression is restricted to developing islet cells. This hypothesis is fully supported by the almost complete absence of *Rfx6* transcripts at E10.5 (data not shown) and total ablation at E15.5 (Fig. 2G,H) in the *Ngn3*-deficient pancreas, which lack islet cells.

Rfx6 expression in developing islet cells is independent of *NeuroD*, *Pax4* and *Arx4*

As expected from the strong enrichment of *Rfx6* in *Ngn3*/eYFP-positive cells, 47% ($n=633$) of *Ngn3*-positive cells co-expressed *Rfx6* at E15.5 (see Fig. 2D-F). Accordingly, the *Rfx6* expression

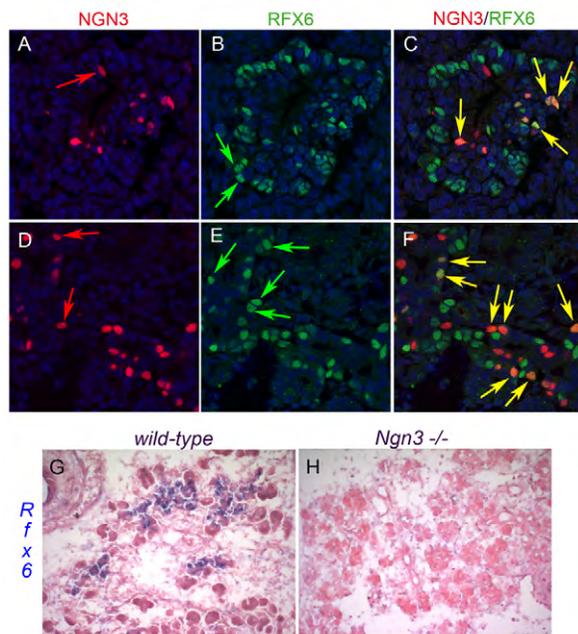


Fig. 2. *Ngn3*-dependent expression of *Rfx6* in the mouse islet lineage. (A-F) Double immunofluorescence showing partial overlapping expression of *Ngn3* (red) and *Rfx6* (green) at E10.5 (A-C) and E15.5 (D-F). Yellow, green and red arrows point to double-labelled, single *Rfx6*-positive and single *Ngn3*-positive cells, respectively. Nuclei are stained with DAPI. (G,H) Loss of *Rfx6* expression (in situ hybridization in blue) in an *Ngn3*-deficient pancreas (H) compared with a wild-type pancreas (G), demonstrating that *Rfx6* is restricted to the islet lineage at E15.5.

pattern largely overlapped with *NeuroD* and *Insm1*, known direct targets of *Ngn3* (Fig. 3A-D), and *Rfx6*-positive cells were essentially postmitotic (see Fig. S2 in the supplementary material). *Rfx6* also overlapped with *Rfx3*, whereas the other members of the *Rfx* family were not found to be expressed in the embryonic pancreas (Fig. 3E,F; data not shown). We next determined the position of *Rfx6* in the hierarchy of transcription factors controlling islet development and islet subtype specification (Fig. 4). At E14.5, *Rfx6* expression was unaffected in the pancreas of *Arx*- and *Pax4*-deficient mice, two key transcription factors regulating the determination of alpha- and beta/delta-cell fate, respectively (Sosa-Pineda et al., 1997; Collombat et al., 2003). Similarly, we could not detect any obvious difference in *Rfx6* expression when *NeuroD* is inactivated (Naya et al., 1997). These data suggest that *Rfx6* acts downstream of *Ngn3* and either upstream of *Arx*, *Pax4* and *NeuroD* or in independent pathways. When examined at E15.5, about 70% of *Rfx6*-positive cells ($n=1469$) were *Ngn3*-negative (Fig. 2D-F), suggesting that *Rfx6* expression is initiated in committed islet progenitor cells and maintained in developing post-*Ngn3* islet cells. In agreement with this hypothesis, glucagon- and insulin-positive cells expressed *Rfx6* in the embryo (Fig. 5A-C). Furthermore, *Rfx6* was maintained in adult islet cells including alpha, beta, delta and PP cells (Fig. 5D-G). Accordingly, *Rfx6* is also expressed in the beta cell lines bTC3 and Min6b1 cells (data not shown). In summary, in the mouse pancreas, *Rfx6* is expressed in islet progenitor cells in the embryo as well as in differentiated adult islet cells.

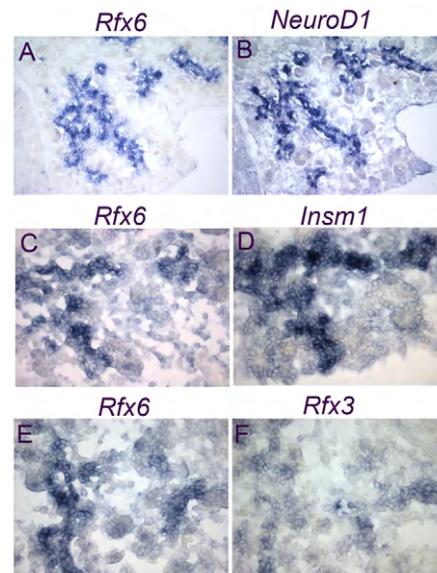


Fig. 3. Overlapping expression of *Rfx6* and other islet transcription factors in the pancreas of mouse embryos. (A-F) In situ hybridization (blue) experiments on adjacent cryosections showing overlapping expression of *Rfx6* with *NeuroD*, *Insm1* and *Rfx3*. All experiments were performed on E15.5 pancreata. Magnifications: 20 \times in A,B; 40 \times in C-F.

Zebrafish *rfx6* is expressed in the pancreatic endocrine progenitors as well as in the mature endocrine cells

To unravel the role of *Rfx6* in pancreas organogenesis, we took advantage of the zebrafish model, where the function of a gene can be easily tackled by the use of morpholinos (MOs) disrupting mRNA splicing or translation (Ekker, 2000). The *Rfx6* ortholog in zebrafish has been identified in Ensembl as ENSDARP00000061121. The predicted zebrafish *Rfx6* protein displays 60% sequence identity with the mouse *Rfx6*. The *Rfx* phylogenetic tree and the conserved synteny between human and zebrafish *rfx6* genomic loci show unequivocally that we have identified the *Rfx6* ortholog in zebrafish (see Fig. S1 in the supplementary material; data not shown). The expression pattern, determined by wholemount in situ hybridization, reveals that the zebrafish *rfx6* gene is expressed in the pancreatic region, as in mice. Its expression starts by 17 hpf, peaks at 24 hpf and persists at least until 72 hpf (Fig. 6A,B; see Fig. S3 in the supplementary material). However, in contrast to the mouse, *rfx6* was exclusively expressed in the pancreatic region, with no expression being detected in the gut. We next determined in which cell types *rfx6* is expressed by performing double-fluorescent in situ hybridization using various endocrine pancreatic markers. As the putative ortholog of *Ngn3* in zebrafish has been described as being expressed only at late stages in the pancreatic region (around 3 days post-fertilization) (Zecchin et al., 2007), we could not use *ngn3* as a marker of endocrine progenitors. We used instead *sox4b*, previously described as expressed predominantly in the precursors

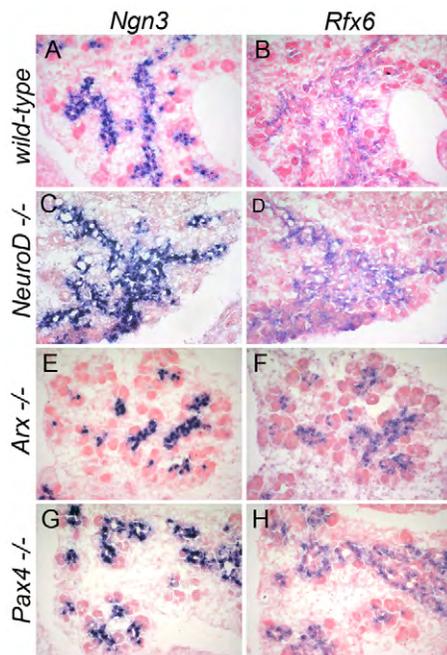


Fig. 4. *Rfx6* expression is unaffected in the pancreas of *Pax4*-, *Arx*- and *NeuroD*-deficient mouse embryos. (A-H) In situ hybridization (blue) for *Rfx6* and *Ngn3* on adjacent pancreas cryosections of wild-type (A,B), *NeuroD*^{-/-} (C,D), *Arx*^{-/-} (E,F) and *Pax4*^{-/-} (G,H) deficient embryos at E14.5.

of endocrine cells (Mavropoulos et al., 2005) and for which we could confirm an absence of colocalization with the hormones at 24 hpf (Fig. 6F). Two other pancreatic markers, *isll* and *neurod*, were used in the double in situ hybridization experiments. *isll* is known to be expressed in postmitotic endocrine cells in mouse embryos (Ahlgren et al., 1997; Thor et al., 1991). The same seems to be true in zebrafish as, at 24 hpf, all mature hormone-expressing cells also express *isll* (Fig. 6E) and there is essentially no colocalization between *isll* and the endocrine marker *sox4b* (Fig. 6C). As for *NeuroD*, it is expressed in the murine *Ngn3* pancreatic precursors as well as in the mature hormone-expressing cells (Huang et al., 2000; Itkin-Ansari et al., 2005; Naya et al., 1997). In the same way, in zebrafish, *neurod* is expressed in the pancreatic progenitor cells as demonstrated by its colocalization with the *sox4b* factor (Mavropoulos et al., 2005), and in the mature hormone-expressing cells (Fig. 6D). It is important to note that the pancreatic endocrine precursors, labelled by *sox4b*, are localised in the ventral part of the pancreatic endoderm, whereas the more differentiated cells are located more dorsally (Fig. 6C and diagram).

The location of *rfx6* transcripts was then compared with that of these three pancreatic endocrine markers, *sox4b*, *neurod* and *isll* (Fig. 6G-N). At stage 18S (18 hpf), *rfx6* showed a perfect colocalization with *sox4b* (Fig. 6G). Progressively, *sox4b* expression became restricted to the ventral part of the *rfx6* expression domain (Fig. 6H,I), the dorsal part corresponding to *isll*-expressing cells (Fig. 6M,N). Finally, a total colocalization between *rfx6* and *neurod* was observed at all stages analysed (ie. 18, 24 and 30 hpf) (Fig. 6J-L). As expected based on *isll* and *neurod* colocalization, we detected *rfx6* transcripts in all endocrine cell types, ie. the *insulin*-, *glucagon*-, *ghrelin*- and *somatostatin*-expressing cells (see Fig. S4 in the supplementary material). Taken together, these data show that the pancreatic

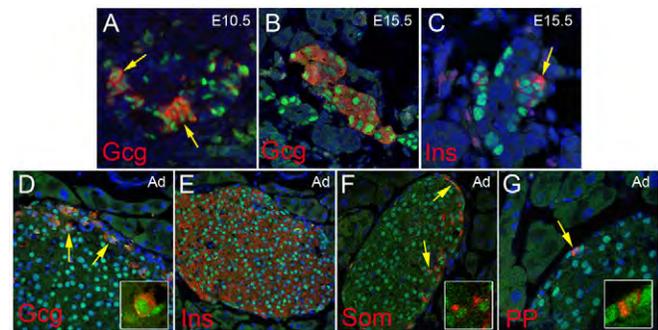


Fig. 5. *Rfx6* expression in mature islet cells in the mouse. (A-G) Double immunofluorescence revealing nuclear expression of *Rfx6* (green) in hormone-expressing cells (red) in developing alpha (A,B) and beta (C) cells, as well as in mature alpha (D), beta (E), delta (F) and PP (G) cells in adult islets of Langerhans. Gcg, glucagon; Ins, insulin; PP, pancreatic polypeptide; Som, somatostatin. Arrows point to double-positive cells; insets show magnifications.

expression pattern of *rfx6* is similar in zebrafish and in mice, as *rfx6* is expressed in the pancreatic endocrine progenitor cells as well as in the mature endocrine cells. By contrast, divergence exists for its expression in the gut.

Impaired endocrine cell differentiation and accumulation of islet progenitor cells in *rfx6* morphants in zebrafish

To assess the role of zebrafish *Rfx6* in pancreatic endocrine development, we abrogated *Rfx6* protein expression in the zebrafish embryo by injecting two distinct antisense MOs. The first MO targets the exon 2 splice donor site and the second targets the exon 2 splice acceptor site, leading to *rfx6* mRNA splicing disruption as shown by RT-PCR (see Fig. S5 in the supplementary material). As expected based on the restricted expression of *rfx6* in the pancreas, the injection of the MOs did not disturb the general morphology of the embryos (data not shown). By contrast, pancreas development was strongly perturbed as injection of either one or the other MO led to the same phenotype: an almost complete depletion of *glucagon*- and *ghrelin*-expressing cells together with a drastic reduction in the number of *somatostatin*-expressing cells (Fig. 7A-L). Conversely, the number of *insulin*-expressing cells was only slightly decreased (Fig. 7, right), although they were no longer clustered in a compact islet (Fig. 7K,L). To understand by which mechanism *Rfx6* affects endocrine cell differentiation, we analysed the expression of *sox4b*, *neurod* and *isll* genes in the morphants. The number of pancreatic endocrine progenitor cells labelled by *sox4b* was drastically increased at 24 hpf upon *rfx6* knock-down (Fig. 7P-R), whereas the number of more differentiated endocrine cells, labelled by *isll*, was significantly reduced (Fig. 7M-O). The lateral view of the developing pancreas at 30 hpf, stained for both *sox4b* and *isll*, highlights the drastic reduction of the *isll* dorsal domain together with the ventral expansion of *sox4b* (Fig. 7 Y-ZII). This increase does not seem to be due to an enhanced proliferation of the progenitors as we could not detect any change in the proportion of *pcna* (proliferating cell nuclear antigen)/*sox4b* double-positive cells (data not shown). On the contrary, the number of *neurod*-expressing cells was not affected in the morphants (Fig. 7S-U), indicating that the total number of endocrine progenitor and differentiated cells was not significantly perturbed. This strongly suggests that the increase of progenitor cells occurs at the expense of more differentiated cells

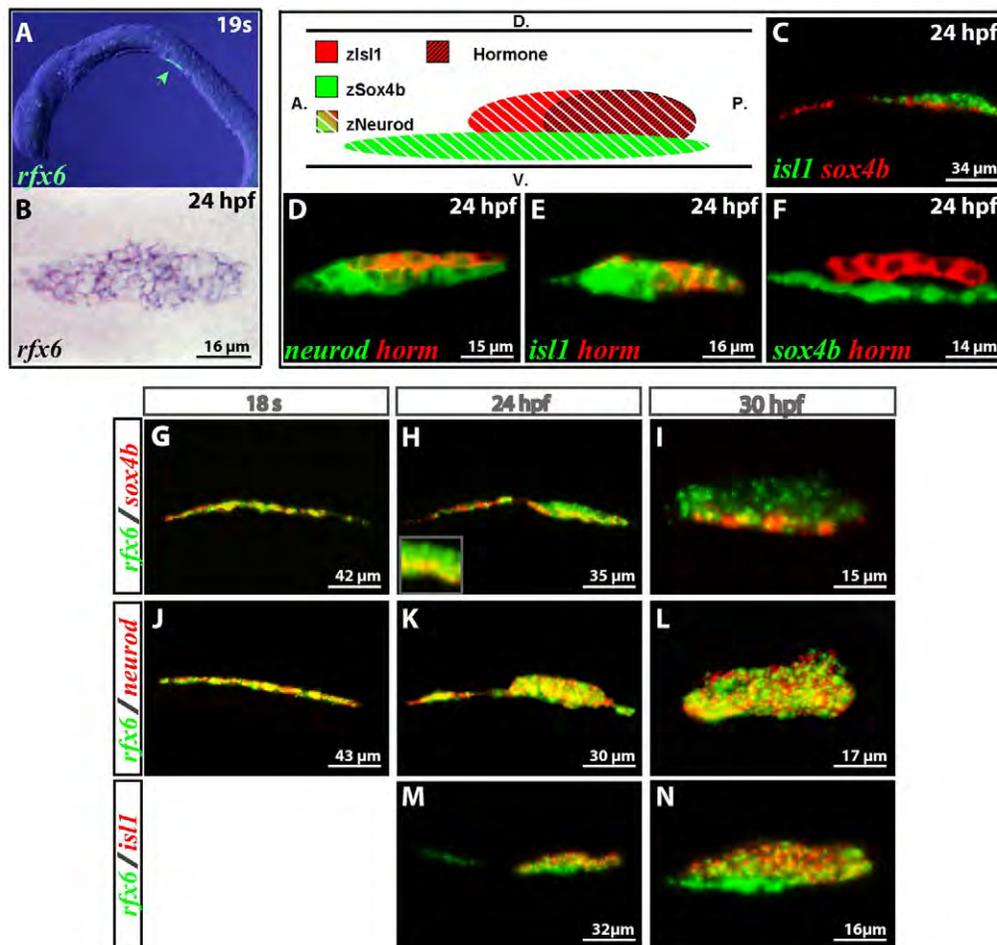


Fig. 6. Zebrafish *rfx6* is expressed in pancreatic endocrine progenitors as well as in mature endocrine cells. (A) Lateral view of fluorescent wholemount in situ hybridization (WISH) performed on stage 19S embryos with an *rfx6* antisense probe (arrowhead). (B) Ventral view of WISH with an *rfx6* probe on 24 hpf embryos. (C–N) Lateral views of the pancreas area from embryos analyzed by double-fluorescent WISH, anterior to the left, dorsal to the top. (C) The expression domains of *isl1* and *sox4b* are distinct. (D–F) Relative location of the hormone-expressing cells in relation to the *neurod*-, *isl1*- and *sox4b*-expressing cells. Hormone expression was detected by WISH using a cocktail of *insulin*, *glucagon*, *somatostatin* and *ghrelin* probes. The diagram shows the respective locations of *sox4b*-, *isl1*-, *neurod*-, and hormone-positive cells within the pancreatic area in a lateral view at 24 hpf. (G–I) At 18s, *rfx6* and *sox4b* expression domains completely overlap, whereas at 24 hpf and 30 hpf *sox4b* is maintained only in the ventral part of *rfx6* domain. (J–L) The expression pattern of *rfx6* completely overlaps the pancreatic *neurod* expression pattern. (M,N) The *isl1*-expressing cells are located in the dorsal part of the *rfx6* expression domain at 24 hpf and 30 hpf.

and that, in *rfx6* morphants, pancreatic endocrine cells are blocked in the progenitor stage. This blockage takes place prior to the expression of the transcription factor *arx*, which is essential for *glucagon*-expressing cell development (V. Verbruggen and B.P., unpublished). Indeed, we also observed a drastic reduction in the number of *arx*-expressing cells in the morphants (Fig. 6V–X). All of these results suggest that Rfx6 is essential for the transition from pancreatic endocrine progenitors to more differentiated *glucagon*-, *ghrelin*- and *somatostatin*-expressing cells.

DISCUSSION

To identify novel downstream effectors of the proendocrine function of the transcription factor Ngn3, we determined the gene expression profile of isolated Ngn3-positive progenitors. Our data are complementary to, and further extend, similar published studies (Gu et al., 2004; White et al., 2008) owing to the combination of an original mouse model with unaltered levels of

Ngn3 transcripts and the use of a very representative Affymetrix array. This is illustrated by the identification of the uncharacterized winged helix transcription factor Rfx6 which has not been found in the above mentioned studies, but was reported recently to be expressed in Ngn3 progenitor cells based on RT-PCR experiments (Miyatsuka et al., 2009). In the present study, we characterize *Rfx6* expression in the mouse and its ortholog in zebrafish, and report its crucial role in the progression of islet cell differentiation in the latter.

Rfx transcription factors in the mouse pancreas

Regulatory factor X (Rfx) proteins are transcription factors conserved from *C. elegans* to mammals. This protein family shares a typical DNA binding domain containing a winged helix motif recognizing a bipartite DNA sequence known as X-box. A recent survey of mammalian genomes has identified seven Rfx genes in mouse and human databases (Aftab et al., 2008). Major findings regarding the

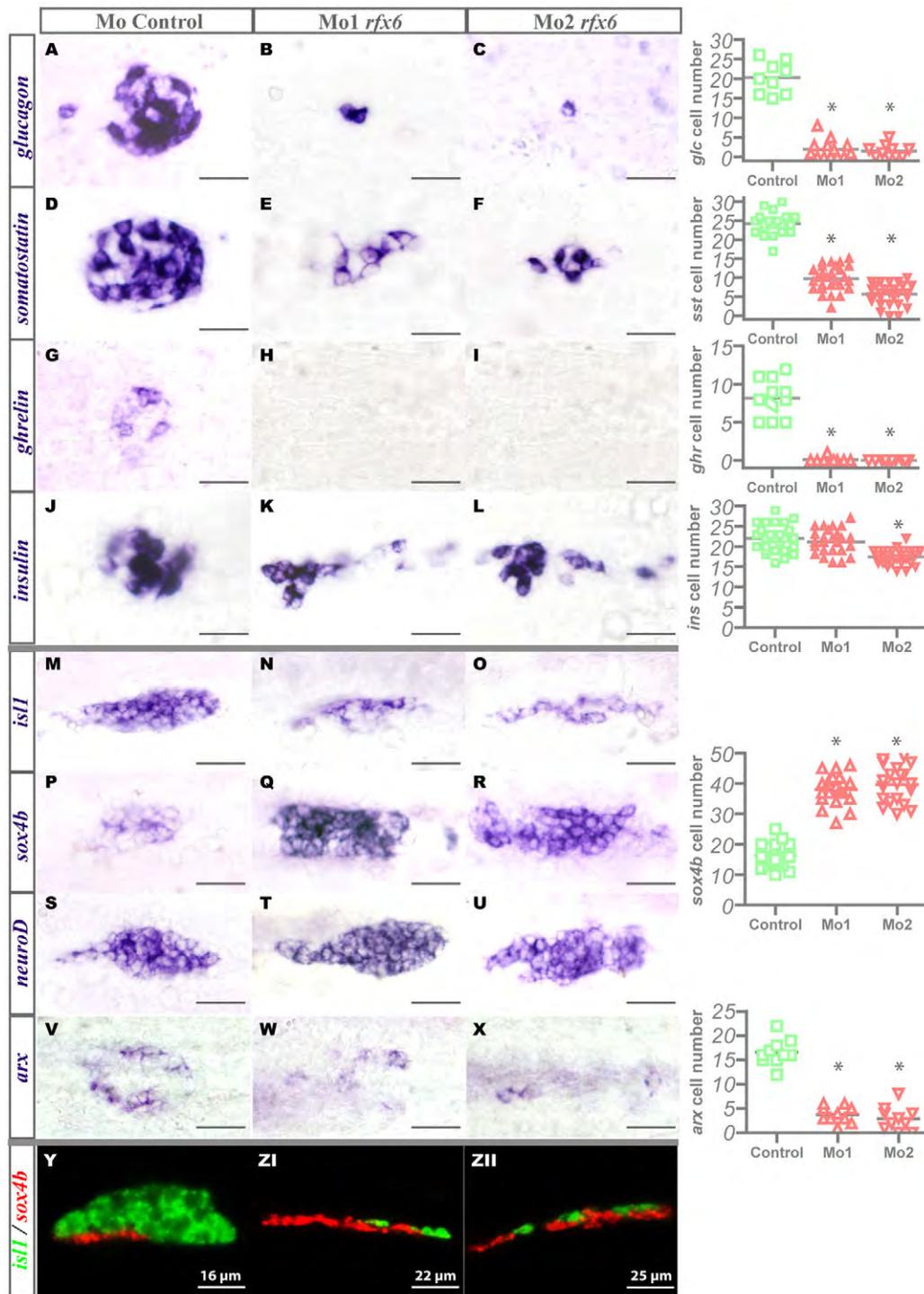


Fig. 7. Impaired endocrine cell differentiation, accumulation of islet progenitor cells and failure of insulin-expressing cells clustering in *rfx6* morphants in the zebrafish. (A–X) Ventral views of the pancreas area from embryos analysed by WISH, anterior to the left. (A–L) Hormone expression in the control and *rfx6* morphants at 30 hpf. (M–X) Pancreatic expression of *isl1*, *sox4b* and *neuroD* at 24 hpf and *arx* at 30 hpf in the control and *rfx6* morphants. (Y–ZII) Lateral views of the pancreas area from 30 hpf embryos analysed by double-fluorescent WISH for *sox4b* and *isl1*, anterior to the left, dorsal to the top. Quantifications (right side of figure) represent the number of positive cells per embryo for controls and morphants. Asterisks (*) indicate that the difference between cell number in control and morphants is statistically significant by Student's *t*-test ($P < 0.001$). Scale bar: 14 μm in A–L; 18 μm in M–X.

function of Rfx proteins arose from studies in invertebrates. *daf-19*, the unique Rfx gene in *C. elegans*, has been shown to be a crucial regulator of ciliogenesis (Swoboda et al., 2000). Similarly, Rfx in *Drosophila* is necessary for ciliated sensory neuron differentiation (Dubruille et al., 2002). In invertebrates, Rfx transcription factors are thought to control the transcription of proteins involved in cilia assembly in a process called intraflagellar transport (IFT). Cilia are organelles found almost ubiquitously in vertebrate cells and are involved in numerous developmental processes and human genetic disorders (ciliopathies) (Gerdes et al., 2009). Recently, *rfx2* has been

shown to control ciliogenesis in the zebrafish pronephros (Liu et al., 2007). In the mouse, the expression of *Rfx3* was reported in the endocrine pancreas and it has been shown that *Rfx3* loss-of-function resulted in impaired islet cell composition and glucose tolerance (Ait-Lounis et al., 2007). This phenotype was associated with abnormal formation of primary cilia on islet cells. However, because the pancreatic deletion of *Kif3a*, a gene involved in IFT, did not result in endocrine failure (Cano et al., 2006), it is not clear whether the islet phenotype in *Rfx3*-deficient mice results from defects in cilia formation. In this study, we report for the first time the pancreatic

expression of *Rfx6* in the mouse and zebrafish, as well as its function in zebrafish. Interestingly, in both the embryonic and adult mouse, *Rfx6* and *Rfx3* have a very similar expression pattern in the islet lineage. Indeed, both are expressed in *Ngn3*-positive endocrine progenitors and expression is maintained in developing and adult islet cells (this study; Ait-Lounis et al., 2007). Furthermore, *Rfx3* was also found enriched (FC 6, FDR 0.051) in our microarray profiles from sorted E15.5 *Ngn3*+/*eYFP*+ cells, whereas *Rfx1*, *Rfx5* and *Rfx7* were not (*Rfx2* and *Rfx4* are not present on the microarray). As expected from the similar pancreatic expression pattern of *Rfx6* and *Rfx3* in the mouse endocrine pancreas, hormone-expressing *Rfx6*-positive cells are ciliated (data not shown). Primary cilia can be seen on endocrine cells in the wild-type zebrafish embryo as well; however, we found that pancreatic cells remain ciliated after *rfx6* knock-down, suggesting that *rfx6* is not required for ciliogenesis in the zebrafish pancreas (see Fig. S6, Movie S1 and S2 in the supplementary material). Nevertheless, it is worth mentioning that *Rfx* proteins are known to dimerize and interact physically with other members of the family (Wolfe et al., 2008). Given that *Rfx3* and *Rfx6* are similarly expressed in the islet lineage, one cannot exclude that *Rfx3* and *Rfx6* could cooperate to regulate common target genes in developing and adult mouse islet cells.

Rfx6 in the hierarchy of transcription factors controlling islet cells differentiation in the mouse

Rfx6 expression in the entire primitive gut epithelium suggests an early function in endoderm specification and/or maintenance in mouse. Such a broad endodermal expression was not revealed for *Rfx3*, although the weak sensitivity of the *Rfx3* antisense probe could preclude the detection of low amounts of transcripts. In the developing pancreatic buds, *Rfx6* is then progressively excluded from multipotent *Pdx1*+/*Ptf1a*+ progenitor cells and becomes restricted to the endocrine cells at E10.5, suggesting that *Pdx1* and/or *Ptf1a* might repress *Rfx6* expression in uncommitted pancreatic endodermal cells as development proceeds. Additionally, both the expression of *Rfx6* in postmitotic *Ngn3*-positive progenitor cells and its maintenance in hormone-expressing cells in the embryo and adult islet lineage, further suggest a dual role in islet cell specification/maturation and function. At all stages analyzed, *Rfx6* is lost in *Ngn3*-deficient pancreata, demonstrating that *Rfx6* is a downstream target of *Ngn3*. Importantly, the *Rfx6* expression pattern overlaps with *Insm1* and *NeuroD*, two previously reported direct target genes of *Ngn3* (Mellitzer et al., 2006; Huang et al., 2000). Furthermore, *Rfx6* and *Insm1* are maintained in adult islet cells and their expression is similarly independent of *Arx*, *Pax4* and *NeuroD* in the embryonic pancreas. These findings suggest that *Rfx6* and *Insm1* would act upstream of *Arx*, *Pax4* and *NeuroD*. Another possibility could be that, in addition to islet subtype specification programs, *Ngn3* would regulate independent generic subroutines and thus *Rfx6* (and *Insm1*) could belong to a novel parallel branch of the *Ngn3*-dependent network. However, whether *Rfx6* and *Insm1* are in the same pathway remains to be determined. At this point, it will be important, in complement to epistasis analysis in knock-out mice, to identify direct target genes of islet transcription factors using ChipSeq technology to decipher regulatory branches controlling cellular subtype or generic properties.

Role of *rfx6* in zebrafish endocrine cell differentiation

To determine *Rfx6* function, we took advantage of the zebrafish system, which has been shown to be an appropriate model to study islet cell development. Indeed, several transcription factors have been demonstrated to have a conserved function regarding endocrine cell

differentiation between the zebrafish and the mouse (Pauls et al., 2007; Zecchin et al., 2004; Song et al., 2007; Yee et al., 2001). Notably however, a functional orthologue for the mouse *Ngn3*, which regulates the islet cell fate decision, has not yet been identified in the fish. However, *sox4b* is found predominantly in endocrine precursor cells. Zebrafish orthologues for all mouse *Rfx* proteins could undoubtedly be identified (see Fig. S1 in the supplementary material). *rfx6*, but not *rfx3*, was found expressed in the islet lineage in a pattern reminiscent of mouse *Rfx6* (islet precursors and mature islet cells), suggesting functional conservation. Importantly however, *rfx6* was not present in the gut endoderm, which might reflect differences in the mechanisms controlling endoderm and/or pancreas specification between the mouse and zebrafish. Injection of two distinct MOs resulted in severe perturbation of islet cell development. *Glucagon*- and *ghrelin*-expressing cells were almost absent and *somatostatin*-expressing cells were drastically reduced (60-77% decrease). Importantly, *arx* expression was strongly reduced in *rfx6* morphants, suggesting that *rfx6* is upstream of *arx* in the regulatory cascade controlling alpha cell differentiation. By contrast to *glucagon*-, *ghrelin*-, and *somatostatin*-expressing cells, the number of *insulin*-expressing cells was only mildly affected (19% reduction with the most efficient morpholino, MO2), but these cells failed to cluster. Thus, the *rfx6* knock-down in zebrafish does not block the differentiation of all pancreatic endocrine cells to the same extent. One interpretation could be that *rfx6* is involved in endocrine subtype specification and particularly in the development of *glucagon*- and *ghrelin*-expressing cells. Another explanation would support a role in islet cell maturation. This latter hypothesis is reinforced by the dramatic increase in the number of endocrine progenitors in *rfx6* morphants concomitantly with a reduction of mature endocrine cells. In this model, the mild reduction of beta cells would be due to the early apparition of *insulin*-expressing cells which appear before the onset of *rfx6* expression. Indeed, in zebrafish, *insulin*-expressing cells first appear around 15 hpf, followed 2 hours later by the *somatostatin*-expressing cells, then by the *glucagon*-expressing cells at 20 hpf (Biemar et al., 2001) and finally the *ghrelin*-expressing cells around 22 hpf (this study; data not shown). As for *rfx6*, transcripts start to be detected only from 17 hpf onwards. We can thus expect that when enough *Rfx6* proteins are present in the cells to control the transition from pancreatic progenitors to more differentiated cells, the differentiation of the first endocrine cells, ie. *insulin*- and *somatostatin*-expressing cells, is already committed.

In conclusion, our gene expression profiling in mouse islet progenitors revealed *Rfx6*, a novel *Ngn3*-dependent winged helix transcription factor. Expression and functional studies demonstrated that *rfx6* is essential for normal islet cell development in zebrafish. Genetic studies in the mouse and experiments designed to identify *Rfx6* target genes are now required to further decipher the function of *Rfx6* in the islet lineage.

Acknowledgements

We thank P. Collombat and A. Mansouri (Max Planck Institute, Goettingen) for providing *Pax4*- and *Arx*-deficient mouse embryos, J. Lee (Geron) for sharing the *NeuroD* knock-out mice, V. Lobstein, M. Messmer, M. Poulet, C. Collin, C. Pardanaud-Glavieux and B. Schubaur for excellent technical assistance, C. Thibault and D. Dembele from the IGBMC Microarray and Sequencing Platform and R. B. Helweg-Larsen from the Microarray Center, Rigshospitalet, Copenhagen. We are grateful to the GIGA Zebrafish, Transgenics, Imaging and Flow Cytometry facilities and to the IGBMC Imaging Centre. The sequencing results were obtained thanks to V. Dhennin, Genotranscriptomics Platform, GIGA, University of Liège, <http://www.giga.ulg.ac.be/>. This work was supported by the Institut National de la Santé et de la Recherche Médicale (INSERM), and by grants from the National Institutes of Health (NIH) Beta Cell Biology Consortium to G.G. (U19-DK061244 and U19-DK072495), the Juvenile Diabetes Research Foundation (JDRF) to G.G. and P.R. (AIP Cellules

souches A03139MS), the Association Française des Diabétiques (AFD) to G.G., the European Union Integrated Project 6th FP Beta Cell Therapy to G.G., P.R., and B.P. (LSHB-CT-2005-512145), and the Belgian State Interuniversity Attraction Poles Program (SSTC, PAI). J.S. is a recipient of a fellowship from the University of Strasbourg and Association pour la Recherche sur le Cancer. A.B. is a recipient of a Fellowship from the Fondation pour la Recherche Médicale. L.F. holds a doctoral fellowship from the Fonds National pour la Recherche scientifique (FNRS). M.L.V. and B.P. are each 'Chercheur Qualifié' of the FNRS. Deposited in PMC for release after 12 months.

Competing interests statement

The authors declare no competing financial interests.

Supplementary material

Supplementary material for this article is available at

<http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.041673/-DC1>

References

- Aftab, S., Semenc, L., Chu, J. S. and Chen, N. (2008). Identification and characterization of novel human tissue-specific RFX transcription factors. *BMC Evol. Biol.* **8**, 226.
- Ahlgren, U., Pfaff, S. L., Jessell, T. M., Edlund, T. and Edlund, H. (1997). Independent requirement for ISL1 in formation of pancreatic mesenchyme and islet cells. *Nature* **385**, 257-260.
- Ait-Lounis, A., Baas, D., Barras, E., Benadiba, C., Charollais, A., Nlend, N. R., Liegeois, D., Meda, P., Durand, B. and Reith, W. (2007). Novel function of the ciliogenic transcription factor RFX3 in development of the endocrine pancreas. *Diabetes* **56**, 950-959.
- Argenton, F., Zecchin, E. and Bortolussi, M. (1999). Early appearance of pancreatic hormone-expressing cells in the zebrafish embryo. *Mech. Dev.* **87**, 217-221.
- Biemar, F., Argenton, F., Schmidtke, R., Epperlein, S., Peers, B. and Driever, W. (2001). Pancreas development in zebrafish: early dispersed appearance of endocrine hormone expressing cells and their convergence to form the definitive islet. *Dev. Biol.* **230**, 189-203.
- Cano, D. A., Sekine, S. and Hebrok, M. (2006). Primary cilia deletion in pancreatic epithelial cells results in cyst formation and pancreatitis. *Gastroenterology* **131**, 1856-1869.
- Claiborn, K. C. and Stoffers, D. A. (2008). Toward a cell-based cure for diabetes: advances in production and transplant of beta cells. *Mt. Sinai J. Med.* **75**, 362-371.
- Collombat, P., Mansouri, A., Hecksher-Sorensen, J., Serup, P., Krull, J., Gradwohl, G. and Gruss, P. (2003). Opposing actions of Arx and Pax4 in endocrine pancreas development. *Genes Dev.* **17**, 2591-2603.
- D'Amour, K. A., Bang, A. G., Eliazer, S., Kelly, O. G., Agulnick, A. D., Smart, N. G., Moorman, M. A., Kroon, E., Carpenter, M. K. and Baetge, E. E. (2006). Production of pancreatic hormone-expressing endocrine cells from human embryonic stem cells. *Nat. Biotechnol.* **24**, 1392-1401.
- Devos, N., Deflorian, G., Biemar, F., Bortolussi, M., Martial, J. A., Peers, B. and Argenton, F. (2002). Differential expression of two somatostatin genes during zebrafish embryonic development. *Mech. Dev.* **115**, 133-137.
- Dubruille, R., Laurencon, A., Vandaele, C., Shishido, E., Coulon-Bublex, M., Swoboda, P., Couble, P., Kernan, M. and Durand, B. (2002). Drosophila regulatory factor X is necessary for ciliated sensory neuron differentiation. *Development* **129**, 5487-5498.
- Ekker, S. C. (2000). Morphants: a new systematic vertebrate functional genomics approach. *Yeast* **17**, 302-306.
- Gerdes, J. M., Davis, E. E. and Katsanis, N. (2009). The vertebrate primary cilium in development, homeostasis, and disease. *Cell* **137**, 32-45.
- 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. *Proc. Natl. Acad. Sci. USA* **97**, 1607-1611.
- Gu, G., Dubauskaite, J. and Melton, D. A. (2002). Direct evidence for the pancreatic lineage: NGN3+ cells are islet progenitors and are distinct from duct progenitors. *Development* **129**, 2447-2457.
- Gu, G., Wells, J. M., Dombkowski, D., Preffer, F., Aronow, B. and Melton, D. A. (2004). Global expression analysis of gene regulatory pathways during endocrine pancreatic development. *Development* **131**, 165-179.
- Heller, R. S., Jenny, M., Collombat, P., Mansouri, A., Tomasetto, C., Madsen, O. D., Mellitzer, G., Gradwohl, G. and Serup, P. (2005). Genetic determinants of pancreatic epsilon-cell development. *Dev. Biol.* **286**, 217-224.
- 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. *Mol. Cell. Biol.* **20**, 3292-3307.
- Itkin-Ansari, P., Marcora, E., Geron, I., Tyrberg, B., Demeterco, C., Hao, E., Padilla, C., Ratineau, C., Leiter, A., Lee, J. E. et al. (2005). NeuroD1 in the endocrine pancreas: localization and dual function as an activator and repressor. *Dev. Dyn.* **233**, 946-953.
- Jenny, M., Uhl, C., Roche, C., Duluc, I., Guillermin, V., Guillemot, F., Jensen, J., Kedinger, M. and Gradwohl, G. (2002). Neurogenin3 is differentially required for endocrine cell fate specification in the intestinal and gastric epithelium. *EMBO J.* **21**, 6338-6347.
- Johansson, K. A., Dursun, U., Jordan, N., Gu, G., Beermann, F., Gradwohl, G. and Grapin-Botton, A. (2007). Temporal control of neurogenin3 activity in pancreas progenitors reveals competence windows for the generation of different endocrine cell types. *Dev. Cell* **12**, 457-465.
- Korz, V., Edlund, T. and Thor, S. (1993). Zebrafish primary neurons initiate expression of the LIM homeodomain protein Isl-1 at the end of gastrulation. *Development* **118**, 417-425.
- Korz, V., Sleptsova, I., Liao, J., He, J. and Gong, Z. (1998). Expression of zebrafish bHLH genes ngn1 and nrd defines distinct stages of neural differentiation. *Dev. Dyn.* **213**, 92-104.
- Kroon, E., Martinson, L. A., Kadoya, K., Bang, A. G., Kelly, O. G., Eliazer, S., Young, H., Richardson, M., Smart, N. G., Cunningham, J. et al. (2008). Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells in vivo. *Nat. Biotechnol.* **26**, 443-452.
- Lee, C. S., Perreault, N., Brestelli, J. E. and Kaestner, K. H. (2002). Neurogenin 3 is essential for the proper specification of gastric enteroendocrine cells and the maintenance of gastric epithelial cell identity. *Genes Dev.* **16**, 1488-1497.
- Liu, Y., Pathak, N., Kramer-Zucker, A. and Drummond, I. A. (2007). Notch signaling controls the differentiation of transporting epithelia and multiciliated cells in the zebrafish pronephros. *Development* **134**, 1111-1122.
- Madsen, O. D. and Serup, P. (2006). Towards cell therapy for diabetes. *Nat. Biotechnol.* **24**, 1481-1483.
- Mavropoulos, A., Devos, N., Biemar, F., Zecchin, E., Argenton, F., Edlund, H., Motte, P., Martial, J. A. and Peers, B. (2005). sox4b is a key player of pancreatic alpha cell differentiation in zebrafish. *Dev. Biol.* **285**, 211-223.
- Mellitzer, G., Martin, M., Sidhoum-Jenny, M., Orvain, C., Barths, J., Seymour, P. A., Sander, M. and Gradwohl, G. (2004). Pancreatic islet progenitor cells in neurogenin 3-yellow fluorescent protein knock-add-on mice. *Mol. Endocrinol.* **18**, 2765-2776.
- Mellitzer, G., Bonne, S., Luco, R. F., Van de Castele, M., Lenne-Samuel, N., Collombat, P., Mansouri, A., Lee, J., Lan, M., Pipeleers, D. et al. (2006). IA1 is NGN3-dependent and essential for differentiation of the endocrine pancreas. *EMBO J.* **25**, 1344-1352.
- Milewski, W. M., Duguay, S. J., Chan, S. J. and Steiner, D. F. (1998). Conservation of PDX-1 structure, function, and expression in zebrafish. *Endocrinology* **139**, 1440-1449.
- Miyatsuka, T., Li, Z. and German, M. S. (2009). Chronology of islet differentiation revealed by temporal cell labeling. *Diabetes* **58**, 1863-1868.
- Murtaugh, L. C. (2007). Pancreas and beta-cell development: from the actual to the possible. *Development* **134**, 427-438.
- 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 Dev.* **11**, 2323-2334.
- Pauls, S., Zecchin, E., Tiso, N., Bortolussi, M. and Argenton, F. (2007). Function and regulation of zebrafish nkx2.2a during development of pancreatic islet and ducts. *Dev. Biol.* **304**, 875-890.
- Ploner, A., Calza, S., Gusnanto, A. and Pawitan, Y. (2006). Multidimensional local false discovery rate for microarray studies. *Bioinformatics* **22**, 556-565.
- 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. *Dev. Biol.* **270**, 443-454.
- Schwitzgebel, V. M., Scheel, D. W., Conners, J. R., Kalamaras, J., Lee, J. E., Anderson, D. J., Sussel, L., Johnson, J. D. and German, M. S. (2000). Expression of neurogenin3 reveals an islet cell precursor population in the pancreas. *Development* **127**, 3533-3542.
- Shapiro, A. M., Lakey, J. R., Ryan, E. A., Korbitt, G. S., Toth, E., Warnock, G. L., Kneteman, N. M. and Rajotte, R. V. (2000). Islet transplantation in seven patients with type 1 diabetes mellitus using a glucocorticoid-free immunosuppressive regimen. *New Engl. J. Med.* **343**, 230-238.
- 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. *J. Biol. Chem.* **278**, 38254-38259.
- Song, J., Kim, H. J., Gong, Z., Liu, N. A. and Lin, S. (2007). Vhnf1 acts downstream of Bmp, Fgf, and RA signals to regulate endocrine beta cell development in zebrafish. *Dev. Biol.* **303**, 561-575.
- 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.
- Swoboda, P., Adler, H. T. and Thomas, J. H. (2000). The RFX-type transcription factor DAF-19 regulates sensory neuron cilium formation in *C. elegans*. *Mol. Cell* **5**, 411-421.
- Thor, S., Ericson, J., Brannstrom, T. and Edlund, T. (1991). The homeodomain LIM protein Isl-1 is expressed in subsets of neurons and endocrine cells in the adult rat. *Neuron* **7**, 881-889.

- Watada, H., Scheel, D. W., Leung, J. and German, M. S.** (2003). Distinct gene expression programs function in progenitor and mature islet cells. *J. Biol. Chem.* **278**, 17130-17140.
- White, P., May, C. L., Lamounier, R. N., Brestelli, J. E. and Kaestner, K. H.** (2008). Defining pancreatic endocrine precursors and their descendants. *Diabetes* **57**, 654-668.
- Wolfe, S. A., Vanwert, J. M. and Grimes, S. R.** (2008). Transcription factor RFX4 binding to the testis-specific histone H1t promoter in spermatocytes may be important for regulation of H1t gene transcription during spermatogenesis. *J. Cell. Biochem.* **105**, 61-69.
- Yee, N. S., Yusuff, S. and Pack, M.** (2001). Zebrafish pdx1 morphant displays defects in pancreas development and digestive organ chirality, and potentially identifies a multipotent pancreas progenitor cell. *Genesis* **30**, 137-140.
- Zecchin, E., Mavropoulos, A., Devos, N., Filippi, A., Tiso, N., Meyer, D., Peers, B., Bortolussi, M. and Argenton, F.** (2004). Evolutionary conserved role of ptf1a in the specification of exocrine pancreatic fates. *Dev. Biol.* **268**, 174-184.
- Zecchin, E., Filippi, A., Biemar, F., Tiso, N., Pauls, S., Ellertsdottir, E., Gnugge, L., Bortolussi, M., Driever, W. and Argenton, F.** (2007). Distinct delta and jagged genes control sequential segregation of pancreatic cell types from precursor pools in zebrafish. *Dev. Biol.* **301**, 192-204.

A

```

RFX6  LQWLIEDNYIVCEQVCLFRRCILYAHYLDPCRHEKLEPACAATFQKTIQKFPLLTARRLQTRGHSYHYHGIQIIESAYYHSVYSQKDLTR
RFX4  LQWLEENYEIAESVCIFRSALYMHYLDPCFKNDTQPVNAASFVKIIRQQFPQLTARRLQTRGDSYHYHGIQAVIESQYYDVNYSKKGAAW
RFX3  LQWLIDNYFTAFQVSLFRSTLYNHYL RHQDEHKLDQVVAISFGKLIISIFMGLRARRLQTRGNSYHYHGIQAVKFDSPLNRLQEDMQYMAN
RFX2  LQWLIDNYETAEGVSLFRSSLYNHYL RHQDEHKLEPVAASFGKLIISVFMGLRARRLQTRGNSYHYHGIQAVKFDSPLNRLQEDTQYMAN
RFX1  LQWLIDNYETAEGVSLFRSTLYCHYLLHQDEGKLEPVAASFGKLIISVFMGLRARRLQTRGNSYHYHGIQAVKFDSPLNRLMEDQQHMAM
RFX5  YRWIRNHLEEHMDTCLFKQSVYDARKYCESLACDRPLSTANFGKIIIEIFPDIKARRLQTRGDSYCYSGIIRRTLVSMPPPLPGLDLKQSE
RFX7  FSWIRNTLEIHPETSLEKQEVYDEKSYCDNLG-YHPLSADFGKIMKNVFNMKARRLQTRGKSYCYSGLRKIAFVHMPTLPNLDLHKTG

```

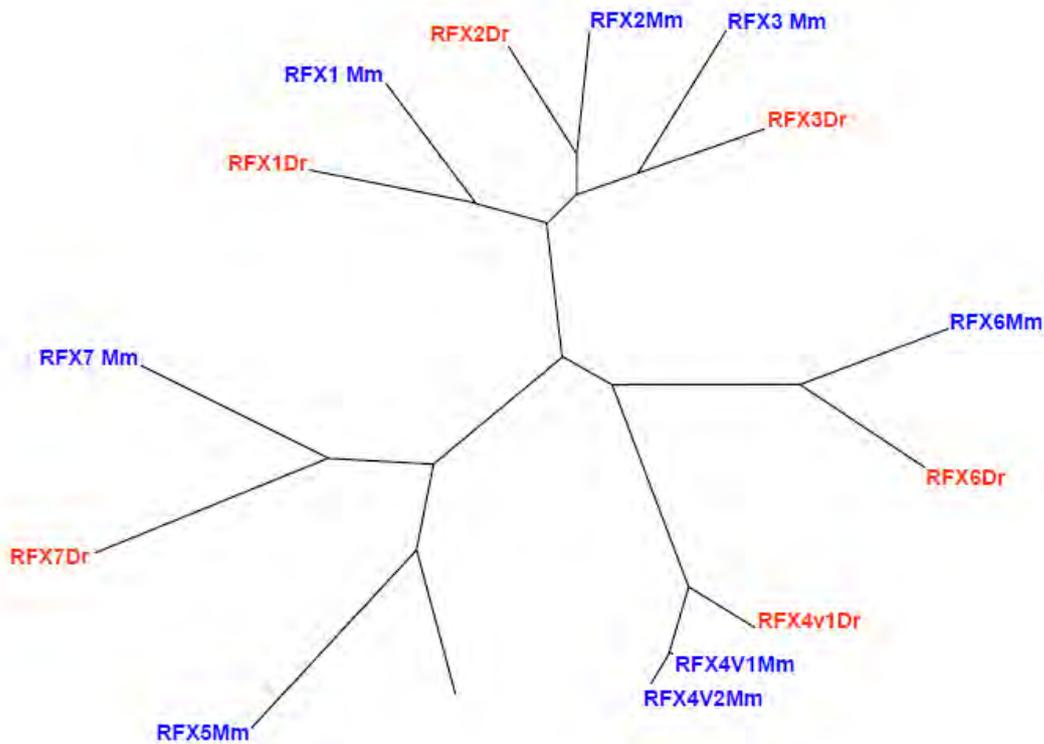
B

Fig. S1. Mouse and zebrafish Rfx proteins. (A) Alignment of mouse Rfx DNA binding domains using ClustalW. (B) Comparison of mouse Rfx proteins and zebrafish orthologues. The dendrogram was generated by comparing mouse and zebrafish Rfx proteins using ClustalW and Treeview. Mouse sequences used for comparison: Rfx1 Mm (NP_033081), Rfx2 Mm (NP_033082), Rfx3 Mm (NP_035395), Rfx4 Mm (NP_001020089), Rfx5 Mm (NP_059091), Rfx6 Mm (NP_001152861), Rfx7 Mm (NP_001028708). Zebrafish sequences used for comparison: Rfx1 Dr (ENSARP00000039332), Rfx2 Dr (ENSARP00000010260), Rfx3 Dr (ENSARP00000041235), Rfx4 Dr (ENSARP00000032295), Rfx5 Dr (ENSARP00000086706), Rfx6 Dr (ENSART00000061122), Rfx7 Dr (ENSARP00000102179). Dr, *Danio rerio*; Mm, *Mus musculus*.

RFX6 Ki-67

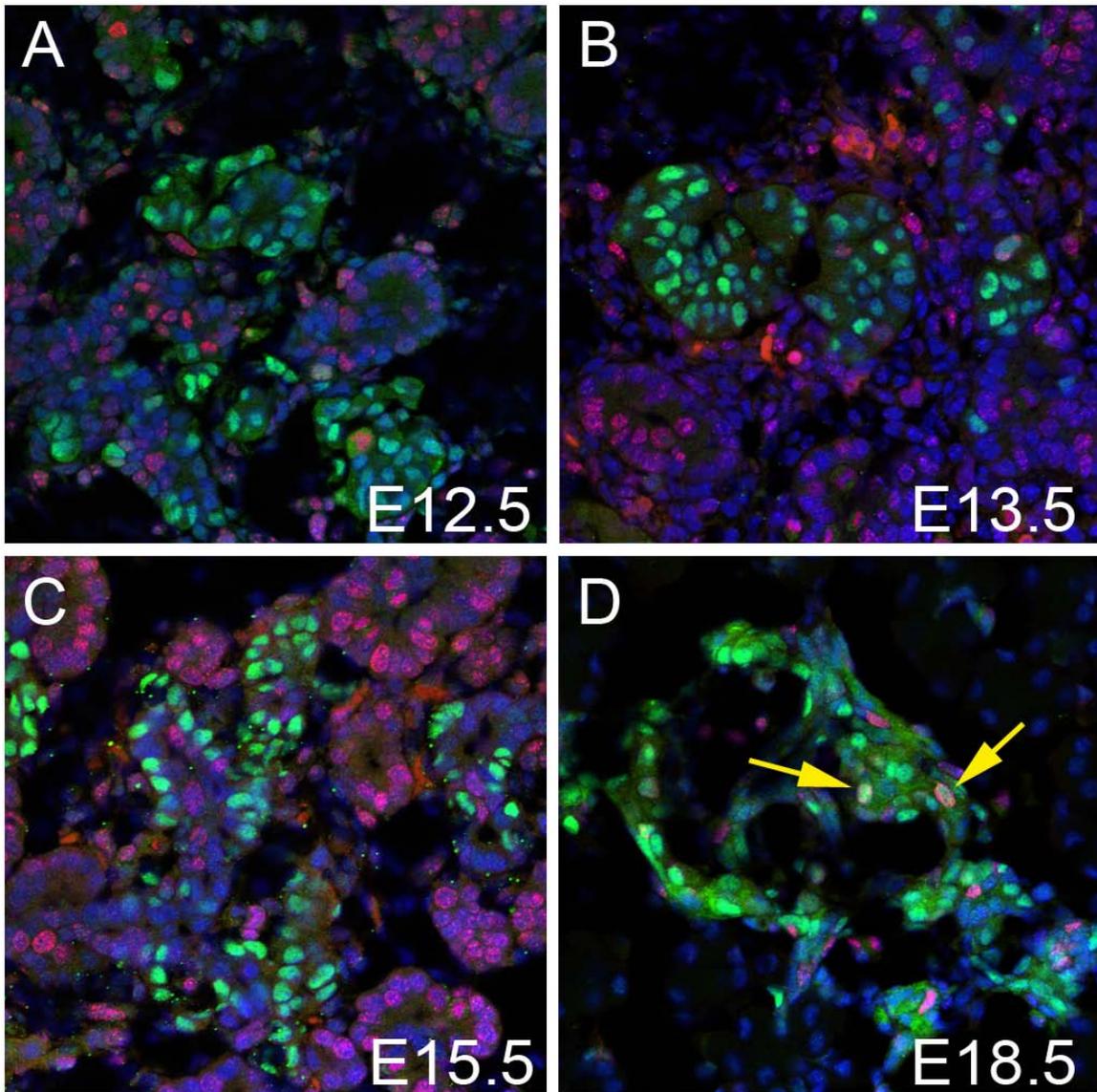


Fig. S2. Rfx6-expressing cells are postmitotic in the mouse pancreas. Double immunofluorescence for Rfx6 and Ki-67 on cryosections of mouse pancreata at different developmental stages showing that Rfx6 is mainly expressed in non-dividing Ki-67-negative cells. Arrows point to rare double-positive cells.

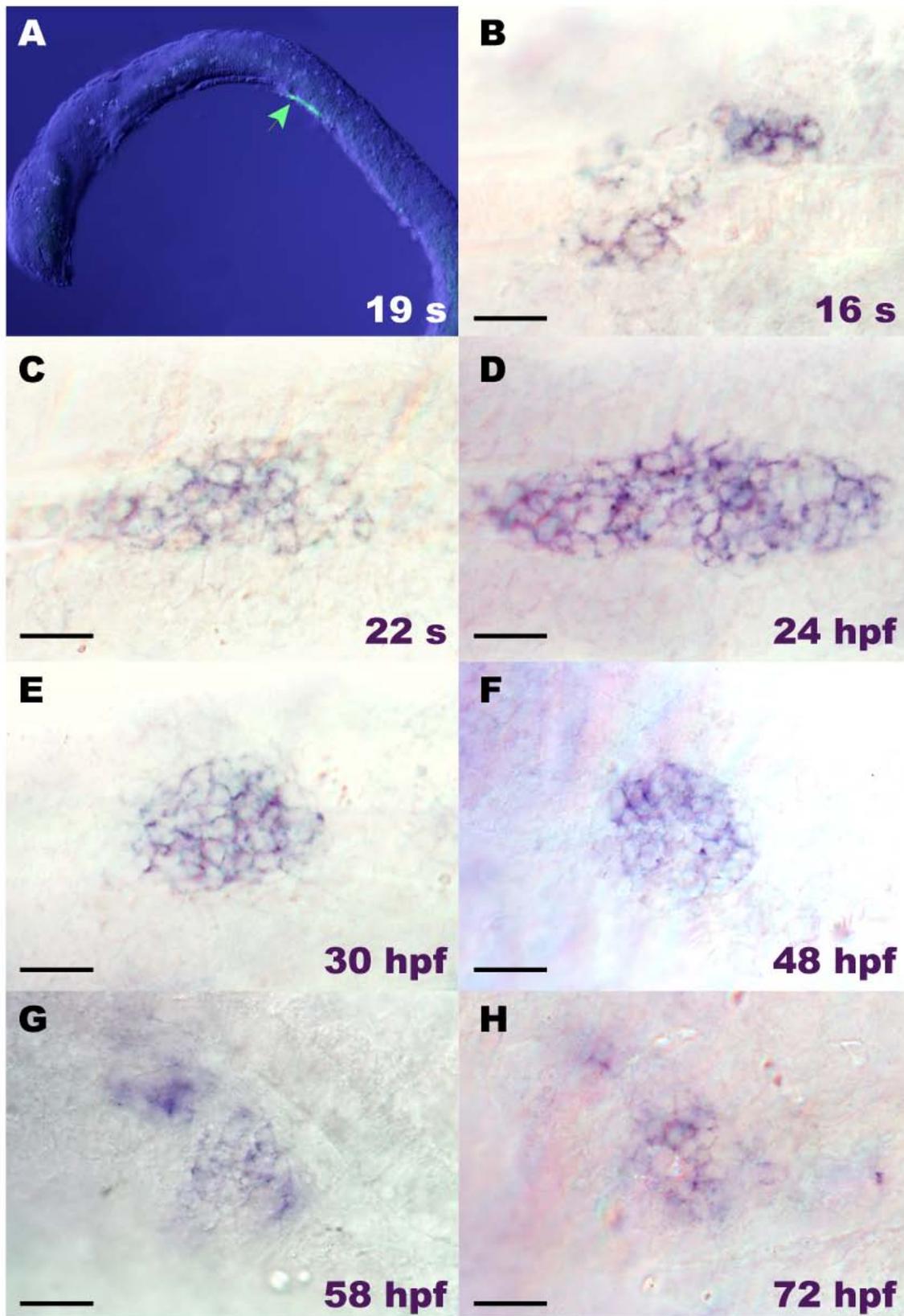


Fig. S3. *rfx6* is exclusively expressed in the pancreas area during early zebrafish development. (A) General view of embryo stained by fluorescent wholemount in situ hybridization (WISH) for the expression of *rfx6*. (B-H) Pancreatic expression starts around 17 hpf, peaks around 24 hpf and persists at least until 72 hpf. Ventral views, anterior to the left. Scale bars: 10 μ m. Panels A and D are the same as panels A and B in Fig. 6.

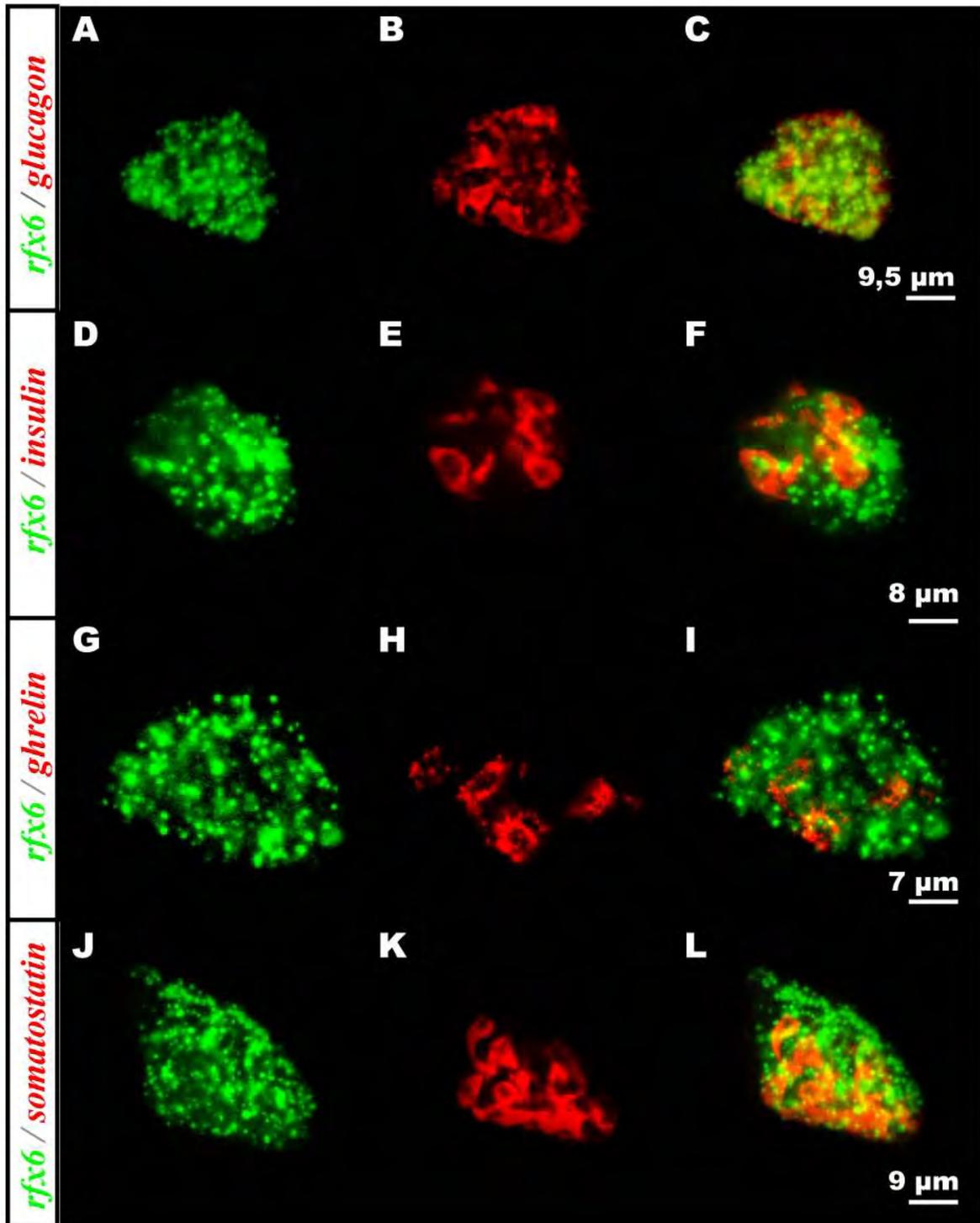


Fig. S4. *rfx6* is expressed in all of the hormone-expressing cells in zebrafish. Ventral views of the pancreas area from 30 hpf embryos analysed by double-fluorescent wholemount in situ hybridization (WISH), anterior to the left. The right column presents the overlay of the left and middle panels.

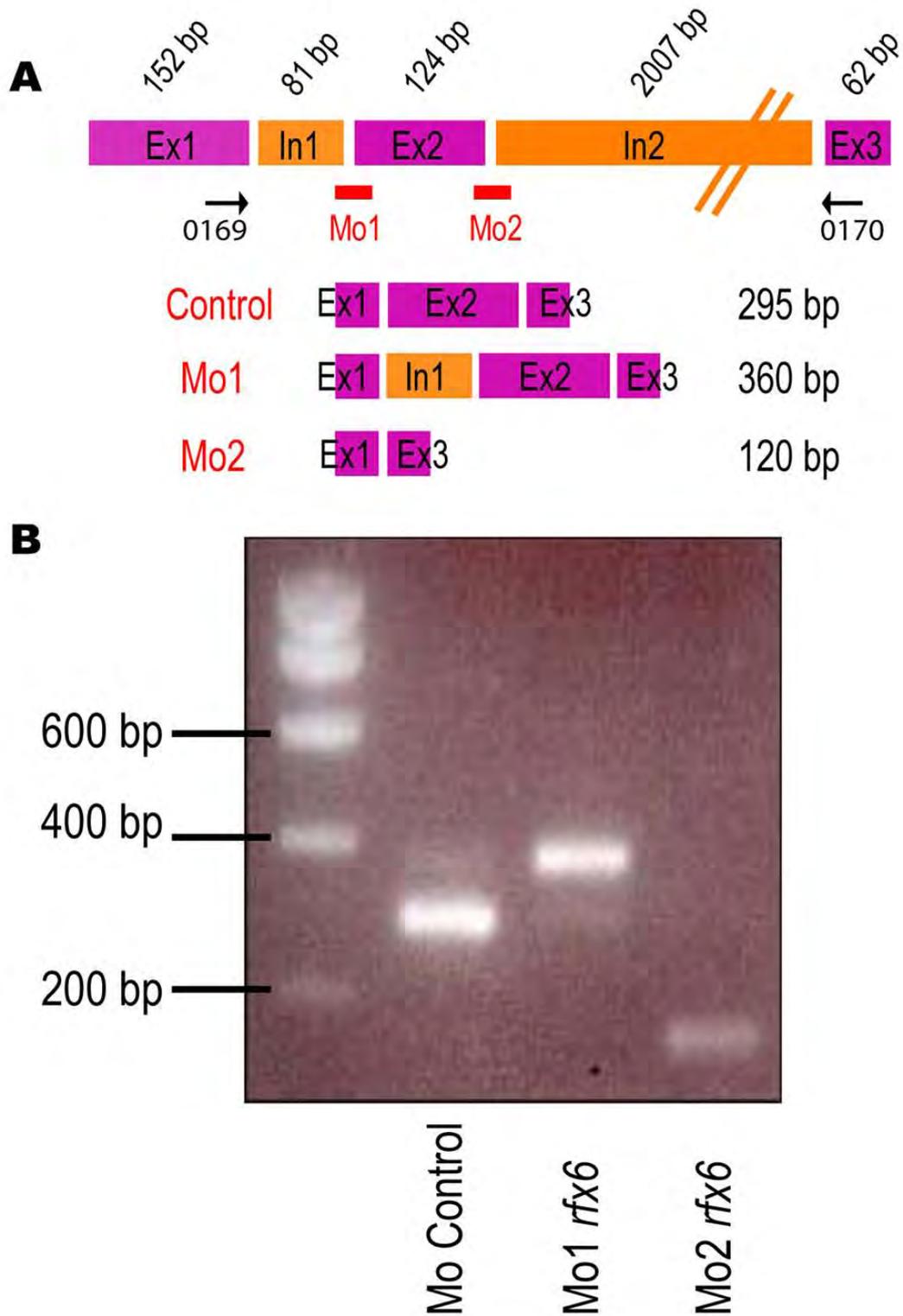


Fig. S5. Targeting efficiency of *rfx6* morpholinos. (A) Partial genomic structure of the *rfx6* gene. Splice sites targeted by morpholinos (MOs) are shown in red. The arrows indicate the location of the primers used for PCR (0169: CATCAAGGACAAGAAGAAGCAGAC and 0170: TTCCCGGAGTAAACAGAGTGATAA). (B) RT-PCR analysis of *rfx6* mRNA structure in the control (lane 1) and morphants (lanes 2 and 3) using 0169 and 0170. In the MO1 morphants, we can observe the apparition of a 360 bp band corresponding to a transcript containing intron1. In the MO2 morphants, the amplicon is reduced in size (120 bp), indicating a deletion of exon2. The identity of the amplicons has been confirmed by sequencing.

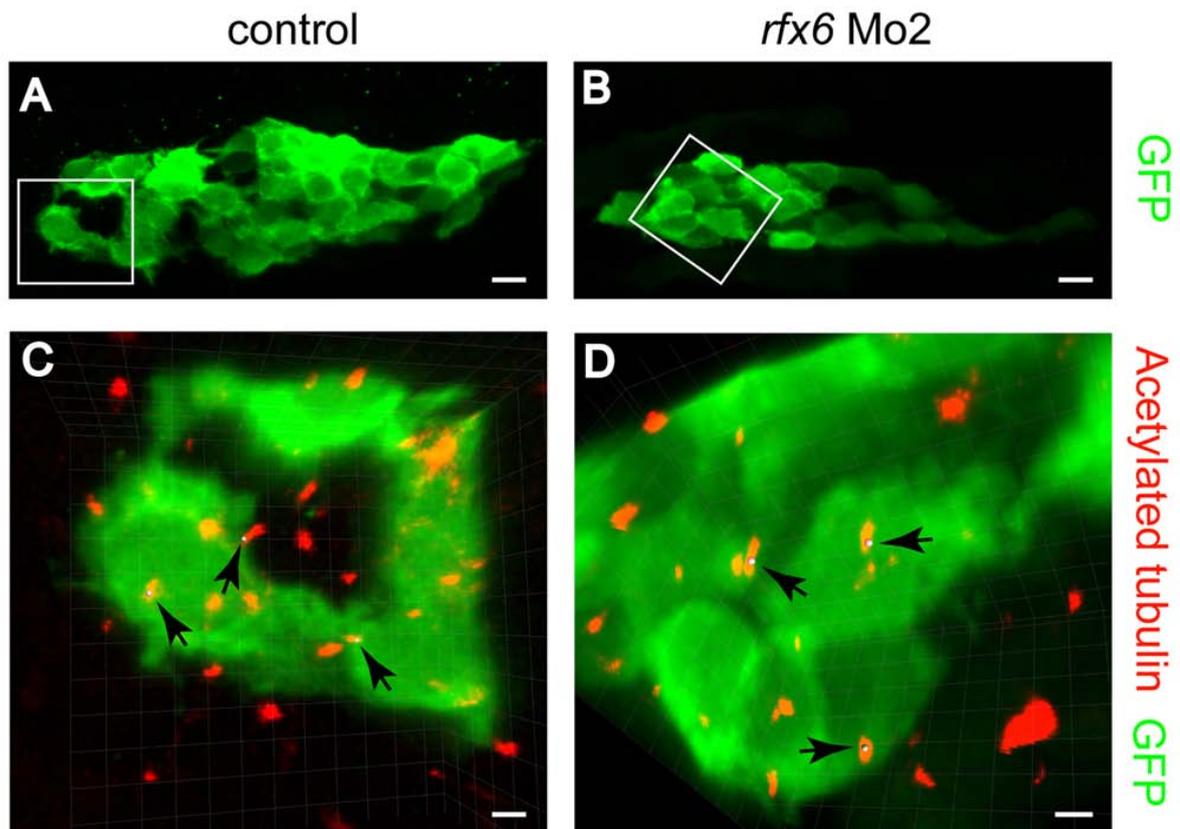


Fig. S6. Acetylated-tubulin labelling reveals ciliated cells within the developing pancreas of wild-type and morphant zebrafish. Experiments to visualize primary cilia were performed on transgenic zebrafish embryos *tg(pax6:GFP)* where pancreatic endocrine cells can be detected by GFP expression (Delporte et al., 2008). (**A,B**) Pancreatic cell labelling using the *tg(pax6:GFP)* transgenic line underlines the developing pancreas in control (A) and *rfx6* Mo2 morphants (B). (**C,D**) Magnifications of the area in white boxes in A and B with the acetylated-tubulin-labelled cilia in red and *pax6:GFP* in green. Arrows in C and D point to cilia. The same cilia are marked with a grey dot visible in Movies 1 and 2. Scale bars: 10 mm in A,B; 2 mm in C,D.

Delporte, F. M., Pasque, V., Devos, N., Manifroid, I., Voz, M. L., Motte, P., Biemar, F., Martial, J. A. and Peers, B. (2008). Expression of zebrafish *pax6b* in pancreas is regulated by two enhancers containing highly conserved cis-elements bound by PDX1, PBX and PREP factors. *BMC Dev. Biol.* **8**, 1-19.

Table S1. Transcription factors or transcriptional regulators upregulated in Ngn3-positive islet progenitor cells

AffyID	Gene symbol	Gene title	FC	FDR
1436092_at	<i>Mafa</i>	Transcribed locus, moderately similar to XP_574723.1 PREDICTED: similar to LRRGT00097 (<i>Rattus norvegicus</i>)	386.55	0.007
1419271_at	<i>Pax6</i>	paired box gene 6	275.2	0.008
1426413_at	<i>Neurod1</i>	neurogenic differentiation 1	171.99	0.011
1425886_at	<i>Fev</i>	FEV (ETS oncogene family)	139.22	0.011
1455123_at	<i>St18</i>	suppression of tumorigenicity 18	130.74	0.009
1450042_at	<i>Arx</i>	aristaless related homeobox gene (<i>Drosophila</i>)	119	0.011
1451598_at	<i>Pax4</i>	paired box gene 4	87.63	0.011
1419185_a_at	<i>Mlxipl</i>	MLX interacting protein-like	75.37	0.012
1457613_at	<i>Rfxdc1/Rfx6</i>	regulatory factor X domain containing 1	65.73	0.011
1421399_at	<i>Insm1</i>	insulinoma-associated 1	61.65	0.010
1422773_at	<i>Myt1</i>	myelin transcription factor 1	53.62	0.010
1422165_at	<i>Pou3f4</i>	POU domain, class 3, transcription factor 4	52.83	0.012
1418176_at	<i>Vdr</i>	vitamin D receptor	51.53	0.012
1447174_at	<i>Dach1</i>	Dachshund 1 (<i>Drosophila</i>)	49.16	0.029
1450428_at	<i>Lhx1</i>	LIM homeobox protein 1	42.43	0.008
1451716_at	<i>Mafb</i>	v-maf musculoaponeurotic fibrosarcoma oncogene family, protein B (avian)	38.73	0.012
1426298_at	<i>Irx2</i>	Iroquois related homeobox 2 (<i>Drosophila</i>)	37.8	0.012
1460299_at	<i>Hlxb9</i>	homeobox gene HB9	34.41	0.012
1429573_at	<i>Dmrtc1a</i>	DMRT-like family C1a	30.05	0.012
1436600_at	<i>Tnrc9</i>	trinucleotide repeat containing 9	29.82	0.012
1432034_at	<i>Neurog3</i>	neurogenin 3	28.41	0.012
1422720_at	<i>Isl1</i>	ISL1 transcription factor, LIM/homeodomain	28.26	0.022
1425828_at	<i>Nkx6.1</i>	NK6 transcription factor related, locus 1 (<i>Drosophila</i>)	24.18	0.036
1421112_at	<i>Nkx2.2</i>	NK2 transcription factor related, locus 2 (<i>Drosophila</i>)	23.87	0.013
1449967_at	<i>Sim1</i>	single-minded homolog 1 (<i>Drosophila</i>)	15.11	0.020
1422174_at	<i>Ipf1</i>	insulin promoter factor 1, homeodomain transcription factor	12.36	0.032
1419224_at	<i>Cecr6</i>	cat eye syndrome chromosome region, candidate 6 homolog (human)	11.22	0.023
1440439_at	<i>AI591476</i>	Expressed sequence AI591476	10.99	0.039
1441484_at	<i>Tcf2</i>	Transcription factor 2	10.75	0.043
1448272_at	<i>Btg2</i>	B-cell translocation gene 2, anti-proliferative	10	0.030
1429269_at	<i>BC068157</i>	cDNA sequence BC068157	9.46	0.041
1440870_at	<i>Prdm16</i>	PR domain containing 16	7.23	0.045
1448928_at	<i>Hdac6 III LOC669168</i>	histone deacetylase 6 III similar to Histone deacetylase 6 (HD6) (Histone deacetylase mHDA2)	6.89	0.043
1430353_at	<i>Glis3</i>	GLIS family zinc finger 3	6.88	0.045
1436483_at	<i>Myt1l</i>	myelin transcription factor 1-like	6.54	0.029
1433894_at	<i>AI591476</i>	expressed sequence AI591476	6.53	0.047
1418582_at	<i>Cbfa2t3h</i>	core-binding factor, runt domain, alpha subunit 2, translocated to, 3 homolog (human)	6.45	0.020
1429633_at	<i>Lcor</i>	ligand dependent nuclear receptor corepressor	6.36	0.043
1434777_at	<i>Mycl1</i>	v-myc myelocytomatosis viral oncogene homolog 1, lung carcinoma derived (avian)	5.92	0.036
1437894_at	<i>Prox1</i>	prospero-related homeobox 1	5.84	0.042
1456824_at	<i>Zfp612</i>	zinc finger protein 612	5.52	0.047
1416983_s_at	<i>Foxo1</i>	forkhead box O1	5.19	0.032
1456021_at	<i>Atf6</i>	activating transcription factor 6	4.74	0.045
1450684_at	<i>Etv1</i>	ets variant gene 1	4.49	0.023
1435236_at	<i>A630018P17Rik</i>	RIKEN cDNA A630018P17 gene	4.23	0.017
1455944_at	<i>Zfp516</i>	zinc finger protein 516	3.33	0.048
1425484_at	<i>Tox</i>	thymocyte selection-associated HMG box gene	2.57	0.044

Comparison of gene expression profiles in Ngn3-positive versus Ngn3-negative cells (E15.5 pancreas). Probe sets found upregulated in Ngn3-positive cells with an FDR>5% were selected and transcriptional regulators were identified using GO annotations (biological process and molecular function) containing transcription. Results are sorted by fold-change (FC). *Mafa* and *Insm1* were hand-annotated as transcription factors. Excluded transcriptional regulators (FDR>5% and <10%) include: *Atf6*, *Rfx3*, *Dach2*, *Zfp568*, *Hhex*, *Zfp354b*, *Tshz1*.

Movie 1. 3D visualization of cilia at the pancreatic cell surface in control embryos at 24 hpf. Acetylated-tubulin-labelled cilia are in red and the pancreatic cell using *pax6:GFP*-labelling is in green. Cilia are marked with a grey dot. Scale box: 2 µm.

Movie 2. 3D visualization of cilia at the pancreatic cell surface in *rfx6* MO2 morphants at 24 hpf. Acetylated-tubulin-labelled cilia are in red and the pancreatic cell using *pax6:GFP*-labelling is in green. Cilia are marked with a grey dot. Scale box: 2 µm.

Summary

The endocrine (islet) cells of the pancreas include alpha and beta cells which secrete the hormones glucagon and insulin respectively, key regulators of glycaemia. During embryogenesis, the transcription factor Neurogenin3 (Ngn3 or Neurog3) controls endocrine fate specification in multipotent pancreatic progenitor cells in the mouse. In *Neurog3* mutant mice, all pancreatic endocrine cells are lost and mice die from diabetes. Despite these findings, our knowledge of the endocrinogenic programs implemented by Ngn3 is still fragmentary. To gain further insight into the Ngn3-dependent regulatory network we determined the transcriptome of purified Ngn3-positive cells, of Ngn3 mutant pancreas and of the genes induced by an ectopic expression of Ngn3 in pancreas. Thus, we had generated a comprehensive list of potential new target genes of Ngn3. We studied one of this gene, the winged helix transcription factor Rfx6, in mouse and zebrafish. In the mouse, Rfx6 is initially found in the gut endoderm and then in post-mitotic islet progenitor cells in the embryo and is maintained in hormone-expressing islet cells. *Rfx6* is dependent on Ngn3 and acts upstream of or in parallel with *NeuroD*, *Pax4*, and *Arx* transcription factors during islet cell differentiation. Loss of function studies in zebrafish (collaboration with the group of M. Voz, University of Liege) revealed that *rfx6* is required for the differentiation of glucagon, ghrelin and somatostatin expressing cells which, are blocked at the progenitor stage. In contrast, beta cells, whose number is only slightly reduced, were no longer clustered in a compact islet. This study unveils Rfx6 as a novel regulator of islet cell development.

Résumé

Les cellules endocrines du pancréas incluent les cellules alpha et beta qui sécrètent le glucagon et l'insuline, deux hormones qui régulent la glycémie. Au cours de l'embryogenèse de la souris, c'est le facteur de transcription Neurogenin3 (Ngn3 ou Neurog3) qui spécifie la destinée endocrine des progéniteurs pancréatiques. En effet, en absence de Ngn3 les cellules endocrines ne se développent pas et les souris meurent de diabète. Afin d'identifier les programmes génétiques qui contrôlent la différenciation endocrine en aval de *Ngn3*, nous avons déterminé le transcriptome des cellules Ngn3-positives, des mutants *Ngn3*, et identifié les gènes induits par une expression ectopique de Ngn3. Cette étude a généré une liste de gènes cibles potentiels de Ngn3 incluant le facteur de transcription Rfx6 qui a été étudié plus avant. Chez la souris, Rfx6 est initialement exprimé dans l'endoderme intestinal embryonnaire puis restreint aux cellules endocrines du pancréas en développement et adulte. Rfx6 dépend de Ngn3 et est donc spécifique du lignage endocrine où il agit en amont ou en parallèle de *NeuroD*, *Pax4* et *Arx*, des régulateurs importants de la différenciation endocrine. L'étude de la perte de fonction chez le poisson zèbre (collaboration avec l'équipe de M. Voz, Université de Liège) a révélé que Rfx6 est requis pour la différenciation des cellules exprimant le glucagon, la ghrelin et la somatostatine, qui sont bloquées au stade de progéniteurs. En revanche le nombre de cellules exprimant l'insuline ne diminue que faiblement mais les cellules ne s'agrègent plus en îlots. En conclusion, ce travail a dévoilé Rfx6 comme un nouveau régulateur de la différenciation endocrine du pancréas.