



THESE

Pour l'obtention du grade de
Docteur en Sciences de l'Université de Strasbourg
Discipline : Sciences de la vie
Spécialité : Aspects moléculaires et cellulaires de la biologie

Présentée par
Perrine STRASSER

Rôle du facteur de transcription RFX6 dans la différenciation et la fonction des cellules β sécrétrices d'insuline : identification et étude de gènes cibles

Soutenue publiquement le 28 septembre 2015

Directeur de thèse :	Gérard Gradwohl – DR INSERM – Strasbourg
Rapporteurs externes :	Patrick Collombat – DR INSERM – Nice Catherine Postic – DR INSERM - Paris
Rapporteur interne :	Maria Elena Torres Padilla – DR INSERM - Strasbourg

A toutes celles et ceux qui sont partis trop tôt,

Remerciements

Gérard, ton encadrement au cours des dernières années m'a permis d'apprendre beaucoup. Tu représentes le pilier parfait pour un doctorant, et je te remercie d'avoir su me guider au cours de cette épreuve qu'est la thèse. Ta présence sans faille, tes conseils et encouragements et surtout ta grande humanité me laisseront un souvenir impérissable de cette aventure. Merci.

Je remercie Maria Elena Torres Padilla, Catherine Postic et Patrick Collombat d'avoir accepté de juger ce travail.

Je remercie également la Fondation pour la Recherche Médicale et l'Association de la Recherche contre le Cancer qui ont soutenu financièrement ce travail pendant 4 ans.

Je remercie aussi tous les anciens membres de l'équipe que j'ai croisés pendant ces 4 années et même un peu plus.

Martine, tu n'as pas été là pendant toute ma thèse et quel dommage ! Ta bonne humeur et ta grande générosité m'ont permis de démarrer dans l'équipe dans les meilleures conditions.

Anthony, merci de m'avoir lancée sur la piste du CHIP. Travailler à tes côtés au démarrage de mon doctorat fut un plaisir et je te souhaite toute la réussite que tu espères professionnellement.

Julie, on a parcouru un peu plus de chemin ensemble dans l'équipe GG et formé une belle équipe ! Très belle continuation à toi en Suisse.

Laure, j'ai démarré ma thèse juste après toi depuis un peu plus de 4 ans. Je te souhaite bon vent dans tes projets futurs !

Aline, j'espère que toutes les personnes qui croiseront ton chemin professionnel se rendront compte de la chance qu'ils ont ! Pour ma part c'est chose faite, et au-delà du fait d'être une collègue exemplaire, tu as su être présente exactement quand il le fallait en tant qu'amie. Merci.

Mégane, Emélie et Julia, si nous en sommes là aujourd'hui c'est aussi grâce à vous. Merci pour votre travail sur le projet huRfx6, et merci de ne pas avoir été découragées par tous les obstacles rencontrés lors de vos stages ! En espérant avoir pu vous apprendre quelques belles choses de la vie en laboratoire. Vous encadrer a été très enrichissant pour moi.

Jean-Marie, merci pour toute ton aide lors de mon doctorat ! Tu fais figure de pilier dans ce cher institut et je suis heureuse d'avoir pu me perfectionner à tes côtés. Bienvenue à toi dans l'équipe GG !

Je tiens à remercier chaleureusement toutes les personnes de la plateforme de Biopuces et Séquençage de l'IGBMC et particulièrement Bernard, Serge, Muriel, et Stéphanie qui ont grandement participé à l'obtention de tous mes résultats.

Tao, cette thèse ne serait pas ce qu'elle est sans ton aide incommensurable pour l'analyse bio-informatique de mes résultats de CHIP. Merci beaucoup pour ta patience.

Merci aussi à Jean-Christophe Ame pour son aide précieuse du côté de l'ESBS.

En parallèle de mon travail de thèse, j'ai eu la chance de découvrir la vie associative aux côtés des membres de l'Association des étudiants et post-doctorants de l'IGBMC. Je vous remercie pour toutes ces belles années.

J'aurai un petit mot particulier pour l'équipe de choc des italiens du premier Sud : Valerio et Simone, merci pour votre bonne humeur, pour ces belles pauses café et pour vos encouragements dans la dernière ligne droite.

David, au-delà d'avoir repris l'association en main à mon départ, tu es un personnage hors du commun et je suis ravie d'avoir pu apprendre à te connaître pendant ma thèse. Bon courage pour ta dernière année ! Et surtout, bienvenue à la Robertsau pour la suite de ton chemin, tu verras, quand on y met un pied, on y reste !

Je remercie aussi en particulier l'équipe du Forum Biotechno 2013 au complet pour cette belle aventure. Grâce à vous, cette journée du 07 juin 2013 aura été une belle réussite !

Yasmine, que d'encouragements ces derniers mois, toujours au bon moment et avec les mots qu'il faut ! C'est un petit bonheur de t'avoir rencontrée il y a déjà quelques années. Merci pour tout de tout cœur et, au fond, c'est toi la championne !

Ahmad, à défaut de me souvenir qu'on était dans la même promo à la fac, sois sûr que je garderai en mémoire nos thèses et tous ces bons moments qu'on a partagés grâce à ton éternelle bonne humeur ! C'est bientôt ton tour !

Arnaud, attention plus que 2 ans devant toi ! Il est où ton papier de première année ? Je te souhaite plein de courage pour le reste de ta thèse, il en faut ! Et merci pour les petites soirées improvisées chez toi, je resterai dans le coin juste pour ça.

Les filles, des pages entières ne suffiraient pas pour vous écrire que ma thèse n'aurait pas eu la même saveur sans notre Pupici team.

Laura, merci pour tes conseils avisés de post-doc, ton calme à toute épreuve et je te souhaite une très belle continuation loin d'ici ! On attend les mini-pupici avec Claire, alors au boulot les Parlog !!

Claire, ma sœur de fac et de thèse. Qui aurait pensé que 8 ans après notre rencontre sur les bancs de la fac, on en serait là ! Merci de m'avoir accompagnée tout au long de ce parcours et de ces épreuves. Courage pour la fin, c'est le pire, mais ça vaut le coup.

Les Pupici, merci pour cette belle amitié sans faille qui ne fait que commencer... 5 ans déjà ! Je vous aime les filles !

Ismail, toi tu es ce que j'appelle le bonus de ma thèse ! Tu m'as poussée à découvrir la vie associative, m'a entraînée à manager le Forum BioTechno 2013, puis amenée à la tête du SPB, rien que ça ! Grâce à toi, j'ai pu apprendre que dans la vie, on n'est jamais à l'abri d'une très très belle surprise. Je compte sur toi pour me faire redécouvrir la vie au Maroc, je l'attends ce voyage.

Les copains de toujours, je n'arriverai jamais à tous vous citer ici, mais vous savez que le cœur y est. Vous m'avez permis de garder un pied dans le monde non scientifique ! Sophie et Christophe, Chouchou et Marie, Laure et Bart, Picard et Alba, Najjar et Laurianne, Claire, Philippe et votre petit Alexandre, Stéphane et ta famille bientôt aggrandie, et enfin Lauranne et Anto, vous êtes tous tout simplement géniaux ! Je souhaite à tout le monde d'avoir dans son cœur des amis comme vous.

Anne, Aurélie, Sophie et Marion, merci pour ces superbes soirées filles à débattre de nos histoires d'école et de comment on va organiser cette belle rencontre avec tous les anciens !

Valou, Tristouille, Minus et toutes les belles personnes qui vous entourent dans votre fabuleux groupe d'amis : merci à tous de m'avoir accueillie parmi vous et surtout de m'avoir montré, après toutes ces années, qu'en pharma il y a quand même des gens bien ! Un petit mot particulier pour toi Propre, je suis de tout cœur avec toi pour ta grande thèse, fonce !

Tout ça ne serait évidemment rien sans une famille formidable et unie ! Alors merci à Gabrielle et Eric, Lucas et Sarah, Julien, Marie-Hélène, Kévin et Gérard.

Les Vallée, merci de m'avoir fait grandir dans un environnement familial si agréable. Julien et Lucas, vous êtes comme des frères pour moi.

Fanny et Margot, vous êtes des merveilles de petites sœurs ! La vraie vie démarre tout doucement pour vous et je vous souhaite de pouvoir réaliser tous vos rêves. Fanny, tu démarres sur les chapeaux de roues en Afrique du Sud ! Margot, plus qu'un pas avant le BAC et ce ne sera que du bonheur !

Papa, tes précieux conseils professionnels et personnels et ton soutien m'ont appris à ne jamais baisser les bras. Tu es un papa formidable par ta présence, ta disponibilité et ton écoute. Je t'aime.

Maman, tu es une maman comme tout le monde rêve d'en avoir une ! Toujours pleine d'énergie, avec une idée à la seconde, qui m'a soutenue dans mes choix et tout au long de mes études. Merci pour ton aide avec ce manuscrit et pour la vie que tu m'offres. Je t'aime.

Et enfin, Thomas, que peut-on dire en seulement quelques mots à la personne qui partage ma vie depuis tant d'années ? Merci pour ton insouciance, ta présence, ta patience, ton écoute et ton amour. Je t'aime et puis c'est tout.

Merci aussi à toutes celles et ceux que j'aurais pu oublier.

A l'image de ces longs remerciements, cette thèse a été pour moi avant tout une aventure humaine incroyable et vous êtes tous, sans exception, des personnes pour qui des remerciements ne suffiront jamais à exprimer à quel point je vous aime !

Abréviations

3HA	triple épitope de l'hémagglutinine du virus de la grippe humaine
Acc1	acetyl-CoA Carboxylase alpha
ADN	acide désoxyribonucléique
ADNc	ADN complémentaire
ADP	adénosine di-phosphate
Amp	ampicilline
AMPc	adénosine mono-phosphate cyclique
AMPK	AMPc-activated protein kinase
ARN/RNA	acide ribonucléique
ARNT	aryl hydrocarbon receptor nuclear translocator
Arx	aristaless related homeobox
ATP	adénosine tri-phosphate
BAC	bacterial artificial chromosome
bHLH	basic helix-loop-helix
CAS9	CrispR associated protein 9
ChIP	immunoprécipitation de chromatine
ChoRE	carbohydrate response element
ChREBP	carbohydrate response element binding protein
CrispR	clustered regularly interspersed palindromic repeats
DBD	DNA binding domain
DID	diabète insulino-dépendant
DIM	dimerization domain
DMT	diabète mitochondrial
DNID	diabète non insulino-dépendant

Dync2li1	dynein, cytoplasmic 2, light Intermediate chain 1
E2F1	E2F Transcription Factor 1
ELISA	enzyme-linked immunosorbent assay
ESC	cellules embryonnaires souches
FACS	fluorescence-activated cell sorting
Fas	fatty acid synthase
FC	fold change
Ffar1	free fatty acid receptor 1
FGF10	fibroblast growth factor 10
FIMO	find individual motif occurrences
Flp	recombinase
Fox	forkhead box
FRT	flippase recognition target
FT	facteur de transcription
G6P	glucose-6-phosphate
Gcg	glucagon
Gck	glucokinase
GEO	gene expression omnibus
GFP	green fluorescent protein
GIP	gastric inhibitory polypeptide
GIPR	gastric inhibitory polypeptide receptor
GLP-1	glucagon like peptide1
Glut	facilitated glucose transporteur
GRACE	glucose response activation conserved element
GWAS	genome wide association study
HBV	virus de l'hépatite B

HEK293T	human embryonic kidney 293 cells
HGF	hepatic growth factor
HIF1	hypoxia-inducible factor 1
HLA	human leucocyte antigen
Hmgn3	high mobility group nucleosomal binding domain 3
Hnf4	hepatocyte nuclear factor 4
IF	immunofluorescence
IFT	intraflagellar transport genes
IGF	insulin growth factor
Ins	insuline
Insm1	insulinoma-associated 1
IP	immunoprécipitation
IPA	ingenuity pathway analysis
IPSC	cellules souche à pluripotence induite
Irs	insulin receptor substrate
Irx	iroquois-class homeodomain protein
kb	kilo base
kDa	kilo Dalton
KO	knock-out
LID	low glucose inhibitory domain
Lpk	Liver-pyruvate kinase
LXR/RXR	liver/retinoid X Receptor
m/v ou v/v	masse/volume ou volume/volume
MACS	model-based analysis for chip-seq
MafA	v-maf avian musculoaponeurotic fibrosarcoma oncogene homolog A
MLX	MAX dimerization protein

MIxipl	MLX interacting protein-like
MODY	maturity onset diabetes of the young
Myt1	myelin transcription factor 1
n/a	not applicable
nan	nanchung
ND	neonatal diabetes
Néo	néomycine
NeuroD1	neurogenic differentiation1
Ngn3	neurogenin3
NLS	signal de localisation nucléaire
NRF1	nuclear respiratory factor 1
nt	nucléotides
OCT-1	octamer transcription factor 1
Pax	paired box
Pb	paire de base
PBS	phosphate buffer saline
PC1/3	proconvertase 1/3
PCR	polymerase chain reaction
Pdx1	Pancreas and duodenal homeobox1
Per	period circadian clock
PKA	cAMP-dependent protein kinase
PNDM	permanent neonatal diabetes mellitus
PP	polypeptide pancréatique
PP2A	protein phosphatase 2A
PPAR	peroxisome proliferator-activated receptor
Pparg	peroxisome proliferator activated receptor gamma

Ptf1a	pancreas transcription factor 1 subunit alpha
qPCR	PCR quantitative
RFP	red fluorescent protein
Rfx	regulatory factor X
SCO	subcommissural organ
Seq	sequencing
Ser	sérine
Slc30a8	solute carrier family 30 (zinc transporter) member 8
Som	somatostatine
Sp	specificity protein
Synd	syndromic diabetes
TF	transcription factor
Thr	thréonine
TNDM	transient neonatal diabetes mellitus
TSS	sites d'initiation de la transcription des gènes
TTIV	transcription traduction in vitro
TTS	site de terminaison de la transcription
Txnip	thioredoxin interacting protein
UCSC	university of california santa cruz
UTR	untranslated region
WB	western blot
WBSCR14	williams beuren syndrome critical region 14
WT	sauvage
X5P	xylulose-5-phosphate
XYLB	xylulokinase

Sommaire

Introduction

A. Anatomie et fonction du pancréas	1
A.1. Organe, histologie et fonctions cellulaires	1
A.2. La cellule beta sécrétrice d'insuline	3
B. Le diabète, une pandémie mondiale	6
B.1. Le diabète insulino-dépendant.....	7
B.2. Le diabète non insulino-dépendant.....	7
B.3. Le diabète gestationnel.....	7
B.4. Les diabètes monogéniques.....	8
B.5. Axes de recherches et thérapies cellulaires pour le diabète	8
C. Développement du pancréas et régulation transcriptionnelle	11
C.1. Spécification pancréatique	12
C.2. Spécification endocrine	13
C.3. Spécification des sous-types endocriniens.....	14
D. Le facteur de transcription RFX6.....	17
D.1. La famille RFX.....	17
D.2. Fonctions diverses des membres de la famille.....	18
D.3. RFX6 et le pancréas.....	21

Matériels et méthodes

A. Matériel biologique	24
A.1. Culture et transfection cellulaire	24
A.1.a. Cellules Min6b1.....	24
A.1.b. Cellules HEK293T.....	24
A.1.c. Transfection cellulaire	24
A.2. Purification des ilots de Langerhans.....	25

B. Immunoprécipitation de chromatine	26
B.1. Immunoprécipitation de chromatine à partir de cellules Min6b1	26
<i>B.1.a. Préparation de chromatine</i>	<i>26</i>
<i>B.1.b. Immunoprécipitation</i>	<i>27</i>
<i>B.1.c. Purification de l'ADN immunoprécipité</i>	<i>28</i>
B.2. Immunoprécipitation de chromatine à partir d'îlots de Langerhans de souris ..	28
<i>B.2.a. Préparation de chromatine</i>	<i>28</i>
<i>B.2.b. Immunoprécipitation</i>	<i>29</i>
<i>B.2.c. Purification de l'ADN immunoprécipité</i>	<i>30</i>
C. Analyses bio-informatiques	30
C.1. Préparation des bibliothèques d'ADN et séquençage	30
C.2. Identification et annotation des pics de ChIP Seq	31
C.3. Recherche de motifs de séquences	32
C.4. Analyse d'occurrence et de co-occurrence de motifs	32
D. Test fonctionnels	33
D.1. Tests de transactivation	33
D.2. Tests de gel retard	37
E. Oligonucléotides et anticorps	40

Résultats

A. Recherche des gènes cibles du facteur de transcription RFX6	43
A.1. Clonage, expression et immunoprécipitation de la protéine de fusion Flag-3HA-RFX6	43
A.2. Développement d'une souris transgénique 3HA-RFX6-mCherry	47
<i>A.2.a. Stratégie de clonage</i>	<i>47</i>
<i>A.2.b. Obtention d'une souris transgénique et expression du transgène 3HA-Rfx6-mCherry</i>	<i>57</i>
A.3. Recherche des gènes cibles de RFX6 dans la lignée cellulaire beta Min6b1 ..	59
<i>A.3.a. Stratégie expérimentale</i>	<i>59</i>

A.3.b. <i>Identification et analyse des sites de liaison de RFX6 dans les cellules Min6b1</i>	61
A.3.c. <i>Identification des gènes ciblés dans les cellules Min6b1 et régulés dans les cellules beta pancréatiques par RFX6</i>	67
A.3.d. <i>Analyse d'occurrence et de co-occurrence de motifs dans les cibles de RFX6 dans les cellules Min6b1</i>	79
A.3.e. <i>Rfx6 maintains the functional identity of adult pancreatic beta cells / Piccand et al. 2014, Cell Reports</i>	87
A.4. Recherche des gènes cibles de RFX6 dans les ilots de Langerhans de souris	88
A.4.a. <i>Stratégie expérimentale</i>	88
A.4.b. <i>Identification et analyse des sites de liaison de RFX6 dans les ilots pancréatiques de souris</i>	88
A.4.c. <i>Analyse d'occurrence et de co-occurrence de motifs dans les cibles de RFX6 dans les ilots pancréatiques de souris</i>	91
A.4.d. <i>Identification des gènes ciblés dans les ilots pancréatiques et régulés dans la cellule beta par RFX6</i>	96
A.4.e. <i>Détermination de réseaux fonctionnels parmi les gènes ciblés dans les ilots pancréatiques et régulés dans la cellule beta par RFX6</i>	106
B. Mlxipl est une cible principale de RFX6 dans la cellule beta adulte	113
B.1. Mlxipl est exprimé dans différents organes	113
B.2. Régulation de l'activité de MLXIPL en fonction du glucose	114
B.3. Rôle de MLXIPL dans le métabolisme des lipides et du glucose	116
B.4. MLXIPL et le pancréas	118
B.5. RFX6 lie le premier intron de Mlxipl.....	121
B.6. Rfx6 régule positivement l'expression de Mlxipl dans la cellule beta adulte .	125
B.7. La Xbox2 de Mlxipl promeut sa transactivation par la protéine RFX6 de souris et humaine	126
B.8. Les mutations ponctuelles humaines de Rfx6 affectent la fonction de transactivation du gène Mlxipl	128
B.8.a. <i>Clonage des gènes Rfx6 humains mutés</i>	130
B.8.b. <i>Expression des protéines RFX6 humaines mutées</i>	132

B.9. Test de liaison à l'ADN entre RFX6 et son site de liaison dans l'intron 1 de Mlxipl.....	133
<i>B.9.a. Expression in vitro de la protéine RFX6 humaine</i>	133
<i>B.9.b. Mise au point des tests de gel retard</i>	137
B.10. Etude du lien entre la stimulation de la sécrétion de l'insuline en réponse au glucose et le couple Rfx6/Mlxipl	140
B.11. Étude du transport cytoplasme/noyau de MLXIPL et RFX6 en fonction de la concentration en glucose.....	141
B.12. Étude de la prolifération des cellules Min6b1 en fonction de la concentration en glucose	144

Discussion et perspectives

A. Rfx6 est un modulateur clé de la sécrétion d'insuline et du maintien de l'homéostasie du glucose	147
A.1. RFX6 cible des gènes essentiels à la sécrétion d'insuline et à l'homéostasie du glucose dans les ilots pancréatiques	147
A.2. Les cibles de RFX6 sont déterminées <i>in vitro</i> et <i>in vivo</i>	148
A.3. RFX6 et ses partenaires	150
B. RFX6 régule directement l'expression de Mlxipl dans la cellule beta pancréatique	152
B.1. Contrôle de la fonctionnalité des protéines RFX6 humaines mutantes	153
B.2. Les gènes régulés par Mlxipl contribuent-ils au phénotype de la souris delta beta ?	154
C. RFX6 et ses perspectives	155
C.1. RFX6 et cancer.....	155
C.2. RFX6 dans l'intestin.....	156

Introduction

Introduction

A. Anatomie et fonction du pancréas

A.1. Organe, histologie et fonctions cellulaires

Le pancréas est un organe abdominal situé derrière l'estomac (**Figure 1**). Il est qualifié de glande mixte puisqu'il assure des fonctions exocrines et endocrines.

Le pancréas exocrine comprend les cellules acinaires et canalaire et représente 90 à 95% du pancréas (Edlund, 2002). Les cellules acinaires synthétisent le suc pancréatique composé d'ions HCO_3^- et d'enzymes pancréatiques (protéases, lipases, amylases) intervenant dans la digestion. La sécrétion de ces enzymes est régulée par des hormones gastriques et intestinales telles que la ghréline, la sécrétine ou encore la cholécystokinine (Murphy & Bloom, 2006). Les cellules canalaire tapissent les canaux pancréatiques qui transportent et déversent ces enzymes pancréatiques dans le canal de Wirsung puis dans la lumière du tube digestif au niveau du duodénum via la papille duodénale (Cleveland et al, 2012).

Le pancréas endocrine représente 1 à 2% de l'organe mature. Il est organisé en amas ou îlots de Langerhans qui sont constitués de cinq types cellulaires : les cellules alpha, beta, delta, PP et epsilon, qui sécrètent respectivement les hormones glucagon, insuline, somatostatine, polypeptide pancréatique et ghréline. Chez la souris, les cellules beta sont situées au centre de l'îlot de Langerhans et sont entourées d'une couronne de cellule alpha. Par contre, chez l'homme, les cellules alpha et beta sont réparties dans tout l'îlot (Bosco et al, 2010). Les cellules delta représentent 5 à 10% du pancréas, et sécrètent la somatostatine. Cette sécrétion est stimulée par le GLP-1 (Glucagon like peptide-1) et provoque l'inhibition de la sécrétion de glucagon par les cellules alpha (Drucker, 2013). Les cellules PP et epsilon représentent moins de 2% du pancréas. Le polypeptide pancréatique participe au contrôle de la satiété et la ghréline stimule l'appétit et inhibe la sécrétion d'insuline chez l'adulte lorsqu'elle est sécrétée par l'estomac et l'intestin (Chandra & Liddle, 2014a; Chandra & Liddle, 2014b; Wierup et al, 2013).

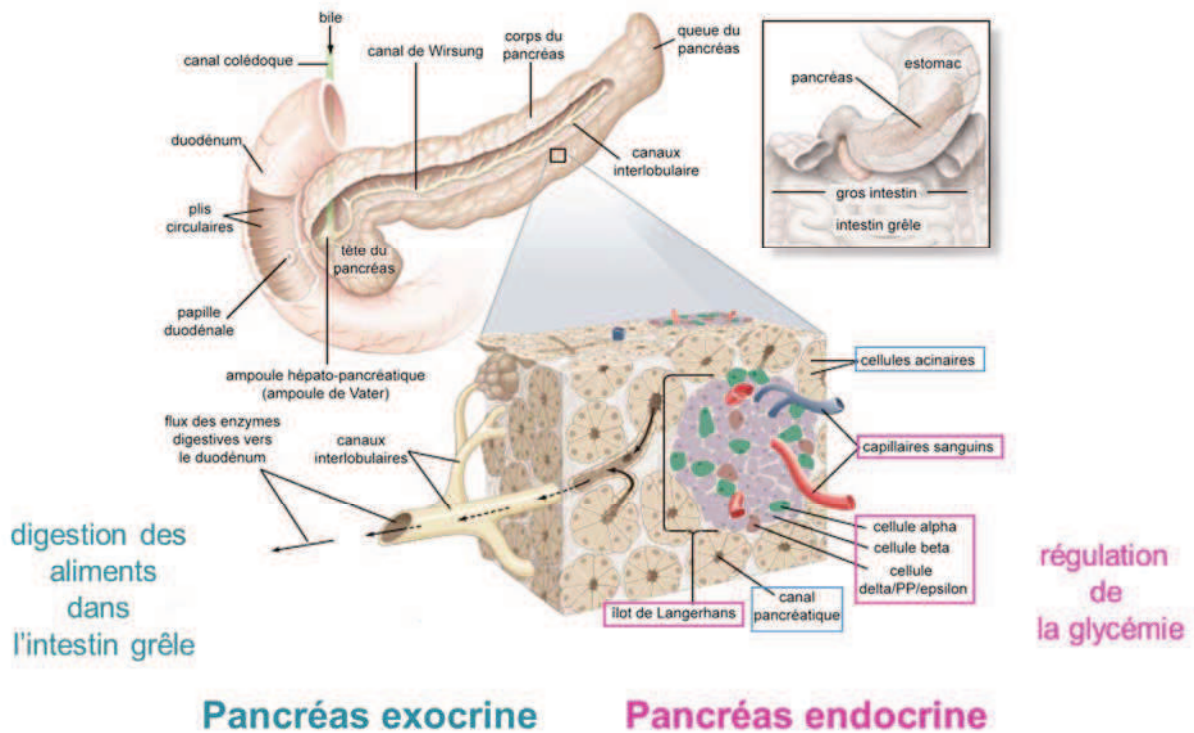


Figure 1 : Anatomie, histologie et fonctions du pancréas humain. Le pancréas, placé sous l'estomac, est divisé en trois parties (tête, corps et queue) et communique avec le duodénum au niveau de la papille duodénale. Le pancréas exocrine est constitué des cellules acinaires et canaliculaires. Le pancréas endocrine est composé des îlots de Langerhans et de capillaires sanguins. Ces îlots sont constitués de cinq sous-types cellulaires : les cellules alpha, beta, delta, PP et epsilon qui sécrètent respectivement le glucagon, l'insuline, la somatostatine, le polypeptide pancréatique et la ghréline. Le pancréas exocrine participe à la digestion des aliments qui a lieu dans l'intestin grêle et le pancréas endocrine permet la régulation de la glycémie. (Adapté d'Encyclopaedia Britannica, 2003.)

La fonction principale du pancréas endocrine est la régulation de la glycémie par la sécrétion d'insuline et de glucagon par les principales cellules endocrines du pancréas (respectivement 60 à 80% pour les cellules beta et 15 à 20% pour les cellules alpha). Ces hormones sont respectivement hypo- et hyperglycémiantes et sécrétées dans la circulation sanguine en réponse à des variations de la glycémie.

A.2. La cellule beta sécrétrice d'insuline

C'est par un équilibre hautement régulé de l'homéostasie du glucose que l'organisme fonctionne dans les conditions idéales (**Figure 2**). Pour cela, les cellules alpha sécrètent le glucagon, hormone hyperglycémiant, en réponse à une glycémie basse. Le glucagon stimule la production hépatique de glucose par glycogénolyse et néoglucogenèse et stimule la disponibilité du glucose dans les organes périphériques tels que le tissu adipeux ou les muscles par lipo- et protéolyse (Stephen L. Aronoff, 2004). D'autre part, l'effet inverse est observé en présence d'une glycémie haute lors de laquelle la sécrétion d'insuline hypoglycémiant par les cellules beta est stimulée. Dans ce cas, le foie assimile le glucose et le stocke sous forme de glycogène et les organes périphériques utilisent le glucose comme source d'énergie.

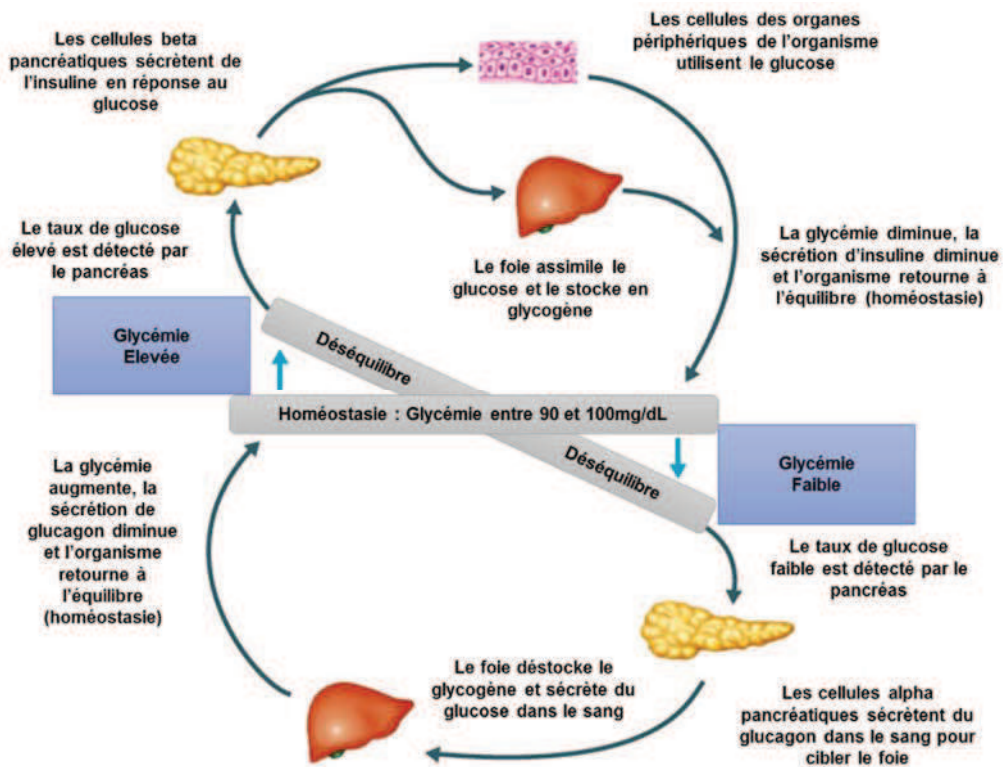


Figure 2 : Régulation de l'homéostasie du glucose par la sécrétion d'insuline et de glucagon. Une glycémie élevée stimule la sécrétion d'insuline par le pancréas. L'insuline induit l'assimilation du glucose par les organes périphériques et provoque la formation de glycogène par le foie. Inversement une glycémie faible entraîne la sécrétion de glucagon par le pancréas qui stimule alors la sécrétion de glucose dans le sang par le foie. Ceci permet le maintien de l'équilibre glycémique dans l'organisme entre 90 et 100 mg/dL de glucose dans le sang. (Adapté de Pearson Education, 2004)

La cellule beta pancréatique synthétise, stocke et sécrète l'insuline. Cette hormone constitue la seule hormone hypoglycémisante de l'organisme. Elle est constituée d'une chaîne C et de deux chaînes A et B de 21 et 30 acides aminés liées par deux ponts disulfures ainsi qu'un pont disulfure intra-chaîne A. L'insuline est tout d'abord synthétisée sous forme de pro-insuline puis est métabolisée par les prohormones convertases PC1/3 et 2 en insuline mature après élimination du peptide C. L'insuline est alors stockée dans des granules intracellulaires et sécrétée en cas de besoin.

L'expression du gène de l'insuline est régulée par différents facteurs de transcription dans la cellule beta. Les principaux sont indiqués dans la **Figure 3**. Ce sont notamment Pdx1, NeuroD1 (Beta2) et MafA qui par leur liaison synergique au niveau du promoteur du gène de l'insuline activent son expression (Melloul et al, 2002; Zhao et al, 2005). L'expression de Pdx1, NeuroD1 et MafA est en outre régulée par d'autres facteurs de transcription tels que HNF4 alpha et beta, FOXA2 ou PAX 6 (Cerf, 2006).

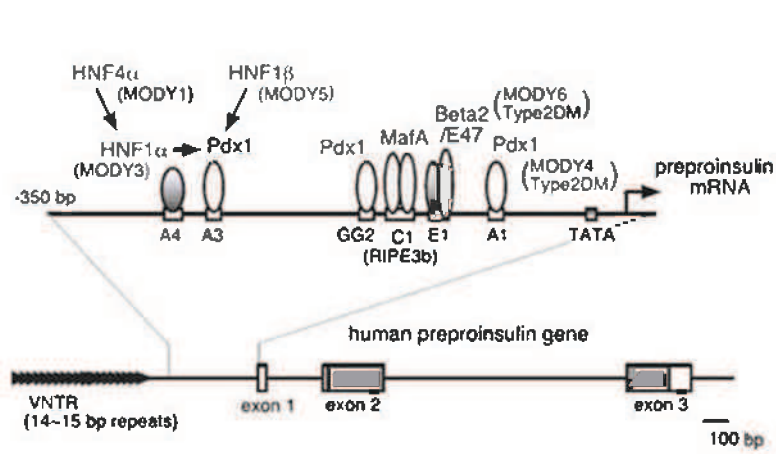


Figure 3 : Eléments régulateurs *cis* et *trans* du gène de l'insuline chez l'homme. MODY : maturity-onset diabetes of the young, Type2DM : type 2 diabetes mellitus. (Aramata et al, 2007)

Dans les cellules beta, l'assimilation du glucose, par le transporteur Glut2 et via les voies de glycolyse et de respiration cellulaire, induit une augmentation de l'ATP intracellulaire. Celle-ci conduit à la fermeture des canaux potassium puis à la dépolarisation membranaire. Une ouverture des canaux calciques de la cellule a alors lieu et permet la sécrétion d'insuline (**Figure 4**).

La sécrétion d'insuline a lieu en deux phases suite à l'ingestion d'aliments. Un premier pic de sécrétion de 5 à 7 minutes puis une sécrétion moins importante sur une

période durant jusqu'à 4h permet le maintien de l'homéostasie du glucose dans l'organisme en continu (Geloneze et al, 2006).

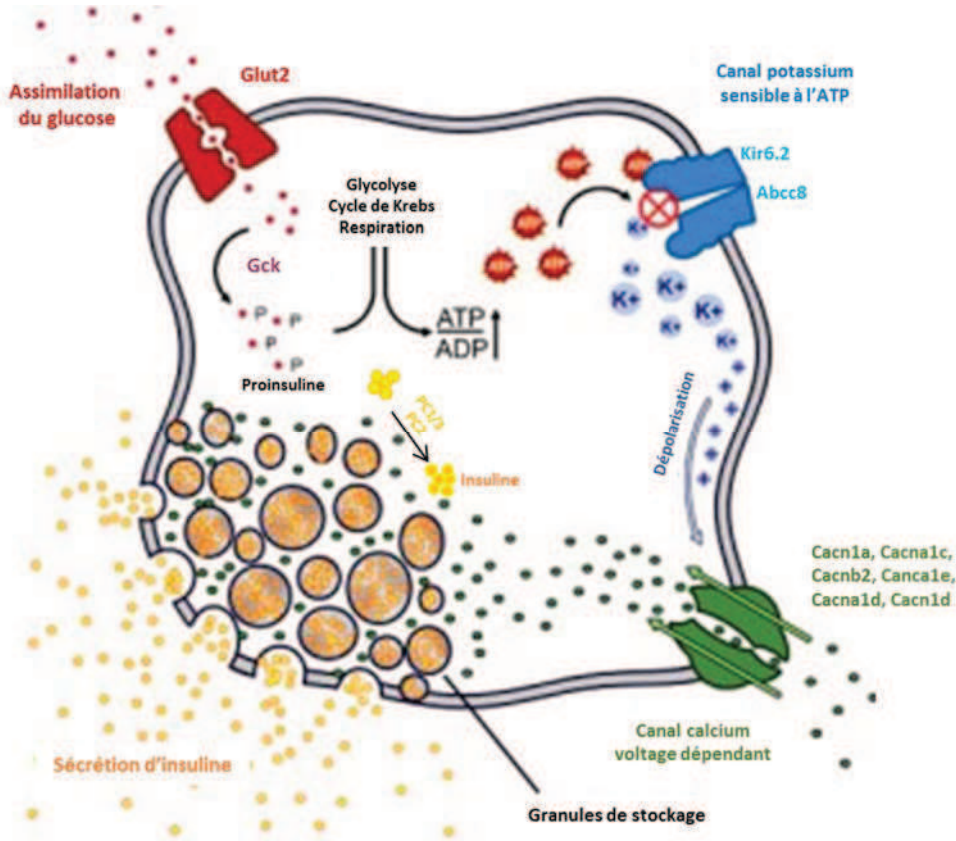


Figure 4 : Mécanisme de sécrétion d'insuline dans la cellule beta pancréatique. Le glucose entre dans la cellule beta grâce au transporteur Glut2, puis il est métabolisé par la glycolyse et le cycle de Krebs produisant des intermédiaires énergétiques qui conduisent à la production d'ATP à travers la chaîne respiratoire. Ceci entraîne une augmentation du ratio ATP/ADP et provoque la fermeture des canaux potassium sensibles à l'ATP. Une dépolarisation membranaire a alors lieu et permet l'entrée de calcium dans la cellule beta. Les granules de stockage de l'insuline, qui est produite à partir de proinsuline par la prohormone convertase PC1/3, fusionnent alors avec la membrane cellulaire et l'insuline est sécrétée dans le sang. (Adapté de betacellbiology.org)

C'est par cette sécrétion d'insuline hautement régulée que l'équilibre glycémique est maintenu dans l'organisme.

B. Le diabète, une pandémie mondiale

La principale pathologie associée au pancréas est le diabète, maladie liée à une élévation anormale du taux de glucose dans le sang, ou hyperglycémie, causée par un défaut de production et/ou d'utilisation de l'insuline. Dans le monde une personne sur douze est atteinte de diabète et une personne sur deux touchée par cette maladie ne le sait pas. La Fédération Internationale du Diabète prévoit 205 millions de diabétiques de plus qu'aujourd'hui dans le monde (**Figure 5**) en 2035. L'Organisation Mondiale de la Santé qualifie ainsi cette maladie de pandémie mondiale « émergente et imputable aux augmentations rapides enregistrées dans la fréquence du surpoids, de l'obésité et de la sédentarité ».

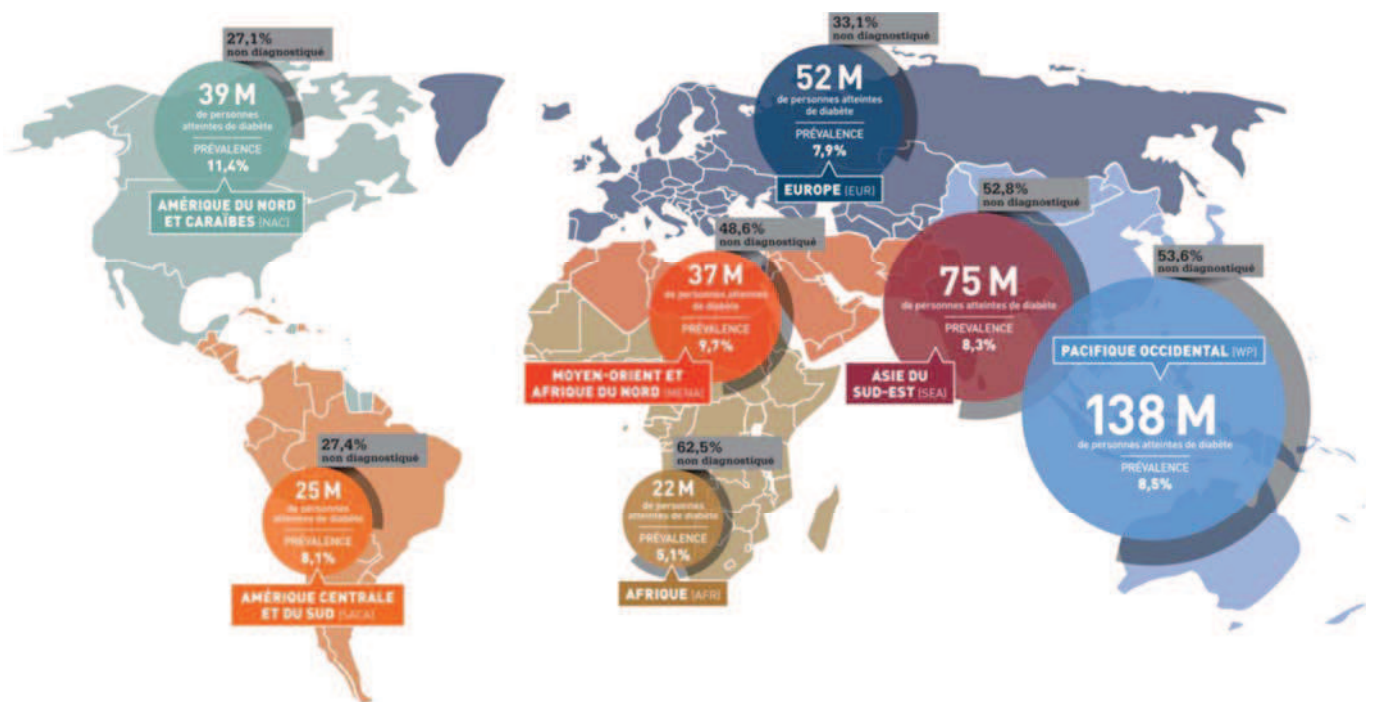


Figure 5 : Population mondiale touchée par le diabète et prévalence de la maladie. 388 millions de personnes sont diabétiques dans le monde et la prévalence moyenne est de 8,4%. On compte 46,3% de diabétiques non diagnostiqués dans le monde. (Fédération Internationale du Diabète, données 2014)

Il existe deux formes principales de diabète.

B.1. Le diabète insulino-dépendant

Le diabète de type 1 ou Diabète Insulino-Dépendant (DID) est une maladie d'origine auto-immune qui conduit à la destruction de plus de 90% des cellules beta des îlots de Langerhans et à l'absence de synthèse d'insuline. Cette maladie est multifactorielle et sa cause n'est pas clairement identifiée (Fuhlbrigge & Yip, 2014). Il survient généralement de manière brutale chez l'enfant ou chez le jeune adulte. Le DID est actuellement le plus souvent traité par injections pluriquotidiennes d'insuline. Ce traitement est lourd et dans de nombreux cas de diabète dérégulé, des complications sévères peuvent apparaître telles que des rétinopathies, néphropathies ou neuropathies. La transplantation d'îlots pancréatiques permet également de traiter le DID (Shapiro et al, 2000) mais bien qu'elle ait montré sa faisabilité et son efficacité à plus ou moins long terme en restaurant une insulino-indépendance chez le patient diabétique, cette approche reste limitée à certains patients. En effet, le nombre de pancréas disponibles ne permet pas de proposer cette thérapeutique à tous les patients et la mise en place d'un traitement antirejet lourd limite le nombre de receveurs potentiels (Barton et al, 2012; Bellin et al, 2012).

B.2. Le diabète non insulino-dépendant

Le diabète de type 2 ou Diabète Non Insulino-Dépendant (DNID) est le plus répandu (90% des cas) et résulte d'une mauvaise utilisation de l'insuline par l'organisme. Il est caractérisé par une résistance des muscles, du tissu adipeux et du foie à l'action de l'insuline. La sédentarité et une alimentation déséquilibrée ou trop riche peuvent être à l'origine de cette maladie. Par ailleurs, de nombreux gènes de prédisposition au DNID ont été révélés par des études de génomes à grand échelle (Bonfond et al, 2010; Dorajoo et al, 2015). Le DNID est actuellement traité par des comprimés antidiabétiques et par un régime alimentaire strict et des activités physiques régulières. A long terme, ce diabète peut devenir insulino-nécessitant et l'insulinothérapie devient alors incontournable.

B.3. Le diabète gestationnel

Le diabète gestationnel constitue un troisième type de diabète qui est caractérisé par une hyperglycémie chronique apparaissant pendant la grossesse. Ce type de diabète se développe car l'organisme de la mère ne peut former et utiliser suffisamment d'insuline lors de la grossesse. Un contrôle glycémique régulier et un régime faible en sucre permet le développement normal du bébé. Cependant, le risque de développer un diabète de type 2 est augmenté pour la mère et le bébé par la suite.

B.4. Les diabètes monogéniques

Enfin, les diabètes monogéniques constituent un autre type de diabète relativement rare (1 à 2% des diabètes) et sont exclusivement d'origine génétique (**Tableau 1**). Les mutations des différents gènes impliquées donnent lieu à des diabètes de différentes sévérités, apparaissant à différents âges, et sont associées à des risques de complications plus ou moins sévères (Ashcroft & Rorsman, 2012; Molven & Njolstad, 2011).

Gène	Fonction cellulaire	Type de diabète	Référence
Kir6.2	Sous-unité du canal potassium dépendant à l'ATP impliqué dans la sécrétion de l'insuline	PNDM, TNDM, MODY	Flanagan et al. 2009 ; Gloyn et al. 2004 ; Hattersley and Ashcroft, 2005
Abcc8	Sous-unité du canal potassium dépendant à l'ATP impliqué dans la sécrétion de l'insuline	PNDM, TNDM	Ellard et al. 2007 ; Hattersley and Ashcroft, 2005
Ins	Hormone hypoglycémisante	PNDM, MODY10	Stoy et al. 2010
Gck	Glycolyse	PNDM, MODY2	Froguel et al. 1993
Slc2a2	Récepteur au glucose	PNDM	Yoo et al. 2002
Pdx1	Facteur de transcription	PNDM, MODY4	Stoffers et al. 1997
Glis3	Facteur de transcription	Synd	Senée et al. 2006
Foxp3	Facteur de transcription	Synd	Bennett et al. 2001
Eif2ak3	Phosphorylation du facteur de transcription El2F	Synd	Delépine et al. 2000
Ptf1a	Facteur de transcription	Synd	Sellick et al. 2004
Rfx6	Facteur de transcription	Synd	Smith et al. 2010
Ngn3	Facteur de transcription	ND	Pinney et al. 2011
mutation locus 6q24	/	TNDM	Temple and Shield, 2010
Tcf2	Facteur de transcription	TNDM, MODY5	Servitja and Ferrer, 2004
Hnf4 alpha	Facteur de transcription	MODY1	Yamagata et al. 1996
NeuroD1	Facteur de transcription	TNDM, MODY6	Mlecki et al. 1999
Tcf1	Facteur de transcription	MODY3	Servitja and Ferrer, 2004
Blk	Kinase stimulant la synthèse et la sécrétion d'insuline	MODY11	Borowiec et al. 2009
Cel	Hydrolyse et Absorption du cholestérol et des esters	MODY8	Torsvik et al. 2010
Klf11	Facteur de transcription	MODY7	Neve et al. 2005
Mtt L1,E,K	Protéine de transferts des ARNs mitochondriaux	DMT	Maassen et al. 2002
Pax4	Facteur de transcription	MODY9	Hebener et al. 1998
Wsf1	Protéine transmembranaire du réticulum endoplasmique	Synd	Rohavem et al. 2011
Pax6	Facteur de transcription	PNDM	Solomon et al. 2009
Gata4	Développement du pancréas	PNDM	D'Amato et al. 2010
Hnf1 beta	Facteur de transcription	PNDM, MODY5	Horikawa et al. 1997
Hnf1 alpha	Facteur de transcription	MODY3	Yamagata et al. 1996

Tableau 1 : Tableau récapitulatif des gènes impliqués dans des diabètes monogéniques. La fonction cellulaire, le type de diabète et la publication scientifique de référence sont indiquées pour chaque mutation. MODY = Maturity Onset Diabetes of the Young, ND = Neonatal Diabetes, Synd = syndromic diabetes, TNDM = Transient Neonatal Diabetes Mellitus, PNDM = Permanent Neonatal Diabetes Mellitus, DMT = mitochondrial diabetes. Adapté de (Ashcroft & Rorsman, 2012; Molven & Njolstad, 2011)

B.5. Axes de recherches et thérapies cellulaires pour le diabète

Pour contourner les actuelles limites des traitements existants, différents axes de recherche illustrés en **Figure 6** sont en cours d'étude dans le but de maîtriser la formation de nouvelles cellules beta *in vitro* et à plus long terme de développer de nouvelles thérapies

cellulaires pour les patients diabétiques. Les principales voies étudiées pour cela sont la différenciation de cellules embryonnaires souches (ESC) ou à pluripotence induite (IPSC), la transdifférenciation par reprogrammation d'autres types cellulaires et la stimulation de la prolifération des cellules beta (Pagliuca & Melton, 2013).

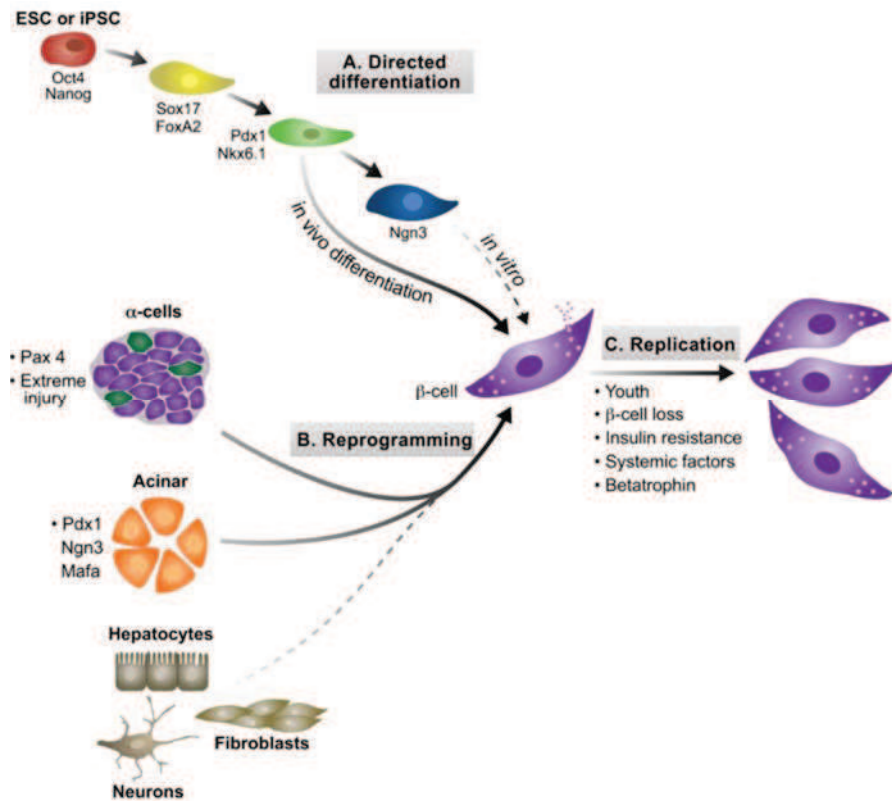


Figure 6 : Stratégies de génération de nouvelles cellules beta sécrétrices d'insuline.

Une première stratégie (A) permet la génération de cellules beta par différenciation de cellules souches embryonnaires (ESC) ou induites (IPSC) à l'aide de facteurs de croissance et de transcription. Une autre stratégie (B) consiste en la reprogrammation de cellules différenciées en cellules beta. De nouvelles cellules beta peuvent également être formées en stimulant la prolifération des cellules beta existantes (C). (Pagliuca & Melton, 2013)

A ce titre, la différenciation de cellules souches embryonnaires humaines en cellules beta en récapitulant les programmes de différenciation embryonnaire a été étudiée (D'Amour et al, 2006; Kroon et al, 2008) et est illustrée en **Figure 7**. Ces études démontrent la possibilité de développer des cellules beta *in vitro* mais celles-ci ne synthétisent pas d'insuline en réponse au glucose. Par contre si les cellules progénitrices formées *in vitro* sont transplantées chez la souris immuno-déficiente, celle-ci peut alors produire des cellules

sécrétrices d'insuline matures. En 2012, Rezania et al. (Rezania et al, 2012) développent également un protocole de différenciation de cellules souches embryonnaires humaines en cellules progénitrices exprimant Pdx1 qui sont ensuite maturées *in vivo* chez la souris en cellules endocrines. Ils proposent ainsi un traitement alternatif des patients diabétiques par des progéniteurs pancréatiques formés *in vitro*. Ainsi, ces études montrent qu'il manque encore une connaissance approfondie des étapes de différenciation terminales des cellules beta sécrétrices d'insuline.

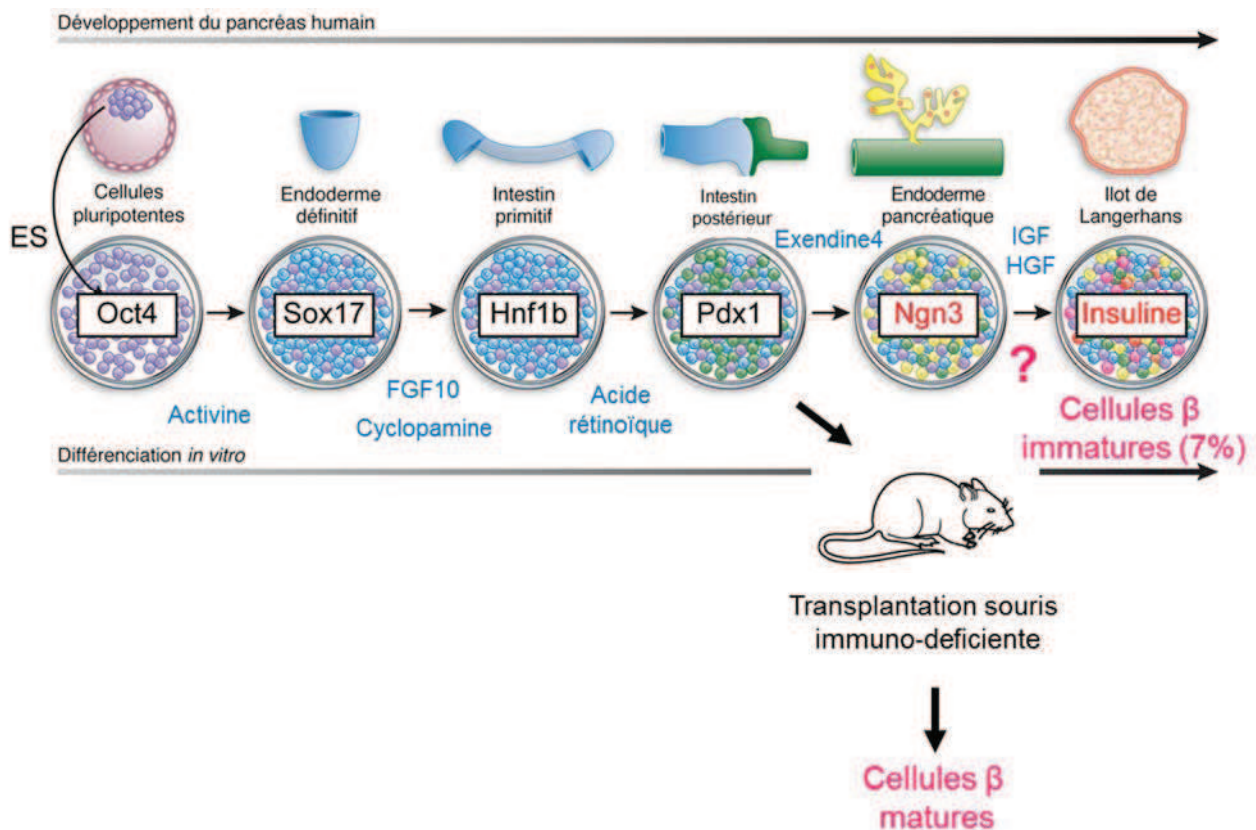


Figure 7 : Génération de cellules beta *in vitro* à partir de cellules souches embryonnaires humaines. Le développement du pancréas humain est mimé pour obtenir des cellules endocrines pancréatiques à partir de cellules souches embryonnaires humaines. Celles-ci sont cultivées en présence de différents facteurs principaux indiqués sur la figure (Activine, FGF10 (fibroblast growth factor 10), Acide rétinoïque, Exendine4, IGF (insulin growth factor) et HGF (hepatic growth factor)) et permettent l'obtention de 7% de cellules beta immatures. Les cellules progénitrices expriment les marqueurs du développement pancréatique indiqués dans les encadrés. Si les cellules exprimant Pdx1 ou Ngn3 sont transplantées chez la souris immuno-déficiente, celle-ci peut alors produire des cellules sécrétrices d'insuline matures. Les étapes de différenciation terminale des cellules beta ne sont à l'heure actuelle pas encore connues. Adapté de (D'Amour et al, 2006; Kroon et al, 2008; Madsen & Serup, 2006)

Plus récemment, un protocole permettant la différenciation *in vitro* de cellules souches pluripotentes humaines en cellules beta, est établi par Pagliuca et al. (Pagliuca et al, 2014). Celui-ci permet d'obtenir des cellules sécrétant des quantités d'insuline en réponse au glucose comparables aux cellules beta adultes humaines et également de réduire l'hyperglycémie chez la souris diabétique.

En 2015, la reprogrammation de cellules souches du mésenchyme pancréatique en cellule beta fonctionnelle *in vitro* est proposée par Guo et al. (Guo et al, 2014) par transfection de ces cellules à l'aide de l'ARNm du marqueur pancréatique Pdx1.

Ainsi, c'est l'élucidation des signaux et cascades de régulation transcriptionnelle contrôlant la différenciation des cellules endocrines du pancréas qui pourrait aboutir, à plus long terme, au développement d'une thérapie cellulaire pertinente pour le diabète.

C. Développement du pancréas et régulation transcriptionnelle

Chez la souris (temps de gestation de 20 jours), le développement du pancréas, illustré en **Figure 8**, débute dès le neuvième jour du développement embryonnaire (E9) par bourgeonnements ventral et dorsal à partir de l'endoderme de l'intestin primitif (Marquez-Aguirre et al, 2015; Murtaugh, 2007; Pan & Wright, 2011). Après croissance de l'épithélium pancréatique, une rotation de l'intestin a lieu et permet la formation du pancréas par fusion des bourgeons. Les premières cellules endocrines apparaissent lors de la première transition avant la fusion des bourgeons pancréatiques, puis lors de la deuxième transition, le nombre de cellules endocrines augmente et la différenciation du tissu exocrine débute. À partir d'E13, la structure du pancréas s'organise en différentes branches (Villasenor et al, 2010) composées de troncs à l'origine de progéniteurs endocrines et canaux et d'extrémités qui donneront les cellules acinaires (Zhou et al, 2007). Au cours de leur différenciation, les cellules endocrines quittent l'épithélium par délamination puis s'agrègent pour former les îlots de Langerhans.

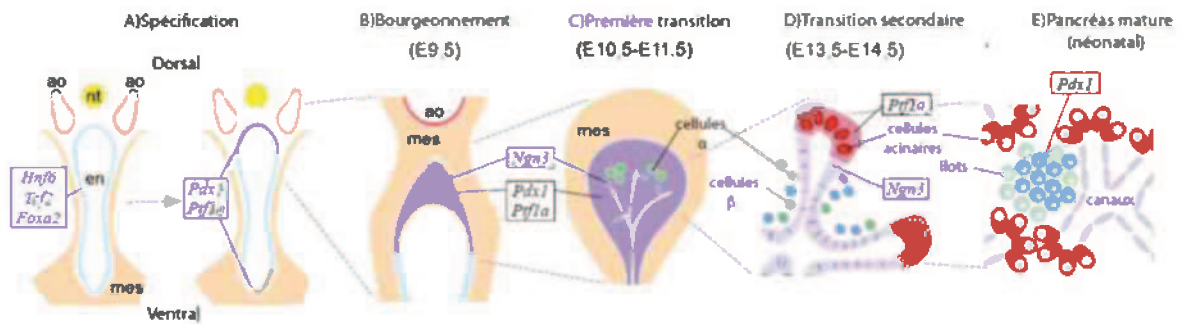


Figure 8 : Coupe transversale schématique embryonnaire du pancréas chez la souris lors du développement, (en : endoderme, mes : mésenchyme, nt : notochorde, ao : aorte). (D'après (Murtaugh, 2007))

Le développement pancréatique est hautement contrôlé par un grand nombre de voies de signalisation cellulaires (non développées ici) et de facteurs de transcription dont les principaux, intervenant dans la formation des cellules beta matures, seront décrits dans ce paragraphe.

C.1. Spécification pancréatique

La cascade de régulation transcriptionnelle complexe intervenant lors du développement du pancréas est présentée en **Figure 9**. Les principaux facteurs de transcription indiqués régulent la différenciation cellulaire pour aboutir au pancréas mature et à la spécification des différentes cellules le composant. Dans un premier temps, les cellules deviennent des progéniteurs pancréatiques multipotents qui sont à l'origine de toutes les cellules de l'épithélium pancréatique. L'expression de deux facteurs de transcription essentiels permet cette spécification pancréatique : Ptf1a et Pdx1.

Pdx1 (Pancreas and duodenal homeobox gene 1) est un FT à homéodomaine exprimé dès E8,5 (Jorgensen et al, 2007). Son inactivation entraîne une agénésie du pancréas exocrine et endocrine et la mort en quelques jours (Jonsson et al, 1994). La délétion hétérozygote de Pdx1 n'est pas létale mais entraîne un diabète expliqué par une baisse du nombre de cellules beta (Okita et al, 2015). L'expression de Pdx1 est dynamiquement régulée par Foxa1 et 2 lors du développement du pancréas (Gao et al, 2008). Ce FT est ainsi à l'origine de l'ensemble des cellules pancréatiques acinaires, canalaire et endocrines.

Ptf1a (Pancreas transcription factor 1 subunit alpha) est un FT à motif basique hélice-boucle-hélice (bHLH), il constitue la sous-unité qui lie l'ADN de la protéine trimérique PTF1

(composée des sous-unités p48 (Ptf1a), p64 et p75). Les souris mutantes pour Ptf1a ne développent pas de tissu exocrine et seulement une ébauche de tissu endocrine et meurent rapidement après la naissance (Krapp et al, 1998). Ptf1a est l'acteur clé du choix de la destinée exocrine ou endocrine des progéniteurs pancréatiques, il est essentiel à la formation du pancréas ventral et à l'origine des cellules pancréatiques différenciées (Kawaguchi et al, 2002b). En 2012, Thompson et al. (Thompson et al, 2012) montrent que Ptf1a stabilise l'identité des progéniteurs pancréatiques en contrôlant l'expression de différents FT dont Mnx1, Pdx1, Nkx6.1 et Onecut1.

Ces deux FT permettent ainsi la formation de progéniteurs pancréatiques à partir de l'endoderme de l'intestin primitif.

C.2. Spécification endocrine

La destinée endocrine des progéniteurs pancréatiques multipotents est ensuite assurée par le FT pro-endocrine **Ngn3** de la famille bHLH.

En effet, lorsque ce gène est inactivé, une mort rapide des souris, qui présentent dans ce cas une agénésie du pancréas endocrine, est observée (Gradwohl et al, 2000). L'intervention précoce de Ngn3 dans le développement du lignage endocrine est également supportée par une absence de co-expression avec les hormones pancréatiques. Par traçage cellulaire, Gu et al. (Gu et al, 2002) montrent que les cellules exprimant Ngn3 sont à l'origine de toutes les cellules endocrines pancréatiques. En 2012, Oropez et Horb (Oropez & Horb, 2012) montrent qu'une expression transitoire de Ngn3 tôt lors du développement chez le xénope favorise la formation des cellules delta et beta.

Ngn3 est également exprimé dans le lignage exocrine et c'est son niveau d'expression qui assure la destinée endocrine des progéniteurs (Wang et al, 2009). En 2012, notre équipe montre, par modulation de la délétion de Ngn3 au cours du développement embryonnaire, que la capacité de différenciation des cellules exprimant Ngn3 évolue. Les cellules délétées pour Ngn3 avant E12,5 conduisent à la formation de cellules acinaires et canalaire. Et lorsque Ngn3 est délété après E12,5, ce sont des cellules canalaire qui sont formées (Beucher et al, 2012).

L'engagement des cellules progénitrices dans la voie endocrine est donc assuré par l'expression de Ngn3. La différenciation endocrine est aussi régulée par ce facteur de transcription, puisqu'il régule des facteurs requis pour cette spécification tels que Insm1 (Breslin et al, 2007) et NeuroD1 (Huang et al, 2000). Enfin, plus récemment, le rôle essentiel de Ngn3 dans la détermination de la destinée endocrine au cours du développement du

pancréas humain est démontré par la technologie CRISPR/Cas9 dans des cellules embryonnaires souches humaines (McGrath et al, 2015).

C.3. Spécification des sous-types endocriniens

La spécification des sous-types cellulaires endocriniens est ensuite permise par l'expression d'un grand nombre de facteurs de transcription clés lors de la différenciation, tels qu'Arx, Pax4, Nkx2.2 et 6.1, NeuroD1, MafA et FoxA2 qui sont notamment déterminant pour la formation des cellules beta sécrétrices d'insuline.

L'expression d'**Arx** (Aristaless related homeobox) et de **Pax4** (Paired box 4) permet respectivement la formation des cellules alpha, epsilon et PP, et delta et beta (Collombat et al, 2007; Collombat et al, 2003). En effet, les souris mutantes pour Arx présentent une absence de cellules alpha au profit des cellules beta et delta alors que l'inverse est observé par surexpression de ce FT. Les souris mutantes pour Pax4 forment des cellules alpha au détriment des cellules beta et PP (Sosa-Pineda et al, 1997). Il a également été démontré qu'il est possible par expression ectopique de Pax4, dans des cellules alpha, de leur conférer une destinée beta (Collombat & Mansouri, 2009). L'étude de doubles mutants Arx/Pax4 montre, de plus, que la perte d'Arx est nécessaire et suffisante pour former des cellules beta à partir de cellules alpha (Courtney et al, 2013). Ces études indiquent que les cellules endocrines matures gardent une plasticité.

Nkx2.2 et 6.1 permettent également la différenciation des cellules endocrines pancréatiques. Les souris mutantes pour Nkx2.2 présentent un nombre réduit de cellules alpha, beta et PP et meurent une semaine après la naissance par retard de croissance et diabète (Sussel et al, 1998). En 2004, Wang et al. (Wang et al, 2004) montrent que Nkx2.2 et Pax4 collaborent pour permettre l'initiation de la différenciation beta. Les souris mutées pour Nkx6.1 présentent également une altération de la différenciation endocrine (Sander et al, 2000). En 2013, Taylor et al. et Schaffer et al. montrent (Schaffer et al, 2013; Taylor et al, 2013) que l'action de Nkx6.1 est requise dans les progéniteurs exprimant Pdx1 pour la différenciation en cellules beta et qu'il réprime la différenciation des cellules acinaires.

NeuroD1 a également été montré comme essentiel dans la mise en place du lignage cellulaire beta. L'absence de ce gène, ciblé par Ngn3, chez les souris mutantes est létale et provoque un diabète néonatal (Huang et al, 2000). Les souris n'exprimant pas NeuroD1 spécifiquement dans les cellules beta présentent une sécrétion d'insuline stimulée par le glucose diminuée et un profil métabolique identique à des cellules beta immatures (Gu et al, 2010).

MafA est exprimé uniquement dans les cellules beta. Les souris mutantes pour ce gène présentent une expression d'insuline diminuée et développent un diabète provoqué par une diminution progressive du nombre de cellules beta (Zhang et al, 2005). Son expression est également perdue dans les souris mutantes pour Nkx6.1, suggérant ainsi son rôle essentiel pour le développement des cellules sécrétrices d'insuline.

Le rôle de **FoxA2** dans la maintenance de la fonction beta est démontré par Gao et al. (Gao et al, 2010). L'inactivation de FoxA2 spécifiquement dans les cellules beta provoque le décès des souris peu après la naissance. Une désorganisation des ilots pancréatiques et une sécrétion d'insuline altérée est également observée. Ceci est spécifiquement dû à une diminution d'expression des sous-unités des canaux potassium sensibles à l'ATP Kir6.2 et Abcc8 (Sund et al, 2001).

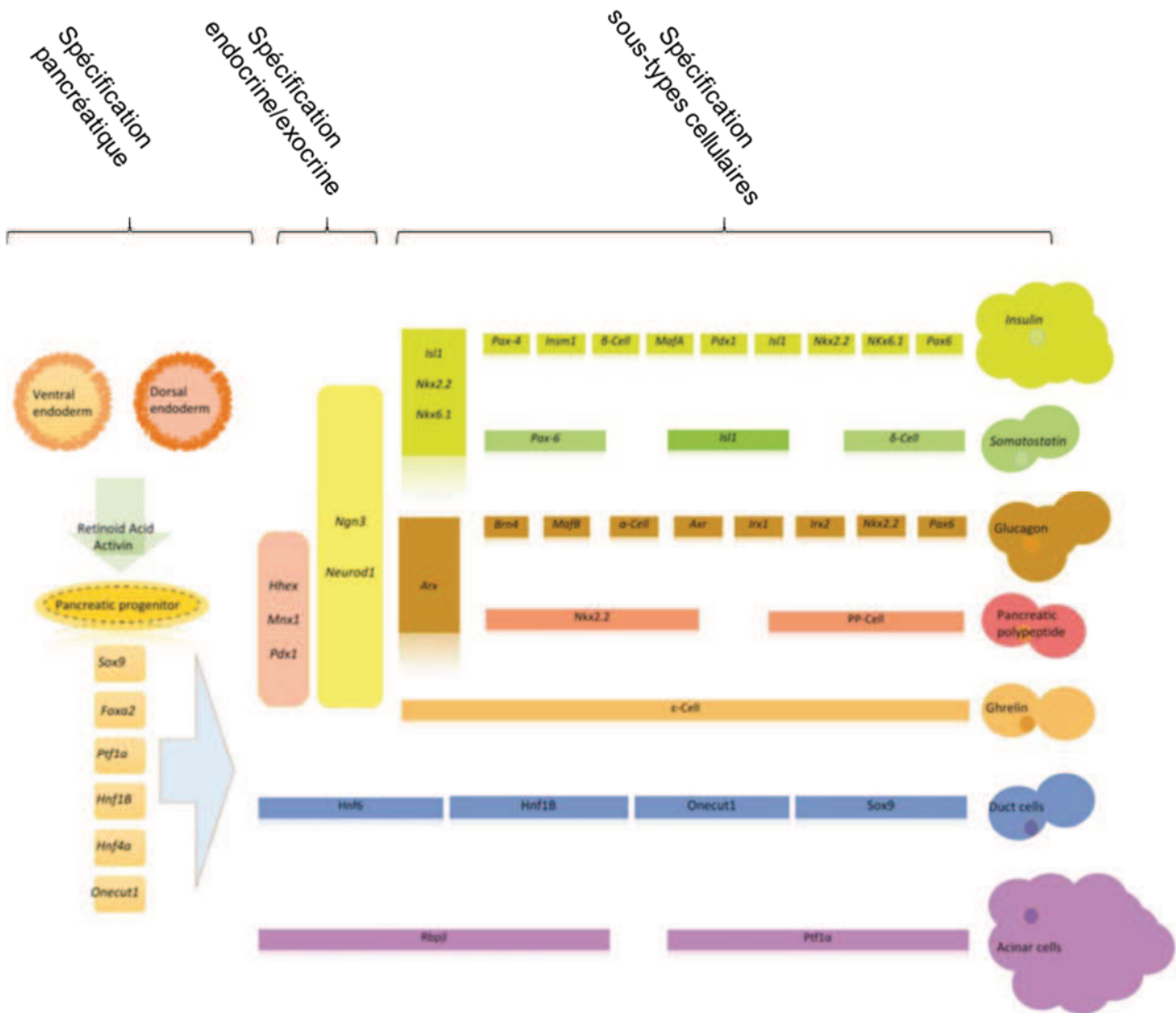


Figure 9 : Cascade de régulation transcriptionnelle contrôlant la différenciation cellulaire lors du développement du pancréas chez la souris. Les progéniteurs pancréatiques sont formés à partir de l'endoderme de l'intestin primitif. Après définition du lignage endocrine, les différents sous-types cellulaires alpha, beta, delta, PP et epsilon sont déterminés. Les cellules acinaires et canaliculaires constituent le pancréas exocrine. Les facteurs de transcription clés impliqués à chaque étape sont indiqués. (D'après (Marquez-Aguirre et al, 2015))

Récemment, Benitez et al. (Benitez et al, 2014) proposent une caractérisation transcriptomique détaillée des populations cellulaires progénitrices et adultes qui sont présentes tout au long du développement du pancréas. Par l'étude transcriptomique de progéniteurs endocrines exprimant Ngn3 ou non, ils découvrent aussi de nouveaux régulateurs du développement et de la fonction du pancréas mature tels que Runx1t1,

Bcl11a, Prdm16 ou Etv1. Ceci montre que, bien que soient déjà connus différents FT et leurs implications dans la régulation transcriptionnelle des étapes précoces de la différenciation du pancréas endocrine, d'autres recherches doivent être développées dans le but d'élucider notamment les mécanismes de régulation des étapes de différenciation terminale des cellules beta sécrétrices d'insuline. Dans ce cadre, afin d'identifier de nouveaux gènes contrôlant la différenciation insulaire en aval de Ngn3, notre équipe a réalisé des expériences de puce à ADN. A partir d'un pancréas embryonnaire de souris transgéniques du laboratoire exprimant la YFP dans les cellules exprimant Ngn3 (Mellitzer et al, 2004), qui constituent les précurseurs immatures de l'ensemble des cellules endocrines des îlots de Langerhans, un enrichissement important du FT Rfx6 a été révélé (Soyer et al, 2010). L'étude du rôle de RFX6 dans la différenciation endocrine du pancréas et de l'intestin a ainsi été initiée dans notre équipe.

D. Le facteur de transcription RFX6

D.1. La famille RFX

La famille des facteurs de transcription RFX (Regulatory Factor X) a été révélée en 1996 par Emery et al. (Emery et al, 1996a) . Elle est retrouvée chez différentes espèces. Sept membres ont été identifiés chez les mammifères (Rfx1 à 7), trois chez la drosophile (Rfx, Rfx1 et 2) et un unique a été mis en évidence chez *C. elegans* (Daf-19). Deux membres supplémentaires de cette famille ont récemment été identifiés chez les vertébrés (Choksi et al, 2014) (**Figure 11**) : Rfx8, conservé et récemment prédit (ENSG00000196460) et Rfx9 qui n'est pas encore référencé.

Ces gènes présentent une structure conservée (Aftab et al, 2008) présentée en **Figure 10** pour les gènes de la souris. Elle est notamment composée d'un domaine de fixation à l'ADN de type « winged helix » composé de deux « wings », de trois hélices alpha et de trois feuillets beta (Brennan, 1993). Ce motif « winged helix » permet aux protéines RFX d'établir un contact avec un motif X-box de 14 nucléotides au niveau de l'ADN. Celui-ci est défini par Laurençon et al. en 2007 (Laurençon et al, 2007) comme suit : RYYNYYN(1,3)RRNRAC. En 2012, un motif de 8 à 11pb de type C-box (séquence riche en Cytosine et Thymidine) complémentaire à la X-box est identifié chez *C. elegans* (Burghoorn et al, 2012). Cette étude suggère un contrôle fin de l'expression des cibles du facteur DAF19 chez *C. elegans* par la présence ou non de ce motif activateur d'expression. Les protéines RFX possèdent également des domaines de dimérisation B, C et DIM (sauf RFX5 et RFX7) leur permettant d'établir des complexes hétéro ou homo-dimères pour leur fixation au niveau du petit sillon de l'ADN. Cependant ces protéines sont tout de même actives à l'état de monomère (Katan-

Khaykovich & Shaul, 1998). RFX 1, 2, et 3 possèdent également un domaine de transactivation au niveau de leur extrémité N-terminale (Aftab et al, 2008).

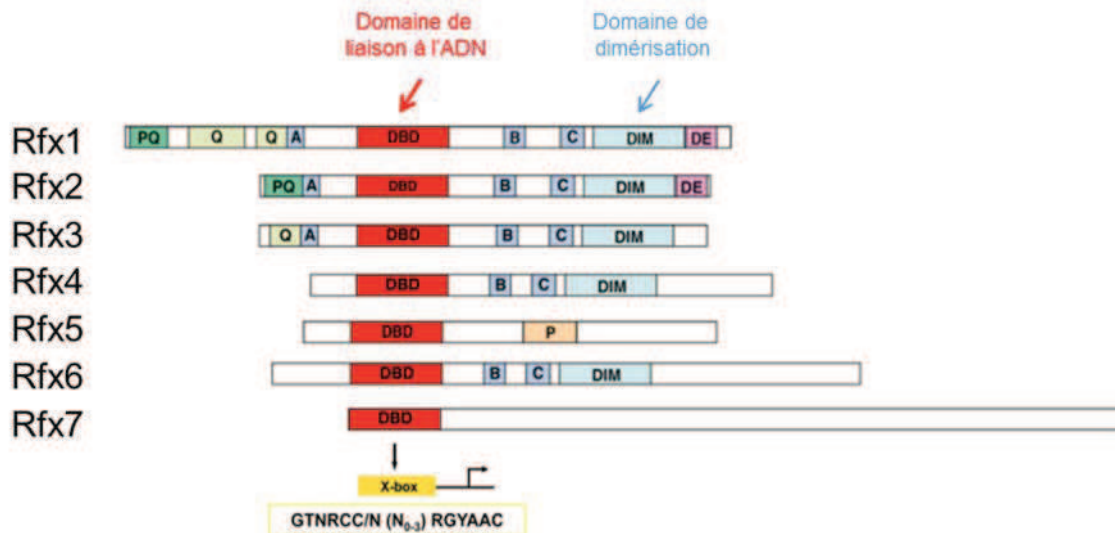


Figure 10 : Structure des facteurs de transcription RFX chez la souris. Tous les membres de la famille RFX possèdent un domaine de liaison à l'ADN conservé, qui fixe le motif X-box au sein de l'ADN (défini par Emery et al. (Emery et al, 1996b) pour RFX1). (DBD = DNA binding domain, DIM = dimerization domain, DE = acidic amino acid rich region, P = proline rich region, PQ = proline and glutamine rich region, Q = glutamine rich region, A,B C = conserved domains of unknown functions). Adapté de (Thomas et al, 2010)

D.2. Fonctions diverses des membres de la famille

Les facteurs de transcription RFX présentent une grande diversité de fonctions connues à ce jour et décrites dans la revue de Choksi et al. (Choksi et al, 2014) spécifiquement concentrée sur leur implication dans la ciliogenèse (**Figure 11**). En effet, les protéines RFX contrôlent notamment la ciliogenèse en activant la transcription de gènes codant pour des protéines de structure et de fonction des cils qui interviennent dans le transport intra-flagellaire (formation des cils), la motilité ciliaire ou encore la coordination du transport et de la régulation du trafic ciliaire. Ce rôle dans la ciliogenèse est maintenu chez différents organismes tels que *C. elegans* (Swoboda et al, 2000), la drosophile (Dubruille et al, 2002) mais également chez la souris (El Zein et al, 2009). La fonction des protéines RFX chez les mammifères est par contre moins connue.

Rfx1 a été le premier membre de cette famille à être identifié et son action sur les gènes du complexe majeur d'histocompatibilité de classe II a été révélée (Reith et al, 1990). Par la

suite le même groupe montre que Rfx1 est nécessaire à l'expression des gènes du virus de l'hépatite B (Reith et al, 1994). Enfin, son rôle dans la réponse aux lésions de l'ADN est découvert en 2005 (Lubelsky et al, 2005). Et la perte de fonction de Rfx1 chez la souris dévoile l'aspect essentiel de ce gène pour le développement puisque les souris mutantes meurent avant E2,5 (Feng et al, 2009).

Le gène **Rfx2** a été caractérisé en 1994 (Reith et al, 1994) et son rôle non essentiel pour le développement précédemment proposé (Chung et al, 2012) mais capital pour la spermatogenèse est démontré chez la souris dans une étude récente (Shawlot et al, 2015). D'autre part, chez le xénope et le poisson zèbre, les embryons mutants pour Rfx2 présentent une fermeture du tube neural ainsi qu'un défaut d'asymétrie droite/gauche de l'organisme (Bisgrove et al, 2012).

Le gène **Rfx4** est quant à lui connu pour son rôle dans le développement du cerveau mais il est également impliqué dans le développement des cils primaires au niveau du système nerveux central chez la souris (Ashique et al, 2009).

Par ailleurs, il a été montré que **Rfx5** régule la transcription des gènes HLA (Human Leucocytes Antigen) de classe II (Hanna & Etzioni, 2014) intervenant dans les mécanismes immunitaires et est impliqué dans le développement du syndrome des lymphocytes dénudés (Bare), qui correspond à une absence de molécules HLA à la surface des lymphocytes T (Nekrep et al, 2002; Reith & Mach, 2001).

Les facteurs de transcription RFX1, 2, 3, 5 et 6 sont exprimés dans le pancréas chez la souris (Smith et al, 2010) mais RFX3 et RFX6 sont les seuls qui ont été identifiés pour leur rôle essentiel dans le développement du pancréas endocrine et la fonction de la cellule beta (Ait-Lounis et al, 2007; Ait-Lounis et al, 2010; Smith et al, 2010; Soyer et al, 2010).

Rfx3 régule l'expression de gènes requis pour l'assemblage et la fonction des cils primaires des cellules (El Zein et al, 2009) et une perte de fonction de ce gène induit un défaut d'asymétrie droite gauche des organes (Bonnafe et al, 2004). Par ailleurs, Rfx3 joue un rôle prépondérant dans la différenciation endocrine des cellules beta du pancréas lors du développement ainsi que dans leur fonction par la régulation de l'expression de la glucokinase (Ait-Lounis et al, 2007; Ait-Lounis et al, 2010). En effet, en 2007, l'étude de souris mutantes pour Rfx3 révèle que la différenciation du pancréas endocrine est affectée, et une réduction du nombre de cellules alpha et beta et une augmentation du nombre de cellules PP sont observées. D'autre part une réduction de la taille et du nombre des cils sur les cellules beta est constatée lors de la perte de fonction de Rfx3 (Ait-Lounis et al, 2007). En 2010, le rôle de Rfx3 dans la régulation de la fonction de la cellule beta est ancré par détermination de son gène cible, la glucokinase (Ait-Lounis et al, 2010). En effet, RFX3 se lie au niveau de deux motifs Pal1 et Pal2 située dans la région promotrice de la glucokinase.

	Organisms	RFX TFs	Protein domains	Expression patterns	Ciliary phenotypes		Key ciliary target genes
					Organism wide	Cilia specific	
Vertebrates	<i>H. sapiens</i> <i>M. musculus</i> <i>X. laevis</i> <i>D. rerio</i>	RFX1		Brain	Homozygous lethal	Not known	<i>ALMS1</i>
		RFX2		Organs of laterality Brain Kidney Testis Epidermis	Left-right asymmetry defects Defective neural tube closure Perturbed HH signaling	Truncated, dysfunctional motile cilia Fewer and truncated immotile cilia	<i>IFT122</i> <i>IFT172</i> <i>WDPCP</i> <i>TTC25</i>
		RFX3		Organs of laterality Brain Pancreas Epidermis	Left-right asymmetry defects Hydrocephalus Malformation of the corpus callosum Perturbed hormone secretion	Truncated, dysfunctional motile cilia Aberrant number of immotile cilia Truncated immotile cilia	<i>Dync2li1</i> <i>Dnah9</i> <i>Dnah11</i>
		RFX4		Brain Testis	Homozygous lethal Reduction/absence of SCO Patterning defects Hydrocephalus	Truncated cilia	<i>IFT172</i>
		RFX6		Pancreas	n/a	n/a	n/a
		RFX8		Not known	Not known	Not known	Not known
		RFX5		Immune system	n/a	n/a	n/a
		RFX7		Not known	Not known	Not known	Not known
		RFX9		Not known	Not known	Not known	Not known
Flies	<i>D. melanogaster</i>	RFX		Brain Chordotonal and external sensory neurons	Sensory behavioral defects	Structurally abnormal cilia	<i>nan</i> <i>iav</i>
		RFX1		Not known	Not known	Not known	Not known
		RFX2		Not known	Not known	Not known	Not known
Nematodes	<i>C. elegans</i>	DAF-19		Ciliated sensory neurons	Dauer formation Dye filling and sensory behavioral defects	Absence of all cilia	Numerous
Fungi	<i>S. cerevisiae</i>	CRT1		Single-celled budding yeast	n/a	n/a	n/a
	<i>S. pombe</i>	SAK1		Single-celled fission yeast	n/a	n/a	n/a

Figure 11 : Expressions et fonctions des facteurs de transcription RFX dans différents organismes. Les membres de la famille RFX de différents organismes sont indiqués et associés à leur expression et fonction. Un schéma de chaque protéine avec ses domaines fonctionnels est indiqué (domaine d'activation en bleu, domaine de liaison à l'ADN en vert, domaine B en rouge, domaine C en mauve, domaine de dimérisation en jaune). Les protéines surlignées en bleu et vert ont été identifiées pour leur rôle dans la ciliogenèse, contrairement aux protéines surlignées en jaune. (ALMS1 (Alstrom syndrome 1), Dnah (dynein axonemal heavy chain), Dync2li1 (dynein cytoplasmic 2 light intermediate chain 1), iav (inactive), IFT (intraflagellar transport genes), n/a (not applicable), nan (nanchung), SCO (subcommissural organ), TFs (transcription factors) TTC25 (tetratricopeptide repeat domain 25), WDPCP (WD repeat containing planar cell polarity effector). (Choksi et al, 2014)

D.3. RFX6 et le pancréas

Le gène **Rfx6** a été caractérisé par étude bio-informatique en 2008 par Aftab et al. (Aftab et al, 2008). L'expression de Rfx6 pendant l'embryogenèse et le développement du pancréas chez la souris a été étudiée dans notre équipe par hybridation *in situ* et immunohistochimie (Soyer et al, 2010). Rfx6, est exprimé au niveau de l'endoderme intestinal à E9,5 (**Figure 12A**), puis son expression est restreinte aux progéniteurs pancréatiques à partir de E10,5 puis endocrines (**Figure 12B et C**), et enfin Rfx6 est exprimé dans les cellules endocrines matures chez l'embryon et la souris adulte (**Figure 12D-G**). Ce patron d'expression spatio-temporel suggère une fonction multiple de Rfx6 dans le développement de l'endoderme et dans la différenciation et la fonction des ilots pancréatiques. En accord avec cette hypothèse, des études d'inactivation de ce gène ont montré une altération du développement du pancréas endocrine chez le poisson zèbre (Soyer et al, 2010). En effet, lorsque l'expression de *Rfx6* est inhibée, une baisse importante du nombre de cellules alpha, epsilon et delta, une réduction plus faible du nombre de cellules beta ainsi qu'une désorganisation des ilots sont observées. De plus, une augmentation du nombre de progéniteurs endocrines a lieu au détriment du développement de cellules différenciées. Les cellules endocrines sont ainsi bloquées au stade de progéniteurs dans les cellules mutées pour Rfx6.

De même, par l'étude de souris mutantes pour Rfx6, Smith et al. démontrent que les cellules endocrines pancréatiques chez la souris sont formées mais qu'elles n'expriment pas d'hormones, excepté pour les cellules PP, et que les souris meurent rapidement suite à des problèmes de nutrition (obstruction intestinale) (Smith et al, 2010). Ils observent également qu'en absence de Rfx6, l'expression de Ngn3 est inchangée alors que des gènes agissant en aval de Ngn3 sont affectés. L'expression de Pax4, qui intervient dans la spécification beta, est augmentée en absence de Rfx6 et les gènes impliqués dans la transcription de l'insuline comme Pdx1, NeuroD1 ou MafA sont réprimés. Dans cette étude, il est également montré que RFX6 peut lier l'ADN sous forme d'homo- ou d'hétérodimère avec RFX3 *in vitro*. Par contre son absence n'a aucune conséquence sur les cils primaires des cellules beta.

En 2011, l'étude de Rfx6 chez le xénope (Pearl et al, 2011) confirme que son expression apparait tout au long du développement et est ensuite restreinte aux cellules endocrines. En absence de Rfx6, l'expression de marqueurs du développement ou de la fonction endocrine pancréatique tels que NeuroD1, l'insuline, le glucagon ou la somatostatine, est perdue.

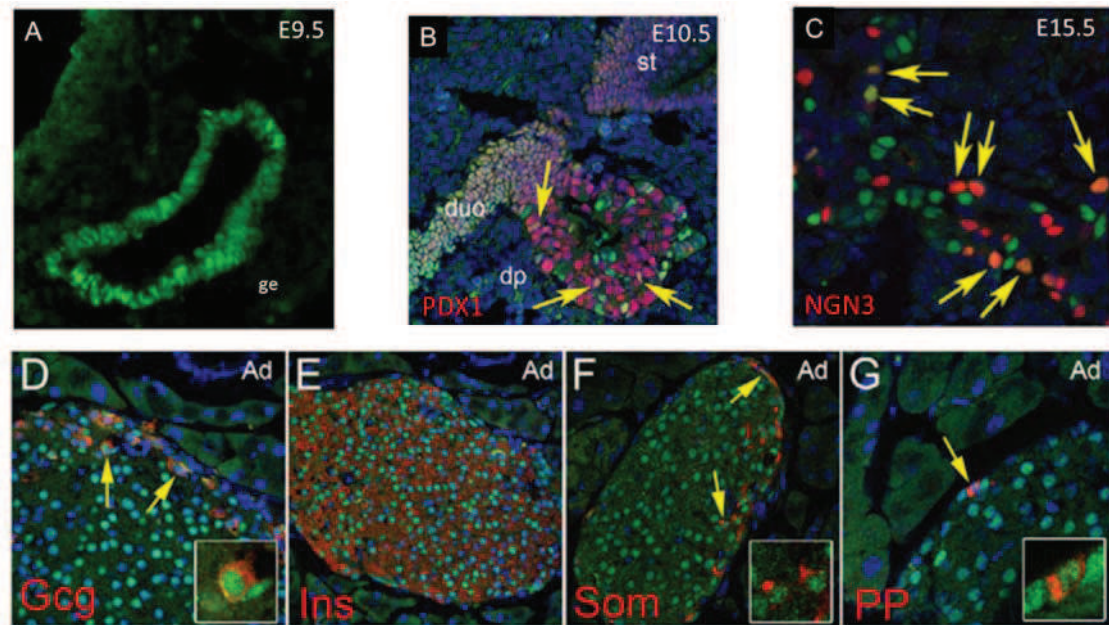


Figure 12 : RFX6 est exprimé pendant le développement du pancréas et dans les cellules pancréatiques adultes. Doubles immunofluorescences révélant l'expression de Rfx6 (en vert) tout au long du développement du pancréas dans l'endoderme intestinal (A), une sous partie des progéniteurs pancréatiques ((B) Pdx1, en rouge) et dans les progéniteurs endocrines ((C) Ngn3, en rouge). Rfx6 est aussi exprimé dans les cellules pancréatiques alpha ((D) glucagon Gcg, en rouge), beta ((E) insuline Ins, en rouge), delta ((F) somatostatine Som, en rouge) et PP ((G) polypeptide pancréatique PP, en rouge) au stade adulte. Les flèches jaunes indiquent les cellules doublement marquées. (ge = gut endoderm, st = stomach, duo = duodenum, dp = dorsal pancreas, Ad = Adulte). Adapté de (Soyer et al, 2010)

L'ensemble de ces données a permis d'identifier Rfx6 comme un acteur majeur de la différenciation et de la maturation des cellules endocrines pancréatiques. Ceci est également confirmé par Liao et al. (Liao et al, 2013) qui étudient l'expression de micro-ARN lors de la différenciation de cellules embryonnaires souches humaines en cellules beta *in vitro*. Ils montrent que l'expression de Rfx6 est régulée par les micro-RNA miR30d et let-7e lors de la différenciation terminale des progéniteurs pancréatiques en cellules beta et qu'elle augmente spécifiquement lors des derniers stades de différenciation.

Chez l'homme, diverses mutations du gène Rfx6 sont identifiées chez 9 patients et résumées dans le **Tableau 2**. En 2010, Smith et al. (Smith et al, 2010) remarquent que le phénotype des souris mutantes pour Rfx6 est similaire à celui de patients atteints de diabète

néonatal, d'atrésie de l'intestin grêle et de malabsorption intestinale. Les auteurs regroupent alors ces symptômes sous le terme « Mitchell Riley syndrome » (OMIM#601346) suite à la découverte de cette pathologie en 2004 par Mitchell et al. (Mitchell et al, 2004). Le séquençage du locus Rfx6 a ainsi révélé différentes mutations dans le gène Rfx6 qui semblent être à l'origine de ce syndrome chez l'homme. La diversité des symptômes observés peut notamment être expliquée par le rôle de Rfx6 à la fois lors du développement du pancréas endocrine et dans la cellule beta adulte ((Chandra et al, 2014; Piccand et al, 2014) et paragraphe Résultats A.3.e) mais également dans l'intestin (Suzuki et al, 2012).

Mutations	Phénotype	Localisation	Référence
Perte d'un site donneur (intron 6, remplacement d'un T par un G) et d'un site accepteur (intron1, remplacement d'un A par un G) d'épissage, hétérozygote	Diabète néonatal, atrésie intestinale, hypoplasie de la vésicule biliaire, diarrhées, atrésie hépatique, vivant à 9 ans	Intron 1 et 6	Mitchell et al. 2004
Perte de site donneur d'épissage, remplacement d'un T par un C (2 patients)	Diabète néonatal, atrésie intestinale, hypoplasie de la vésicule biliaire, diarrhées, atrésie hépatique, nutrition parentérale, décédé à 5 et 6 mois	Intron 2	Mitchell et al. 2004
R181Q	Diabète néonatal, atrésie intestinale, problèmes hépatiques, décédé à 2,5 mois	Exon 4	Smith et al. 2010
S217P	Diabète néonatal, atrésie intestinale, agénésie de la vésicule biliaire, diarrhées, vivant à 6 ans	Exon 6	Chappell et al. 2008
Délétion avec décalage du cadre de lecture	Diabète néonatal, atrésie intestinale, hypoplasie de la vésicule biliaire, nutrition parentérale, décédé à 3 mois	Exon 7	Martinovici et al. 2010
K260T (perte d'un site donneur d'épissage)	Diabète néonatal, atrésie intestinale, agénésie de la vésicule biliaire, diarrhées, problèmes hépatique, nutrition parentérale, décédé à 5 mois	Exon 7	Concepcion et al. 2014
Délétion de 9 nt et Insertion d'un G (perte d'un site accepteur d'épissage)	Diabète néonatal, atrésie intestinale, agénésie de la vésicule biliaire, diarrhées, nutrition parentérale, vivant à 21 mois	Intron 7	Spiegel et al. 2011
V506G	Diabète néonatal, atrésie intestinale, agénésie de la vésicule biliaire, nutrition parentérale, vivant à 6 ans	Exon 14	Chandra et al. 2014
R726X + R866X (2 patients)	Diabète diagnostiqué à 3 et 6 ans, atrésie intestinale, agénésie de la vésicule biliaire, vivants à 9 ans	Exon 17 et 18	Sansbury et al. 2015

Tableau 2 : Mutations du gène Rfx6 humain et phénotypes associés. La localisation de la mutation dans le gène Rfx6 humain et la publication scientifique de référence sont indiquées.

Matériels et méthodes

Matériels et méthodes

A. Matériel biologique

A.1. Culture et transfection cellulaire

A.1.a. Cellules Min6b1

Les cellules Min6 ont été établies à partir de tumeurs des cellules beta qui se forment dans des souris transgéniques exprimant l'antigène T du virus SV40 sous le contrôle du promoteur de l'insuline (Miyazaki et al, 1990). Les cellules de la sous lignée Min6b1 utilisée dans ce travail secrètent de l'insuline suite à des stimuli de glucose de façon dose-dépendante (Lilla et al, 2003). Les cellules Min6b1 miment les propriétés des cellules beta puisqu'elles gardent un phénotype différencié en culture. Elles sont cultivées à 37 °C et à 5 % de CO₂ en milieu DMEM (Dulbecco Minimum Eagle Medium 11965-084 Gibco Life Technology) comprenant 25mM de glucose, et supplémenté avec 15% (v/v) de SVF (sérum de veau fœtal), 71µM de beta-mercaptoéthanol, 1% (m/v) de pénicilline et 1% (m/v) de streptomycine.

A.1.b. Cellules HEK293T

Les cellules HEK293T ont été générées en 1973 par transformation de cellules rénales embryonnaires humaines par un adénovirus humain de type 5 (Graham et al, 1977). Ces cellules se cultivent et sont transfectées aisément, elles ont également un taux de croissance élevé, et constituent ainsi un modèle d'étude idéal. Ces cellules sont cultivées à 37 °C et à 5 % CO₂ en milieu DMEM comprenant 5,5mM de glucose et supplémenté avec 10% (v/v) de SVF (sérum de veau fœtal), 1% (m/v) de pénicilline et 1% (m/v) de streptomycine.

A.1.c. Transfection cellulaire

Les cellules sont transfectées soit avec de l'ADN plasmidique soit avec des ARN interférents. Après un jour d'incubation en milieu de culture sans antibiotique, les cellules sont transfectées en utilisant de la lipofectamine (Lipofectamine™ 2000, Invitrogen), qui permet, par formation de liposomes cationiques contenant l'ADN ou l'ARN d'intérêt, de faire entrer ce dernier dans la cellule par endocytose. Après une incubation de 24h à 37°C et à 5% de CO₂, le milieu de transfection est remplacé par le milieu de culture habituel. Un jour plus tard, les expériences souhaitées peuvent être réalisées.

ADN plasmidique

La transfection des cellules est réalisée à 60% de confluence, avec 60µL de lipofectamine et 24µg d'ADN plasmidique dans des boîtes de culture cellulaire de 10 cm² (BD Falcon 100x20mm) selon le protocole du fournisseur Invitrogen. Pour les expériences d'immunofluorescence décrites dans ce manuscrit, la transfection des cellules, cultivées directement sur des lamelles, sera réalisée dans des boîtes de culture cellulaire de 4 puits (Nunclon™ Δ Surface) à l'aide de 0,8µg d'ADN plasmidique et 2µL de lipofectamine par puits.

ARN interférence

Les cellules Min6b1 expriment de façon endogène Rfx6, ce qui permet l'utilisation d'ARN interférents dirigés contre le gène de Rfx6 dans ce système expérimental. Les cellules Min6b1 sont transfectées par des ARN interférents (siRNA) dirigés contre Rfx6 (Mouse Rfx6 (320995) siRNA – SMARTpool L- L-05699901 Dharmacon ou ON-TARGETplus non targeting pool D-00181010 Dharmacon ON-TARGETplus). La transfection des cellules Min6b1 est réalisée à 60% de confluence, dans des boîtes de culture cellulaire de 6 puits (Multiwell Falcon®) à l'aide de Lipofectamine RNAimax selon les indications du fournisseur (Lipofectamine RNAimax 13778030). Différentes concentrations de siRNA ont été testées et c'est à 75nM de siRfx6 que l'expression de Rfx6 est diminuée de manière significative et que les expériences souhaitées sont réalisées.

A.2. Purification des ilots de Langerhans

La purification des ilots de Langerhans débute par l'anesthésie de souris CD1 (Charles River) à l'aide d'une injection de kétamine et xylazine (kétamine 100mg/mL + xylazine 20mg/mL). L'abdomen de la souris est alors passé à l'éthanol 70% (v/v) puis ouvert. Dans un second temps et après rinçage au PBS (Phosphate Buffer Saline : NaCl 0,8%, Na₂HPO₄ 0,5%, KCl 0,02%, KH₂PO₄ 0,02% (m/v) le péritoine est ouvert pour permettre l'accès aux organes internes. Après dégagement de la zone comportant la vésicule biliaire et le canal cholédoque, ce dernier est clampé au niveau de sa jonction dans l'intestin grêle. 5mL d'une solution de collagénase (Collagénase type 5 C9263 Sigma 2mg/mL, HEPES Gibco Life Technology 10µM, DNase1 DN25 Sigma 3,3mg/L, RNAsine Roche 13,2U/mL, HBSS [MgCl₂ 0,49mM ; MgSO₄ 0,41mM ; KCl 5,36mM ; KH₂PO₄ 0,44mM ; NaCl 136,88mM ; Na₂HPO₄ 0,33mM ; CaCl₂ 1,26mM ; NaHCO₃ 4,17mM ; Glucose 5,55mM]) sont ensuite injectés lentement dans la vésicule biliaire ou à défaut directement au départ du canal cholédoque pour permettre le remplissage du pancréas. Le pancréas injecté est ensuite retiré en entier de la souris qui sera sacrifiée. Le pancréas est incubé pendant 38 min dans 2 mL supplémentaires de solution de collagénase pour permettre la digestion du tissu. Chaque

échantillon est ensuite agité énergiquement environ 5 secondes pour finaliser une bonne dissociation du tissu. La réaction de digestion est stoppée par ajout de 30mL de tampon QB à 4°C (Quenching buffer : HEPES Gibco Life Technology 25mM, BSA MP Biomedicals 0,5% (m/v), HBSS) puis le tube est à nouveau agité. Le tissu est alors centrifugé 2 min à 241g et 4°C. Un second lavage avec 30mL de tampon QB est réalisé. Le culot est ensuite resuspendu dans 10mL de tampon QB puis filtré grossièrement pour permettre l'élimination des gros débris de tissu. Un second volume de 10mL de tampon QB est utilisé pour rincer le filtre. Après centrifugation durant 2 min à 241g et 4°C, le culot bien sec est resuspendu dans une solution d'Histopaque 1100 à 4°C (45% Histopaque 1077 Sigma 10771, 55% Histopaque 1119 Sigma 11191). Après centrifugation 20 min à 377g et 4°C, les ilots pancréatiques sont maintenant dans le surnageant, auquel il est ajouté 25mL de tampon QB. Les ilots sont culottés 4 min à 377g et 4°C puis lavés avec un second volume de tampon QB de 20mL. Après une dernière centrifugation de 3 min à 241g et 4°C, les ilots sont resuspendus dans 10mL de milieu de culture (RPMI 1640 Gibco 11875-093, HEPES 25mM, L-glutamine 2mM, 10% (v/v) FCS, 1% (m/v) pénicilline, 1% (m/v) streptomycine). Enfin, les ilots sont récoltés sous une loupe à l'aide d'une pipette (p200µL) dans le but d'éliminer un maximum de tissu exocrine encore présent, puis repris dans 10mL de milieu de culture. Après 24h de culture à 37°C, 5% CO₂ le milieu de culture est changé et 48h plus tard, les ilots peuvent être dénombrés. Une moyenne de 150 à 250 ilots pancréatiques purifiés par souris est obtenue après 3 jours de culture.

B. Immunoprécipitation de chromatine

B.1. Immunoprécipitation de chromatine à partir de cellules Min6b1

B.1.a. Préparation de chromatine

Des cellules Min6b1 à 80% de confluence (cultivées dans des boîtes de culture cellulaire de 10 cm² (BD Falcon 100x20mm)), sont soumises à un premier lavage au PBS (Phosphate Buffer Saline : NaCl 0,8%, Na₂HPO₄ 0,5%, KCl 0,02%, KH₂PO₄ 0,02% (m/v)) puis les cellules sont fixées 10 min à température ambiante et sous agitation dans une solution de PFA 1% (Paraformaldéhyde Electron Microscopy Sciences 16% 15710) dans le tampon PBS. Après élimination de la PFA, les cellules sont incubées 5 min à température ambiante dans une solution de PBS-Glycine (PBS ; Glycine 125mM (AnalaR NormaPur)) sous agitation. Deux lavages au PBS sont effectués puis les cellules sont décollées mécaniquement. La suspension cellulaire est alors centrifugée 5 min à 671g et 4°C. Le culot est dissocié, puis du tampon de gonflement (Swelling buffer : 25mM Tris HCl pH8, 1,5mM

MgCl₂, 10mM KCl, 0,1% NP40 Sigma, 1mM DTT, Protéase Inhibitor Cocktail PIC 1x Roche) est ajouté (6x le volume du culot). La solution est passée au Dounce de type A (25 passages) permettant de broyer les cellules et leurs membranes sans que les noyaux ne soient abimés. Après homogénéisation des cellules, aspiration et refoulement à l'aide d'une pipette (p1000 µL), le dénombrement des cellules est effectué à l'aide d'une Cellule de Neubauer.

La suspension cellulaire est alors centrifugée à 671g 5 min 4°C puis le culot est resuspendu dans du tampon de lyse (1% SDS, 10mM EDTA, 50mM Tris HCl pH 8, PIC 1x) pour atteindre une concentration de 2x10⁷ cellules/mL. Une incubation des cellules de 30 min minimum est réalisée à 4°C dans ce tampon sous agitation, ce qui permet la lyse nucléaire. Les échantillons sont ensuite soniqués à l'aide d'un sonicateur Covaris E210 à une concentration de 2.10⁷ cellules/mL et 500µL de volume par tube pendant 4 min. Puis la préparation soniquée est centrifugée pendant 10 min à 4°C et 10621g. Le surnageant, qui est conservé à 4°C, contient la chromatine fragmentée qui sera utilisée pour la suite du protocole.

L'efficacité de sonication est tout d'abord vérifiée sur un aliquot (50µL) de chromatine fragmentée. Pour cela, l'aliquot est incubé en présence de 200mM de NaCl et de 100µg/mL de protéinase K (Sigma) pendant une nuit à 65°C et sous agitation (1,6g). Le lendemain, une purification à l'aide de phénol-chloroforme et précipitation à l'éthanol de l'ADN est réalisée. L'échantillon est resuspendu dans du Tris 10mM en fin de purification. 500ng à 2µg d'ADN sont alors déposés sur gel d'agarose et la taille des fragments de chromatine peut être vérifiée. La concentration de cette chromatine purifiée est également estimée au Nanodrop 2000 et permettra d'estimer la quantité de chromatine qui sera choisie au moment de l'immunoprécipitation.

Des protéines A ou G-Sépharose (GE Healthcare) sont ensuite saturées avec 100µg/mL de BSA et 400µg/mL d'ARNt de levure pendant au moins 3h sous agitation à 4°C. Celles-ci sont resuspendues dans une solution contenant 25µg/mL de BSA et 400µg/mL d'ARNt de levure. La chromatine fragmentée (450µL) conservée à 4°C est enfin incubée pendant 4 à 16h avec 25µL de protéines A ou G-Sépharose saturées.

B.1.b. Immunoprécipitation

Après centrifugation de 3 min à 3824g et 4°C, le surnageant de chromatine est dilué à 50µg/mL dans le tampon de CHIP (0,01% SDS Euromedex, 1,1% Triton, 1,2mM EDTA, 16,7mM Tris HCl pH 8, 167mM NaCl, PIC 1x) à l'aide de la concentration estimée auparavant. Un aliquot de 10% de chromatine avant immunoprécipitation est conservé à 4°C (=INPUT). Une immunoprécipitation de chromatine est réalisée sur 50µg de chromatine

équivalente à 4.10^6 cellules Min6b1 pendant une nuit en présence de l'anticorps d'intérêt (anti-Rfx6 2767 ou anti-HA 12CA5, IGBMC 1/200^e).

B.1.c. Purification de l'ADN immunoprécipité

La chromatine immunoprécipitée est ensuite incubée pendant 4h à 4°C avec 50µL de protéines A ou G Sépharose. Après centrifugation pendant 5 min à 956g à 4°C, le surnageant, qui constitue la fraction non liée, est éliminé. Les protéines A ou G Sépharose comportant la chromatine d'intérêt sont ensuite lavées à trois reprises pendant 5 min à l'aide de 1 mL de chacun des tampons de lavages I (0,1% SDS, 1% Triton, 2mM EDTA, 20mM Tris pH 8, 150mM NaCl, PIC 1x), II (0,1% SDS, 1% Triton, 2mM EDTA, 20mM Tris pH 8, 500mM NaCl, PIC 1x) et III (1mM EDTA, 10mM Tris pH 8). Enfin, les échantillons sont incubés dans du tampon d'éluion (1% SDS, 100mM NaHCO₃) pendant 30 à 60 min à 50°C sous agitation (3g). Après centrifugation pendant 3 min à 3824g à 4°C, le surnageant, qui constitue la fraction liée, est récupéré. La fraction liée ainsi que l'input sont ensuite incubés en présence de 200mM de NaCl, 12,5mM d'EDTA et 100µg/mL de protéinase K pendant une nuit à 65°C sous agitation (1,6g).

Le lendemain, l'ADN est purifié à l'aide de phénol-chloroforme et précipité à l'isopropanol en présence de 20µg/mL de Glycogène. L'échantillon est resuspendu en fin de purification dans de l'eau ultrapure. La concentration de cette chromatine purifiée est estimée au Nanodrop 2000 ou au Qubit (Life Technology).

B.2. Immunoprécipitation de chromatine à partir d'ilots de Langerhans de souris

B.2.a. Préparation de chromatine

La fixation des ilots pancréatiques purifiés, selon le protocole indiqué dans le paragraphe Matériels et méthodes A.2, est réalisée directement après les 72h de culture. Pour cela, 400 à 1200 ilots (ou IEQ pour ilots équivalents) sont récoltés dans des tubes de 1,5mL (Eppendorf LowBind Tube 022431021). Après centrifugation à 100g pendant 1 min ils sont lavés trois fois dans du PBS 1x froid (NaCl 0,8%, Na₂HPO₄ 0,5%, KCl 0,02%, KH₂PO₄ 0,02% (m/v)) puis fixés 10 min à température ambiante et sous agitation dans une solution de PFA 1% (Electron Microscopy Sciences PFA16% 15710) dans le tampon PBS. La réaction de fixation est stoppée par ajout de 133mM final de Glycine (AnalaR NormaPur) sous agitation pendant 5 min. La suspension d'ilots est alors centrifugée 2 min à 1000g et 4°C puis le culot est lavé deux fois au PBS froid. Un dernier lavage au PBS + PIC 1x (Roche) est effectué avant centrifugation de 2 min à 2000g et 4°C. Le culot d'ilots pancréatique est congelé à l'azote liquide et peut être conservé à -80°C.

Le culot de 400 à 1200 ilots est ensuite resuspendu dans 500µL de tampon de lyse (1% SDS, Triton 2%, 100mM NaCl, 1mM EDTA, 10mM Tris HCl pH 8, PIC 1x) et incubé sur glace pendant 15 min avec agitation régulière. Les ilots sont ensuite passés à travers une aiguille 30G jusqu'à ce que plus aucun ilot ne soit visible à l'œil nu. Les échantillons sont ensuite soniqués à l'aide d'un sonicateur Covaris E210 à une concentration 400 à 1200 ilots dans 500µL de volume par tube pendant 5 min. Puis la chromatine soniquée est centrifugée pendant 5 min à 4°C et 20817g. Le surnageant, qui est conservé à -80°C, contient la chromatine fragmentée qui sera utilisée pour la suite du protocole.

L'efficacité de sonication est tout d'abord vérifiée sur un aliquot (100IEQ) de chromatine fragmentée. Pour cela, l'aliquot est incubé en présence de 200mM de NaCl et de 0,3µg/mL de protéinase K (Sigma) pendant une nuit à 65°C et sous agitation (1,6g). Le lendemain, une purification à l'aide de phénol-chloroforme et une précipitation à l'éthanol de l'ADN sont réalisées. L'échantillon est resuspendu dans 20µL d'eau ultrapure en fin de purification puis incubé à 37°C pendant 30 min sous agitation (3g). L'ADN est alors déposé sur gel d'agarose et la taille des fragments de chromatine peut être vérifiée. La concentration de cette chromatine purifiée est également estimée au Qubit (Life Technology).

Des protéines A/G-Sépharose (GE Healthcare) sont ensuite lavées à trois reprises avec du tampon de CHIP (1 volume de tampon de lyse (1% SDS, Triton 2%, 100mM NaCl, 1mM EDTA, 10mM Tris HCl pH 8, PIC 1x) pour 4 volumes de tampon de dilution (Triton 0,75%, 140mM NaCl, 1mM EDTA, 50mM HEPES pH8, 0,1% Sodium deoxycholate (m/v), PIC 1x)) après centrifugation 2 min à 400g et 4°C.

B.2.b. Immunoprécipitation

La chromatine fragmentée, conservée à -80°C, est décongelée lentement sur glace et diluée au 1/4 dans le tampon de dilution (Triton 0,75%, 140mM NaCl, 1mM EDTA, 50mM HEPES pH8, 0,1% Sodium deoxycholate (m/v), PIC 1x). 20µL de protéines A/G-Sépharose sont ajoutées par échantillon et une incubation de 2h à 4°C sous agitation est réalisée. Après centrifugation de 2 min à 400g et 4°C, un aliquot de 10% de chromatine avant immunoprécipitation est conservé à 4°C (=INPUT). L'immunoprécipitation de chromatine a été testée sur différentes quantités d'ilots pancréatiques (1000, 5000 ou 10000 IEQ) et avec différentes quantités d'anticorps anti-Rfx6 (2767 IGBMC) (1/660^e et 1/200^e). C'est avec exactement 10867 IEQ et une concentration d'anticorps anti-Rfx6 de 1/200^e pendant une nuit à 4°C sous agitation que les meilleurs résultats ont été obtenus.

B.2.c. Purification de l'ADN immunoprécipité

Enfin, après blocage des protéines A/G-Sépharose avec une concentration finale en BSA de 5mg/mL dans le tampon de ChIP pendant une nuit, un volume de 30µL de protéines A/G-Sépharose bloquées est ajouté par immunoprécipitation.

Après centrifugation pendant 2 min à 400g et 4°C, le surnageant qui constitue la fraction non liée est éliminé. Les protéines A/G Sépharose liées à la chromatine d'intérêt sont ensuite lavées une fois à l'aide de 1mL de chacun des tampons de lavages I (0,1% SDS, 1% Triton, 2mM EDTA, 20mM Tris pH 8, 150mM NaCl), II (0,1% SDS, 1% Triton, 2mM EDTA, 20mM Tris pH 8, 500mM NaCl) et III (1mM EDTA, 10mM Tris pH 8, 0,25M LiCl, 1% Sodium deoxycholate (m/v), 1% NP40). Après trois lavages successifs au tampon TE (1mM EDTA, 10mM TrisHCl pH 8) les échantillons sont incubés dans 150µL de tampon d'éluion (1% SDS, 100mM NaHCO₃) pendant 15 min à température ambiante sous agitation puis centrifugés pendant 3 min à 500g. Le surnageant est récupéré. Une seconde éluion et une centrifugation pendant 2 min à 2000g sont réalisées puis les deux éluats sont rassemblés et constituent la fraction liée.

La fraction liée ainsi que l'INPUT sont enfin incubés en présence de 200mM de NaCl et de 0,3µg/mL de protéinase K (Sigma) pendant une nuit à 65°C et sous agitation (1,6g). Le lendemain, une purification à l'aide de phénol-chloroforme et une précipitation à l'éthanol de l'ADN sont réalisées. L'échantillon est resuspendu dans 15µL d'eau ultrapure en fin de purification puis incubé à 37°C pendant 30 min sous agitation (3g). La concentration de cette chromatine immunoprécipitée et purifiée est estimée au Qubit (Life Technology).

C. Analyses bio-informatiques

C.1. Préparation des bibliothèques d'ADN et séquençage

Suite à la préparation d'ADN immunoprécipité, celui-ci est confié à la plateforme Biopuces et Séquençage de l'IGBMC (<http://www-microarrays.u-strasbg.fr/index.php>) pour permettre le séquençage des échantillons. Cette étape démarre par la préparation des bibliothèques d'ADN à l'aide du kit NEXTflex ChIP Seq (5143-02, Bioo Scientific, V12.10) à partir de 10ng d'ADN. Les fragments d'ADN sont ensuite liés à des adaptateurs de séquençage possédant un index (code barre d'au moins 6nt) et permettant leur identification parmi plusieurs échantillons (NEXTflex ChIP-Seq Barcodes - 6, 514120, Bioo Scientific). Enfin, l'ADN est enrichi par PCR (2 min à 98°C ; [30 sec à 98°C, 30 sec à 65°C, 60 sec à 72°C] x 14 cycles; 4 min à 72°C) et purifié à l'aide de billes Agencourt AMPure XP (A63881, Beckman). Après vérification du profil de la bibliothèque par électrophorèse capillaire (Bioanalyzer, Agilent 2100) l'ADN est chargé sur une puce (ou flowcell) à 8pM et la

génération d'ensembles de séquences à l'aide du Cbot (Illumina) est réalisée. Le séquençage de 50pb de part et d'autre de chaque fragment d'ADN généré est effectué par le HiSeq2500 (Illumina). Enfin, la première étape d'analyse à l'aide de RTA 1.17.20/CASAVA 1.8.2 (Illumina) permet l'extraction des intensités des ensembles de séquences à partir des images, l'identification des bases puis le filtrage des séquences en fonction de leur qualité et la suppression des dimères d'adaptateur. La quantité seuil de séquences totales à atteindre est de 25 millions pour permettre une analyse bio-informatique de qualité.

L'ensemble des analyses bio-informatiques décrites par la suite et réalisées sur les données de ChIP Seq produites au cours de ma thèse ont été réalisées en étroite collaboration avec Tao Ye, ingénieur de la plateforme Biopuces et Séquençage de l'IGBMC. Les analyses d'occurrence et de co-occurrence ont été réalisées avec l'aide de Stéphanie Legras, doctorante de la plateforme Biopuces et Séquençage de l'IGBMC.

C.2. Identification et annotation des pics de ChIP Seq

L'analyse bio-informatique des données de ChIP Seq débute par l'alignement des séquences sur le génome de souris (version mm9 Ensembl) à l'aide de l'outil Bowtie 1.0.0 (Langmead, 2010; Langmead et al, 2009) qui permet un alignement de larges ensembles de séquences sur de grands génomes. L'identification des régions enrichies en séquences ou pics de ChIP Seq est permise par l'utilisation du logiciel MACS 1.4.2. (Model-based Analysis of ChIP-Seq) (Zhang et al, 2008). Ce logiciel permet de reconstruire empiriquement les fragments de séquences dans toute leur longueur et les utilise pour améliorer la résolution spatiale des sites de liaison prédits. MACS utilise également une distribution de Poisson dynamique pour capturer efficacement les biais locaux dans la séquence du génome, ce qui permet une prédiction plus sensible et plus robuste (Feng et al, 2012; Feng et al, 2011). Les gènes Ref Seq, les caractéristiques génomiques ainsi que les coordonnées correspondantes ont été téléchargés à partir du navigateur UCSC (<http://genome.ucsc.edu/>).

L'annotation des pics de ChIP Seq est réalisée à l'aide des programmes GPAT (Krebs et al, 2008) ou HOMER (<http://homer.salk.edu/homer/>),(Heinz et al, 2010)). L'association des pics de ChIP Seq avec leurs gènes correspondants détermine la distance entre le sommet du pic et le site d'initiation de la transcription du gène le plus proche (identifié dans les bases de données Ensembl67 (expériences réalisées dans les îlots pancréatiques) ou Ref Seq (expériences réalisées dans les cellules Min6b1)). Puis la région d'occupation du génome par le sommet du pic est déterminée (site d'initiation de la transcription ou TSS (entre -1kb et +100pb), site de terminaison de la transcription (entre -100pb et +1kb), séquence codante CDS, 3'UTR, 5'UTR, exon, intron, région intergénique ou gène non codant).

C.3. Recherche de motifs de séquences

Enfin, l'identification de motifs de séquences enrichies sous les pics de CHIP Seq est réalisée par une analyse MEME-CHIP qui peut être conduite sur un maximum de 50Mb de séquences nucléotidiques (<http://meme.ebi.edu.au/meme/index.html>) (Ma et al, 2014; Machanick & Bailey, 2011). Cette analyse permet ainsi la construction de nouveaux motifs consensus (outils MEME (Bailey et al, 2010; Bailey & Elkan, 1994) et DREME), l'identification de leur position sous les pics (outil CentriMO (Bailey & Machanick, 2012)) ainsi que l'identification des sites connus similaires (outil TomTom) dans les bases de données telles que Jaspar 2014 (<http://jaspar.genereg.net/>). Dans les analyses effectuées lors de ce travail, ce sont dans chaque cas les 100pb centrées autour des sommets de 1000 pics aléatoires parmi le jeu de données d'intérêt qui sont analysées. Si la totalité des pics identifiés pour le jeu de données d'intérêt est inférieure à 1000, c'est alors les 100pb centrées autour des sommets de la totalité des pics qui sont analysées. Pour l'analyse du dernier jeu de données (CHIP RFX6 dans les ilots pancréatiques de souris), une mise à jour du logiciel permettant une prise en charge d'un plus grand nombre de séquences nous a permis de réaliser cette recherche dans la totalité des pics identifiés à l'aide des outils MEME et TomTom.

C.4. Analyse d'occurrence et de co-occurrence de motifs

Le but de cette analyse est de détecter si certains motifs connus sont enrichis dans les séquences analysées, en comparaison de ce qui aurait été obtenu par chance. L'analyse réalisée est décrite en **Figure 13**. Les coordonnées chromosomiques des séquences d'intérêt (500pb centrées sous le sommet des pics pour l'analyse dans les cellules Min6b1 et 200pb centrées sous le sommet des pics pour l'analyse dans les ilots pancréatiques) servent d'apport pour cette analyse. Pour démarrer, les motifs connus sont détectés par l'outil FIMO (Grant et al, 2011) dans les séquences d'intérêt, puis les occurrences et co-occurrences (deux motifs présents dans la même séquence co-occurrent) sont dénombrées. Cette analyse est réalisée une fois sur l'échantillon d'intérêt et n fois sur une sélection de séquences au hasard. La significativité d'occurrence ou de co-occurrence d'un motif est évaluée par le Z score. Un Z score élevé signifie que l'occurrence du motif est plus fréquente que ce qui aurait été obtenu par chance, inversement un Z score peu élevé signifie que l'occurrence du motif est moins fréquente que ce qui aurait été obtenu par chance.

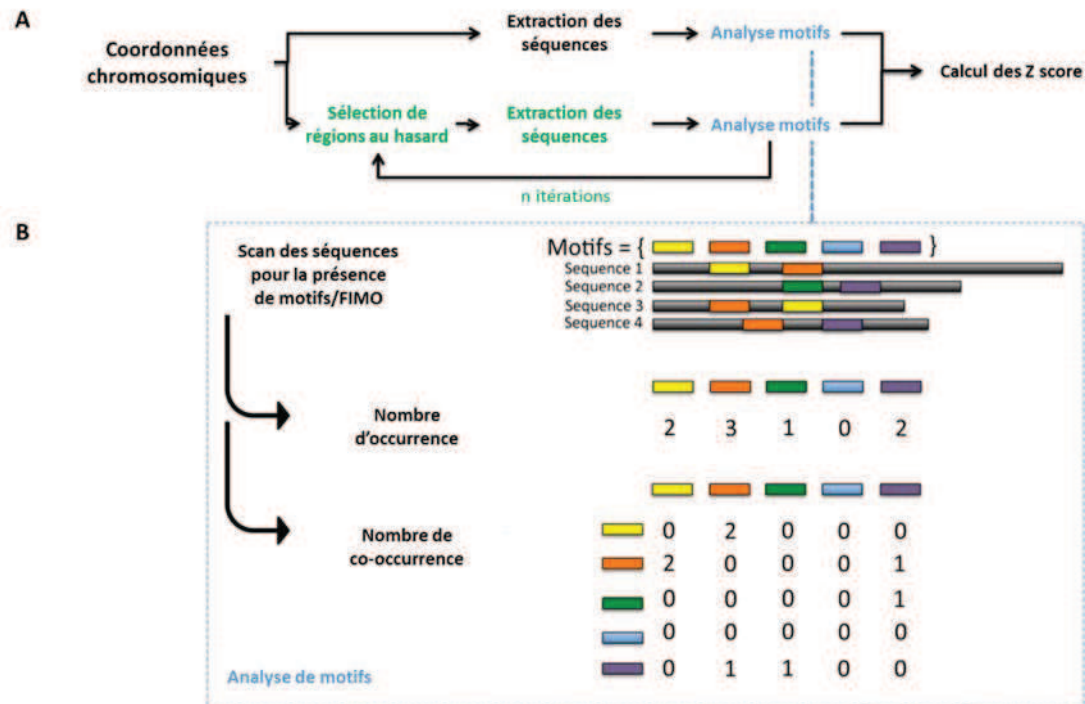


Figure 13 : Diagramme d'analyse d'occurrence et de co-occurrence. Les séquences correspondantes aux coordonnées chromosomiques des pics sont extraites, puis le nombre d'occurrence et de co-occurrence de motifs connus sont comptés. (Figure adaptée de Stéphanie Legras)

D. Test fonctionnels

D.1. Tests de transactivation

Pour les expériences de transactivation, deux éléments régulateurs d'intérêt ont été clonés en amont de la luciférase Firefly dans le vecteur pGL4.23[luc2/minP]. Les stratégies de clonage développées sont présentées en **Figure 14** et **Figure 15**.

Le premier vecteur luciférase construit est le pGL4.23-4Xbox2 (**Figure 14**). Le vecteur pGL4.23 est dans un premier temps linéarisé par digestion de restriction Kpn1 et Nhe1. Simultanément, l'insert double brin 4Xbox2 est créé par hybridation des oligonucléotides sens et antisens. Le vecteur pGL4.23[luc2/minP] et l'insert 4Xbox2 sont ensuite ligués pour obtenir le vecteur pGL4.23-4Xbox2. Les clones recombinants sont identifiés par PCR à l'aide des oligonucléotides screen pcr pGL4.23 sens et antisens (flèches oranges) et présentent un produit de 246pb. Enfin, la présence de l'insert est vérifiée par séquençage de l'échantillon chez GATC Biotech.

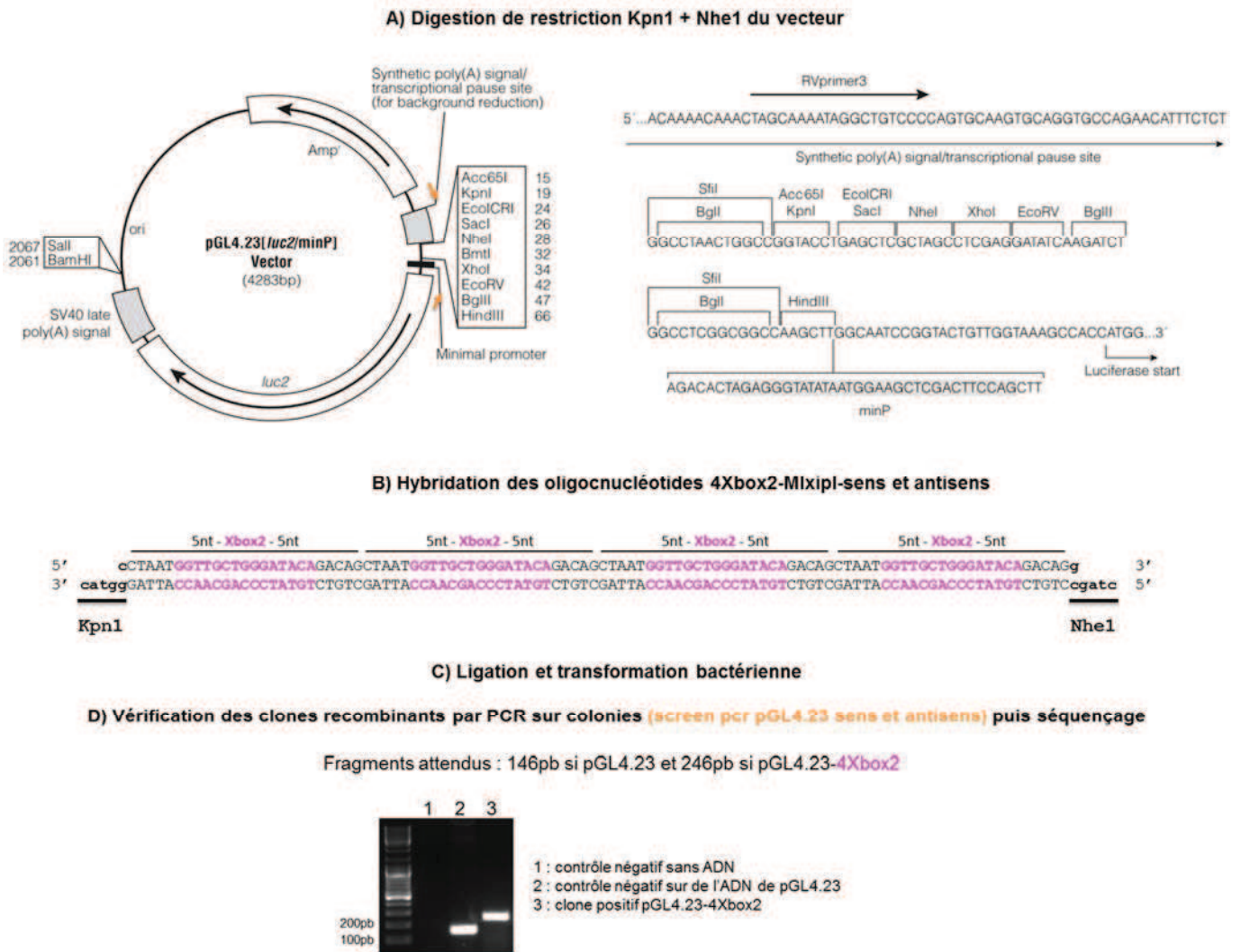
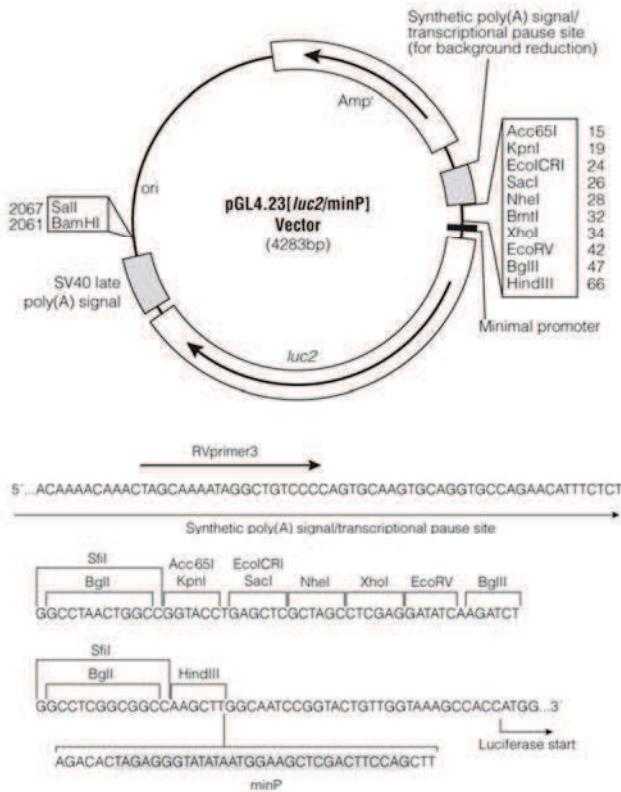


Figure 14 : Stratégie de clonage de 4 copies de la Xbox2 de Mixipl identifiée par ChIP Seq dans le vecteur pGL4.23[luc2/minP]. Les différentes étapes de cette stratégie sont présentées dans les points A) à D). L'identification des clones recombinants est réalisée par PCR sur colonie et séquençage à l'aide des amorces screen pcr pGL4.23 sens et antisens indiquées en orange.

Le second vecteur luciférase construit est le pGL4.23-Xbox123 (**Figure 14**). Le vecteur pGL4.23 est dans un premier temps linéarisé par digestion de restriction Kpn1 et Nhe1. Simultanément, l'insert Xb123 est créé par PCR sur de l'ADN génomique de souris à l'aide des oligonucléotides Mixipl-Xb123 sens et antisens (flèches oranges) possédant respectivement les sites de restriction Kpn1 et Nhe1. Après digestion de restriction, le vecteur pGL4.23[luc2/minP] et l'insert Xb123 sont ensuite ligués pour obtenir le vecteur pGL4.23-Xbox123. Les clones recombinants sont identifiés par restriction et la présence de l'insert est vérifiée par séquençage de l'échantillon chez GATC Biotech.

A) Digestion de restriction Kpn1 + Nhe1 du vecteur

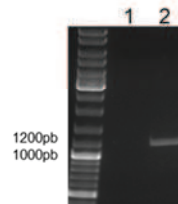


B) Préparation de l'insert Xb123

Région de l'intron 1 de Mxipl

```
GAGAGCTGTCTCAAAAACGGAGGCAGATTCATTTGAGGAACACGTTTGAGTTTGT
CCTCTGGTCTCACAGAGTGTATGCACCCACATPCTCCCTCTCTCCACTTCTGAGG
TTTTCTGGTTCACCCCATTCACACAGCTTTTGAACCTTGGCCCTAGACTTTTTGCA
CCAAACCGGACTCTTCTCTCACGTGCATTCAACCAAATCCTTCTCTCGGGACTTGA
TATTGTCACTGACTCATCCCCTCATCCTAGGCTACCAACCCTCTCGGGGAGTTGA
GGAGGGACAGCAGAAGGCTGGGTAGGACCTGGGGGAACAGAACCTAATGGTTGCT
GGGATACAGACAGAGTCACAAACCTAGGAGTGTCCCCATTTACTCTATATGACAA
CAGGAGCTGGCCAGGAAGGATTAGTGTCCCAACCCAATCATGTTCACTTCTTGA
CAGATGGTATTTATCTAGCTATGGGAGGTAGGGGACCAGCCATCAGGCTAGGGTGGC
CTATTTGGGTAATGAAGCCATTGTCTCCTGAGGTGGCAGCCCTGCCACCCACGT
CCCACACCTGGCAGCCACCATTAGCTCAAAGCAAATACTGGAGCTTGGAGACGAGA
AAGGAGAGGAGGCCAATGACCCATGCACCTCAAGGTAGTGCAGGAAGAGGAGGG
CTCAGATGCTACCCCTCATCCCTTACCCCTCATCCCTCACCCAGGGGAGCTACCCG
AGAGCAGGCCTTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
TATACAGTCTATACTGTCTATACAGTGTGAGGATGGAATCCAGTCTGTCCAGTTC
CTGGGAAAGCTGCCACTGAGCCACGCCCCAGCCCTCACTGGGGATCTAGAAAGT
GCTCTGTTGGCAGAGCCACTCCAGCCCCAGCCTGACTTCTGGATGGCTAAACC
TTTGAAGCTTTGAGTGAGGAGTTTGAAGTGGCCAGAATGACTCGGAGGCAAGATT
GGTTTGTAGACCATGAGGAAGGAAATCAGACTAAAGATAGGACAGGAAGGTGCC
AAGACCTTGATTGTCTTCATCTTGGCCTCTTGGCCAGCTTCCAGGTACTGATAACT
TAGCCTGTTCTCAGAGGAACCTTGA
```

- PCR sur ADN génomique de souris à l'aide des oligonucléotides Mxipl-xb123 sens et antisens : fragment de 1167 pb
- Digestion du produit de PCR par Kpn1 et Nhe1



1 : contrôle négatif sans ADN
2 : PCR sur ADN génomique de souris

C) Ligation et transformation bactérienne

D) Vérification des clones recombinants restriction et séquençage

Figure 15 : Stratégie de clonage de la région de l'intron 1 de Mxipl contenant trois Xbox dans le vecteur pGL4.23[luc2/minP]. Les différentes étapes de cette stratégie sont présentées dans les points A) à D). En B) La PCR sur ADN génomique de souris est réalisée avec les amorces Mxipl-Xb123 sens et antisens comportant respectivement les sites de restriction Kpn1 et Nhe1 et indiquées en orange.

A partir du vecteur pGL4.23-Xbox123, des mutants avec délétion d'une ou plusieurs Xbox ont été produits par mutagenèse dirigée à l'aide du Kit Quick Change XL Site-directed mutagenesis de chez Agilent selon les indications du fabricant. Pour cela, des oligonucléotides énumérés dans le

Tableau 6 et désignés à l'aide de l'interface Web dédiée d'Agilent ont été utilisés. Ces mutants seront testés en transactivation dans un futur proche.

Les différents vecteurs d'expression de facteurs de transcription qui ont été ou seront testés en transactivation sont présentés dans le **Tableau 3**.

Plasmide	Description	Source
pcDNA3	Commercial	Invitrogen
pcDNA3-HA-mNgn3	Insertion de la séquence codante de HA-mNgn3 dans les sites BamHI et EcoRI du pcDNA3	Cloné au laboratoire
pcDNA3-hRfx6	Insertion de la séquence codante de huRfx6 au niveau du site NotI du pcDNA3 voir partie Résultats B.8.a	Cloné au laboratoire par Mégane Denu et Aline Meunier
pcDNA3-hRfx6 F294V	Mutagenèses dirigées du plasmide pcDNA3-hRfx6 voir partie Résultats B.8.a	Clonés dans le laboratoire par Aline Meunier
pcDNA3-hRfx6 R181Q		
pcDNA3-hRfx6 R578P		
pcDNA3-hRfx6 S217P		
pcDNA3-hRfx6 V506G		
pcDNA3-mE2F1	Fajas et al.,2002 (Fajas et al, 2002)	Don de Lluis Fajas
pHD	Müller et al., 1988 (Muller et al, 1988)	Don de Klaus Kaestner
pHD-mFoxa2	Müller et al., 1988 (Muller et al, 1988)	Don de Klaus Kaestner

Tableau 3 : Liste des plasmides d'expression de facteurs de transcription utilisés pour les tests de transactivation.

Les tests de transactivation sont réalisés en cellules HEK293T et avec l'aide du kit Dual-Luciferase Reporter Assay System de Promega. Les cellules sontensemencées à raison de 20000 cellules par puits dans des plaques de culture de 96 puits (Falcon) dans du

milieu de culture sans antibiotique. 24h plus tard, les cellules sont co-transfectées avec 0,37ng de vecteur codant pour la luciférase Renilla (pGL4.75[hRLuc/CMV], Promega), 125ng de vecteur codant pour la luciférase Firefly sous contrôle des éléments régulateurs d'intérêt (pGL4.23[luc2/minP], Promega) et 10ng de vecteur d'expression codant pour les facteurs de transcription d'intérêt par puits. Après 48h de culture les cellules sont lysées et l'émission de lumière est mesurée selon les instructions du kit. 10µL de lysat cellulaire (sur 20µL au total) sont déposés sur une plaque 96 puits à fond blanc et plat (23300, Berthold Technologies) et l'émission de lumière est mesurée à l'aide du luminomètre lecteur de microplaques Centro XS³ LB960 (Berthold Technologies) équipé de deux injecteurs et associé au logiciel Microwin pour le paramétrage des expériences et l'acquisition des résultats. A l'aide du premier injecteur, 25 µL de réactif LARII sont injectés par puits, la solution est mélangée pendant 2s et la lumière produite par la luciférase Firefly est mesurée pendant 10s. 25 µL de réactif Stop & Glo sont ensuite injectés par puits à l'aide du second injecteur, la solution est mélangée pendant 2s et la lumière produite par la luciférase Renilla est mesurée pendant 10s. Les données sont analysées à l'aide du logiciel Microsoft Excel 2010, et normalisées par les valeurs obtenues pour la luciférase Renilla.

D.2. Tests de gel retard

Les tests de gel retard sont effectués en présence de sonde ADN couplée en 5' au fluorophore Cyanine5.5 dont la longueur d'onde d'excitation maximale est de 673nm et celle d'émission de 707nm. Les sondes utilisées sont listées dans le Tableau 4. Les conditions expérimentales et les sondes ont été désignées et testées en regard des résultats obtenus par les groupes de Mach, Hearing et German (David et al, 1995; Reith et al, 1994; Smith et al, 2010) ayant montré des interactions spécifiques entre RFX1 ou RFX6 et une Xbox présente dans le virus de l'hépatite B. D'autre part, des sondes constituées de la Xbox2 de Mxipl identifiée lors de ce travail ont également été testées.

Sonde	Séquence 5'-3'
Cy55-Xbox2 sens	Cy55-cta <u>atggtt</u> gctgggatacagacag
Xbox2 antisens	ctgtctg <u>tatccc</u> agcaaccattag
Cy55-2Xbox2 sens	Cy55-ctgtctg <u>tatccc</u> agcaaccattagctgtctg <u>tatccc</u> agcaaccattag
2Xbox2 antisens	cta <u>atggtt</u> gctgggatacagacagcta <u>atggtt</u> gctgggatacagacag
Cy55-HBVXbox sens	Cy55-tac <u>cccg</u> ttgcccggcaacggccag
HBVXbox antisens	ctggccg <u>ttgccc</u> ggcaacggggta

Tableau 4 : Liste des sondes ADN utilisées pour les expériences de gel retard. Les nucléotides soulignés constituent les séquences Xbox 2 de Mxipl et la Xbox du virus de l'hépatite B.

Les réactions de liaison sont effectuées entre la sonde ADN à 50nM et 5µL à 10µL de protéine produite par Transcription/Traduction *in vitro* (Kit TnT® Quick Coupled Transcription/Translation System Promega) dans un volume final de 20µL et dans du tampon de liaison (12mM HEPES Gibco Life Technology, 12,5mM MgCl₂, 0,5mM DTT, 0,06mM EDTA, 60mM KCl, 1µg polydIdC Sigma 81349). Après 30 min d'incubation à température ambiante, les échantillons sont séparés sur gel d'Acrylamide 4% natif à 90V pendant 1h à 4°C dans du TBE 0,5x (40 mM Tris-HCl, pH 8,3, 45 mM H₃BO₃, 1 mM EDTA). Les gels sont ensuite révélés au scanner Odyssey (Li-Cor) à 700nm.

Plasmide	Description	Source
pcDNA3	Commercial	Invitrogen
pcDNA3-HA-mNgn3	Insertion de la séquence codante de HA-mNgn3 dans les sites BamHI et EcoRI du pcDNA3	Cloné au laboratoire
pcDNA3-hRfx6	Insertion de la séquence codante de huRfx6 au niveau du site NotI du pcDNA3 voir partie Résultats B.8.a	Cloné au laboratoire par Mégane Denu et Aline Meunier
pcDNA3-huRfx3-Flag	Insertion de la séquence huRfx3-Flag dans les sites HindIII et BamHI du pcDNA3 (à partir du plasmide commercial pUC57-huRfx3-Flag, Genscript)	Cloné au laboratoire par Perrine Strasser
pcDNA3-HA-mRfx6	Insertion de l'étiquette HA en 5' de mRfx6	Cloné au laboratoire
pcDNA3-3HA-hRfx6	Insertion de l'étiquette 3HA au niveau des sites PpuMI et EcoRV du pcDNA3-hRfx6 voir partie Résultats B.9.a et Figure 62	Cloné au laboratoire par Julia Grans

Tableau 5 : Liste des plasmides utilisés pour les tests de gel retard.

E. Oligonucléotides et anticorps

Nom	Séquence 5'-3'	Expérience
Rfx6(5')	TTTAACTCCCACCTGCAGGCC	Stratégie de clonage de pCMVTag2A-Flag-3HA-Rfx6
T3	ATTTAACCTCACTAAAGGGA	Stratégie de clonage de pCMVTag2A-Flag-3HA-Rfx6
PCR1/end-Rfx6	CAGCAGCTTTCTAGAGACTTCTTT	Stratégie de clonage de pCMVTag2A-Flag-3HA-Rfx6-2A-mCherry
PCR1/2A-EcoRI	TGGTCTGGGTTCTCCTCCACGTGCGCTGCCTGCTTCAGCAGG- -CTGAAGTTGGTTGCTCCGCTGCCAGTGTCTCCAGTGTCTGCCCAT	Stratégie de clonage de pCMVTag2A-Flag-3HA-Rfx6-2A-mCherry
PCR2/2A-mC	GGCAGCGGAGCAACCAACTTCAGCCTGCTGAAGCAGGCAGGCG- -ACGTGGAGGAGAACCAGGACCAATGGTGAAGCAAGGGCAGGAGGAT	Stratégie de clonage de pCMVTag2A-Flag-3HA-Rfx6-2A-mCherry
PCR2/end mC-EcoRI	CCGAATCTTACTTGTACAGGTCGTCCAT	Stratégie de clonage de pCMVTag2A-Flag-3HA-Rfx6-2A-mCherry
Flag-EcoRI	CCGAATTCGCGGCCGCCACCAT	Stratégie de clonage de pCMVTag2A-Flag-3HA-Rfx6-2A-mCherry
PCR1 sens	AGAGCTCCTTCAAGCCGGGGCGGCGTCCGAGATCC- -GGGCGGTGGGAAGGATGATTACAAGGATGACGACGATA	Clonage des bras homologues au BAC-Rfx6 RP23 415D6
PCR1 anti-sens	AGAGTGCTTCGTTTTTATTCTTGGACCCCTTAGGACTC- -AGCACCGCAGCCTCGTATGTTGTGTGGAATTGTGAG	Clonage des bras homologues au BAC-Rfx6 RP23 415D6
PCR2 sens	CAGGTTCCCCAGCCCGCCCGCCCGCCCGGGG- -CTCTGGGCTCTACACAGAGCTCCTTCAAGCCGGGGCGGCG	Clonage des bras homologues au BAC-Rfx6 RP23 415D6
PCR2 anti-sens	CCTGCTGAAACGAATGTTCCGACAGAGCATCCTA- -CTTTATGTGATTACAGAGTCTCGTTTTTATTCTTGGGA	Clonage des bras homologues au BAC-Rfx6 RP23 415D6
Xbox1 deletion For	GGCCGGTACCGAGAGGGAGGAGGATTCATT	Mutagenèse dirigée du vecteur pGL4.23-Xbox123
Xbox1 deletion Rev	AATGAATCTGCCTCCCTCTCGGTACCGGCC	Mutagenèse dirigée du vecteur pGL4.23-Xbox123
Xbox2 deletion For	CCCTGGGGGAACAGAACCCTAATGACAGAGGTCAC	Mutagenèse dirigée du vecteur pGL4.23-Xbox123
Xbox2 deletion Rev	GTGACCTCTGTCAATTAGGTTCTGTTCGCCAGGG	Mutagenèse dirigée du vecteur pGL4.23-Xbox123
Xbox3 deletion For	GTAAGTATAACTTAGCCTCTTGTAGCTAGCCTCGAGG	Mutagenèse dirigée du vecteur pGL4.23-Xbox123
Xbox3 deletion Rev	CCTCGAGGCTAGCTCAAGAGGCTAAGTTATCAGTAC	Mutagenèse dirigée du vecteur pGL4.23-Xbox123
Mlxip1-Xb123 sens	CCGGTACCGAGAGACTGTCTCAAAAACG	Stratégie de clonage dans le vecteur pGL4.23
Mlxip1-Xb123 anti-sens	CCGCTAGCTCAAGGTTCTCTGAGAACA	Stratégie de clonage dans le vecteur pGL4.23
screen pcr pGL4.23 sens	AGTGCAGGTGCCAGAACATT	Stratégie de clonage dans le vecteur pGL4.23
screen pcr pGL4.23 anti-sens	AACAGTACCGGATTGCCAAG	Stratégie de clonage dans le vecteur pGL4.23
Génotypage PCR1 F	CCAACTTGCAAAACCCAGATT	Génotypage de la souris transgénique Flag-3HA-Rfx6-2A-mCherry
Génotypage PCR1 R	ATTGGTCTGGGTTCTCCTC	Génotypage de la souris transgénique Flag-3HA-Rfx6-2A-mCherry
Génotypage PCR2 F	CACGCTTCAAAAGCGCACGT	Génotypage de la souris transgénique Flag-3HA-Rfx6-2A-mCherry
Génotypage PCR2 R	GAAGGCGATTGCGCTGCGAAT	Génotypage de la souris transgénique Flag-3HA-Rfx6-2A-mCherry
hRfx6 PpuMI anti-sens	TCCAGCTTCGGGACCTTGGC	Stratégie de clonage de pCDNA3-3HA-huRfx6
EcoRV 3HA hRfx6 sens	GCGATATCGCAGGAGGATGGCTTACC	Stratégie de clonage de pCDNA3-3HA-huRfx6
screen pcr pCDNA3-3HA-huRfx6 sens	CCTATTGACGTCAATGACGG	Stratégie de clonage de pCDNA3-3HA-huRfx6
screen pcr pCDNA3-3HA-huRfx6 anti-sens	CCGTGGTAAGCAAACCTCTC	Stratégie de clonage de pCDNA3-3HA-huRfx6
pCDNA3 sens	CCACTGCTTACTGGCTTATCG	Production d'un produit de PCR avant TTV
pCDNA3 anti-sens	GCAACTAGAAGGCCACAGTCG	Production d'un produit de PCR avant TTV
R181Q sens kit	CTCCTAACAAACAAGGCAGCTTGGAAACAAGAGGC	Stratégie de clonage des vecteurs pCDNA3-huRfx6 MUTANTS
R181Q asens kit	GCCTCTTGTCCAAAGCTGCCTTGTGTTAGGAG	Stratégie de clonage des vecteurs pCDNA3-huRfx6 MUTANTS
S217P sens kit	GAAAGGGCTTGACAAGGTTTCTGGAAGCAAGCTAAA	Stratégie de clonage des vecteurs pCDNA3-huRfx6 MUTANTS
S217P asens kit	TTTAGCTTGTCCAGGAAACCTTGTCAAGCCCTTTC	Stratégie de clonage des vecteurs pCDNA3-huRfx6 MUTANTS
F294V sens kit	AGATCCAGCATTTTTTATTACACGTTTGGCAAGGAATGCCTGA	Stratégie de clonage des vecteurs pCDNA3-huRfx6 MUTANTS
F294V asens kit	TCAGGCATTCCTTGCCAAACGTTAATAAAAAATGCTGGATCT	Stratégie de clonage des vecteurs pCDNA3-huRfx6 MUTANTS
V506G sens kit	GTGGAGTTTTTTGGTCTCGTGGAAATGCATAATCTCACCTTGA	Stratégie de clonage des vecteurs pCDNA3-huRfx6 MUTANTS
V506G asens kit	TTCAAGGTGAGATATGCATTCACAGCACCACAAAAAATCCAC	Stratégie de clonage des vecteurs pCDNA3-huRfx6 MUTANTS
R578P sens kit	ATGCTTTCTGGCCAAACCTAATAAAGGGAGCATGG	Stratégie de clonage des vecteurs pCDNA3-huRfx6 MUTANTS
R578P asens kit	CCATGCTCCCTTATTAGGTTGGCCAGAAAGCAT	Stratégie de clonage des vecteurs pCDNA3-huRfx6 MUTANTS
Kpn1Not1 huRfx6 sens	TTGGTACCGCGCCGAGGAGGATGGCCAAAGT	Stratégie de clonage de pCDNA3-huRfx6
Kpn1Not1 huRfx6 anti-sens	TTGGTACCGCGCCGCTTAAAGTGCCTCCAGCTG	Stratégie de clonage de pCDNA3-huRfx6
Mlxip1 Xbox1 F	CTCTGGTCTCACAGAGTGTATGC	ChIP qPCR
Mlxip1 Xbox1 R	GGGTGAACCAAGAAACCTCA	ChIP qPCR
Mlxip1 Xbox2 F	TGTCCCCATTACTCTATATGACA	ChIP qPCR
Mlxip1 Xbox2 R	GGGTTGGGGACACTAATCCT	ChIP qPCR
Mlxip1 Xbox3 F	GCCAGCTTCCAGGTACTGAT	ChIP qPCR
Mlxip1 Xbox3 R	CAAGCAATGGCTTCCACTCT	ChIP qPCR
Mlxip1 F	CAGAAGAGATCAGGGCAAGG	qPCR
Mlxip1 R	CGCACACACACATCTACATTCA	qPCR
Rfx6 F	GGCCATGGAGACTCAATTTAAC	qPCR
Rfx6 R	GGCAGCTTACTCGCATCC	qPCR
Rplp0 F	ACTGGTCTAGGACCCGAGAAG	qPCR
Rplp0 R	TCCACCTTGTCTCCAGTCT	qPCR

Tableau 6 : Liste des oligonucléotides utilisés pour les différentes expériences décrites dans ce manuscrit.

Nom	Espèce	Origine	Expérience	Dilution
Anticorps primaires				
HA 12CA5	souris	IGBMC	Western Blot	1/1000
HA 12CA5	souris	IGBMC	Immunofluorescence	1/1000
HA 12CA5	souris	IGBMC	ChIP	1/200
HA 12CA5	souris	IGBMC	Immunoprécipitation	1/25
Rfx6 souris 2767	lapin	IGBMC Equipe Gérard Gradwohl	Western Blot	1/1000
Rfx6 souris 2767	lapin	IGBMC Equipe Gérard Gradwohl	ChIP	1/200
Flag 12F1b11	souris	ATCC	Western Blot	1/1000
Dapi	-	Sigma	Immunofluorescence	1/50000
RFP SAVPA3216	lapin	AbCys	Immunofluorescence	1/100
RFP SAVPA3216	lapin	AbCys	Western Blot	1/500
DS red	souris	IGBMC	Western Blot	1/1000
DS red 632393	souris	Clonetech	Western Blot	1/500
Mlxipl ab92809	lapin	Abcam	Immunofluorescence	1/100
Mlxipl	lapin	Cayman	Immunofluorescence	1/100
Mlxipl NB-400-135	lapin	Novus	Immunofluorescence	1/100
Rfx6 humain SAB1402062	lapin	Sigma	Western Blot	1/500
Actine 1ATD27	souris	IGBMC	Western Blot	1/1000
Anticorps secondaires				
Anticorps anti-souris couplé au DyLight 649	âne	Jackson	Immunofluorescence	1/500
Anticorps anti-lapin couplé au DyLight 649	âne	Jackson	Immunofluorescence	1/500
Anticorps anti-souris couplé à la peroxydase	chèvre	Jackson	Western Blot	1/20000
Anticorps anti-lapin couplé à la peroxydase	chèvre	Jackson	Western Blot	1/20000

Tableau 7 : Liste des anticorps utilisés pour les différentes expériences décrites dans ce manuscrit.

Résultats

Résultats

Le pancréas est une glande mixte puisqu'il assure des fonctions exocrines et endocrines. La fonction principale du pancréas endocrine est la régulation de la glycémie par la sécrétion d'insuline et de glucagon. Ces hormones sont respectivement hypo- et hyperglycémiantes et sécrétées dans la circulation sanguine en réponse à des variations de la glycémie.

La principale pathologie associée au pancréas est le diabète, maladie liée à une élévation anormale du taux de glucose dans le sang, causée par un défaut de production ou d'utilisation de l'insuline. 388 millions de personnes sont diabétiques aujourd'hui et l'Organisation Mondiale de la Santé indique que le diabète pourrait devenir la septième cause de décès dans le monde d'ici 2030.

Pour développer de nouvelles pistes thérapeutiques, de nombreuses équipes de recherche essaient de différencier des cellules souches embryonnaires ou induites en cellules beta en récapitulant les programmes de différenciation embryonnaire (Paragraphe Introduction B.5). A ce titre, l'élucidation des signaux et cascades de régulations transcriptionnelles, qui contrôlent le développement des cellules beta, est essentiel pour la génération de cellules insulino-sécrétrices fonctionnelles *in vitro* et par conséquent pour le développement d'une thérapie cellulaire pour le diabète.

Dans ce contexte, l'équipe de Gérard Gradwohl s'intéresse aux mécanismes moléculaires et cellulaires contrôlant la différenciation des cellules endocrines dans le pancréas. Différents facteurs de transcription et leurs implications dans la régulation transcriptionnelle des étapes précoces de la différenciation du pancréas endocrine sont ainsi déjà connus (Paragraphe Introduction C). Cependant, d'autres recherches doivent être développées dans le but d'élucider notamment les mécanismes de régulation des étapes de différenciation terminale des cellules beta. Dans ce cadre, l'équipe a initialement démontré le rôle central du facteur de transcription pro-endocrine Ngn3 au cours de l'embryogenèse chez la souris dans la spécification de la destinée endocrine des progéniteurs pancréatiques multipotents (Gradwohl et al, 2000). Afin d'identifier de nouveaux gènes contrôlant la différenciation insulaire en aval de Ngn3, le laboratoire a par la suite réalisé des expériences de puce à ADN (Mellitzer et al, 2004). A partir d'un pancréas embryonnaire de souris transgéniques du laboratoire exprimant la *YFP* dans les cellules Ngn3+, qui constituent les précurseurs immatures de l'ensemble des cellules endocrines des îlots de Langerhans, un enrichissement important (65 fois) du facteur de transcription Rfx6 a été révélé (Soyer et al, 2010).

Le facteur de transcription à domaine « winged helix », Rfx6, a plus récemment été identifié comme un nouveau régulateur de la différenciation endocrine pancréatique en aval de Ngn3 chez le poisson zèbre, la souris et l'homme. De plus, diverses mutations de Rfx6 chez l'homme ont été identifiées et reliées au syndrome de Mitchell Riley notamment caractérisé par un diabète néonatal, une atrophie de l'intestin grêle et une malabsorption intestinale (Paragraphe Introduction D.3).

Dans ce contexte, le rôle précis de RFX6 dans la différenciation et la fonction des cellules beta pancréatiques doit être élucidé. L'objectif de mon doctorat consiste ainsi à identifier les gènes ciblés et régulés par RFX6 dans la cellule beta sécrétrice d'insuline chez l'adulte.

A. Recherche des gènes cibles du facteur de transcription RFX6

A.1. Clonage, expression et immunoprécipitation de la protéine de fusion Flag-3HA-RFX6

Pour découvrir le rôle de RFX6 non connu à ce stade dans les cellules beta adultes sécrétrices d'insuline, la détermination de ses gènes cibles a constitué mon projet de thèse. Pour me permettre de développer une stratégie de recherche basée sur des expériences de ChIP Seq qui seront décrites par la suite, j'ai tout d'abord développé des outils préliminaires à cette approche. Parce qu'aucun anticorps dirigé contre RFX6 efficace en ChIP et qu'aucune cible de RFX6 n'étaient connus à ce moment, l'utilisation d'un ChIP anti-HA-RFX6 à l'aide d'un anticorps anti-HA monoclonal, dont l'efficacité en ChIP avait déjà été démontrée (Strub et al, 2011), a été envisagée dans la lignée cellulaire beta murine Min6b1. Cette lignée cellulaire beta murine sécrète de l'insuline en réponse au glucose (Lilla et al, 2003) et exprime Rfx6 de façon endogène constituant ainsi un excellent modèle d'étude. Dans un premier temps, le clonage, l'expression et l'immunoprécipitation d'une protéine de fusion Flag-3HA-RFX6 ont donc été réalisés.

Pour cette stratégie de clonage, le vecteur d'expression eucaryote pCMVTag2A portant l'insert du gène Rfx6 étiqueté par un élément Flag a été ligué avec un insert 3HA (triple épitope HA de l'hémagglutinine du virus de la grippe humaine). La stratégie de clonage de l'étiquette 3HA dans le vecteur pCMVTag2A-Flag-Rfx6 est présentée en **Figure 16**.

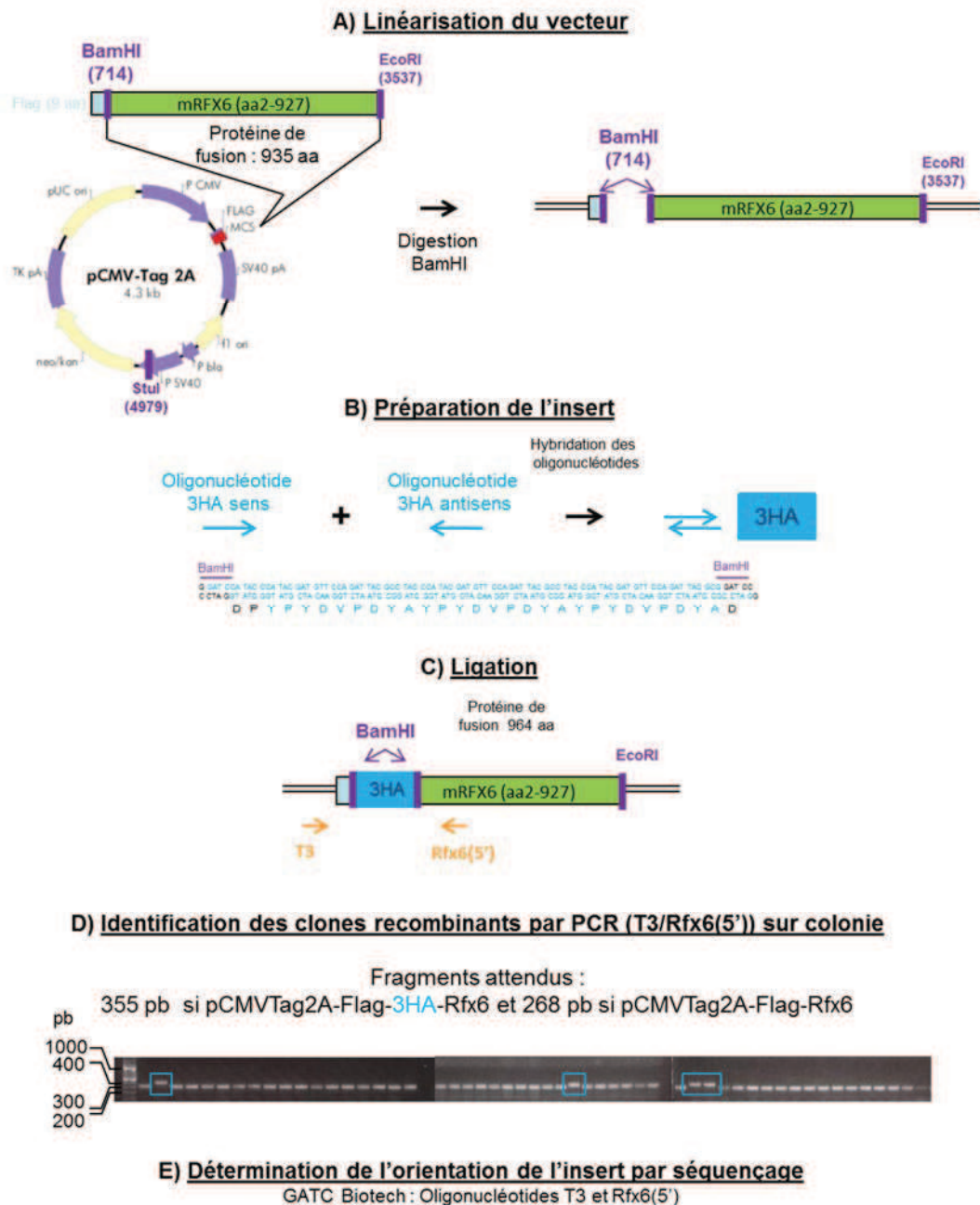


Figure 16 : Stratégie de clonage de l'étiquette 3HA dans le vecteur pCMVTag2A-Flag-Rfx6. Les différentes étapes de cette stratégie sont présentées dans les points **A)** à **E)**. Le vecteur est dans un premier temps linéarisé par digestion de restriction BamHI. Simultanément, l'insert double brin 3HA est créé par hybridation des oligonucléotides sens et antisens. Le vecteur pCMVTag2A-Flag-Rfx6 et l'insert 3HA sont ensuite ligués pour obtenir le vecteur pCMVTag2A-Flag-3HA-Rfx6. Les clones recombinants sont identifiés par PCR à l'aide des oligonucléotides T3 et Rfx6(5') (flèches oranges) et présentent un produit de 355pb. Enfin, l'orientation de l'insert est vérifiée par séquençage de l'échantillon chez GATC Biotech. Les oligonucléotides utilisés dans cette stratégie de clonage sont indiqués dans le **Tableau 6**.

Après obtention d'un clone recombinant possédant l'insert 3HA dans la bonne orientation, l'expression de la protéine de fusion Flag-3HA-RFX6 a été testée dans la lignée cellulaire Min6b1 et est présentée en **Figure 17**. L'expression de la protéine de fusion Flag-3HA-RFX6 (106kDa attendu) dans la lignée cellulaire Min6b1 est ainsi observée (**encadrements bleus, Figure 17B**) pour les trois anticorps testés. Des bandes sont également observées entre 72 et 95kDa dans les pistes 3 à 6, leur présence détectée à l'aide de l'anticorps anti-HA dans les cellules non transfectées suggère qu'il s'agit de bandes aspécifiques. De plus, l'expression endogène de Rfx6 dans les cellules Min6b1 non transfectées est constatée en piste 5 et la transfection des cellules permet d'obtenir une surexpression de Rfx6 observée en piste 6. La transfection des cellules Min6b1 par le vecteur pCMVTag2A-Flag-3HA-Rfx6 permet ainsi l'expression de la protéine de fusion Flag-3HA-RFX6 dans cette lignée cellulaire à des niveaux supérieurs à la protéine endogène.

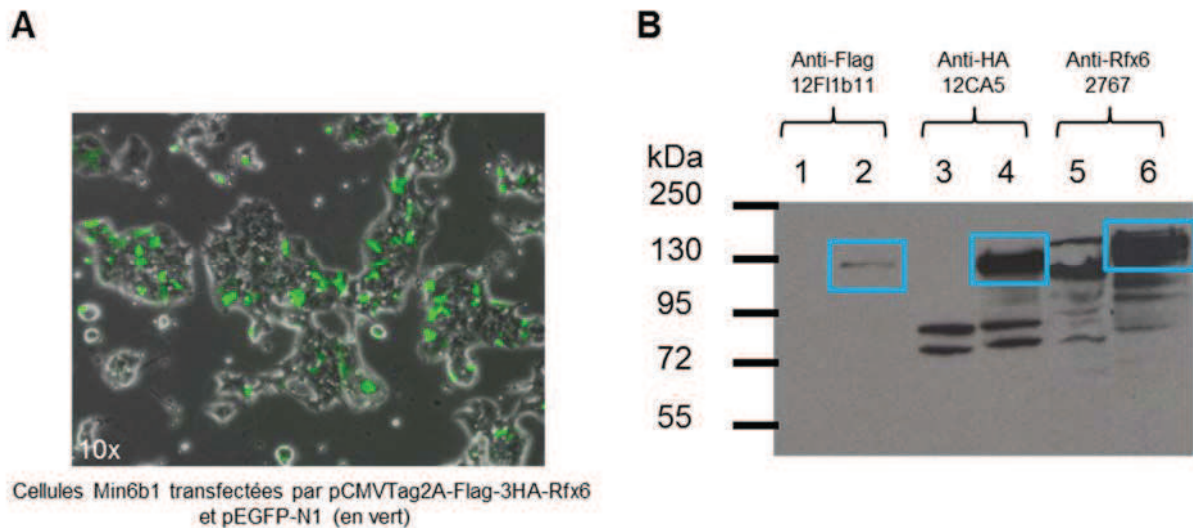


Figure 17 : Expression de la protéine de fusion Flag-3HA-RFX6 dans la lignée cellulaire Min6b1. (A) Photo au microscope à épifluorescence des cellules Min6b1 transfectées par pCMVTag2A-Flag-3HA-Rfx6 et le vecteur témoin de transfection pEGFP-N1 (en vert). (B) Analyse de l'expression de la protéine de fusion Flag-3HA-RFX6 (cadres bleus) par western blot à partir d'extraits protéiques totaux de cellules Min6b1 non transfectées (1-3-5) ou transfectées (2-4-6) par pCMVTag2A-Flag-3HA-Rfx6.

Des expériences d'immunofluorescence (IF) anti-HA, présentées en **Figure 18**, ont également permis de confirmer l'expression de la protéine de fusion Flag-3HA-RFX6 dans les cellules Min6b1 transfectées par le vecteur pCMVTag2A-Flag-3HA-Rfx6 et de révéler sa

localisation nucléaire contrairement à la localisation intracellulaire de la protéine RFX6 endogène dans les Min6b1 qui ne semble pas uniquement nucléaire (voir **Figure 66**) lorsqu'elle est révélée par IF anti-RFX6.

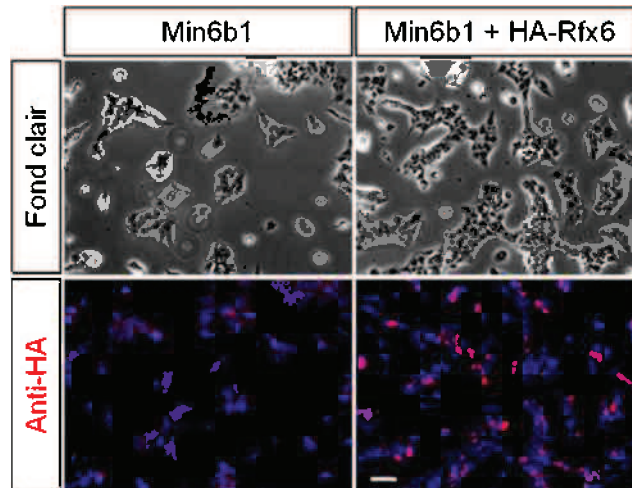


Figure 18 : Localisation nucléaire de la protéine 3HA-Rfx6 dans les cellules Min6b1. Immunofluorescence anti-HA (en rouge) sur des cellules Min6b1 non transfectées (à gauche) ou transfectées par le vecteur pCMVTag2A-Flag-3HA-Rfx6 (à droite). Barre d'échelle = 50µm.

Dans le but de poursuivre la mise au point des conditions expérimentales pour une analyse en ChIP Seq des gènes cibles de RFX6 dans les cellules Min6b1, l'immunoprécipitation de la protéine de fusion Flag-3HA-RFX6 a été testée à l'aide de l'anticorps anti-HA.

Comme attendu, les résultats montrent, que pour les cellules non transfectées (**Figure 19 pistes 1 à 3**), aucune protéine n'est immunoprécipitée (**pistes 3**). Lorsque les cellules sont transfectées par le vecteur pCMVTag2A-Flag-3HA-Rfx6 (**Figure 19, pistes 4 à 6**) la protéine de fusion Flag-3HA-RFX6 (**encadrements bleu, pistes 6**) est par contre immunoprécipitée, et aucune trace ne reste dans le surnageant après immunoprécipitation (**pistes 5**). Ces résultats sont obtenus à la fois avec les western blots anti-RFX6 et anti-Flag.

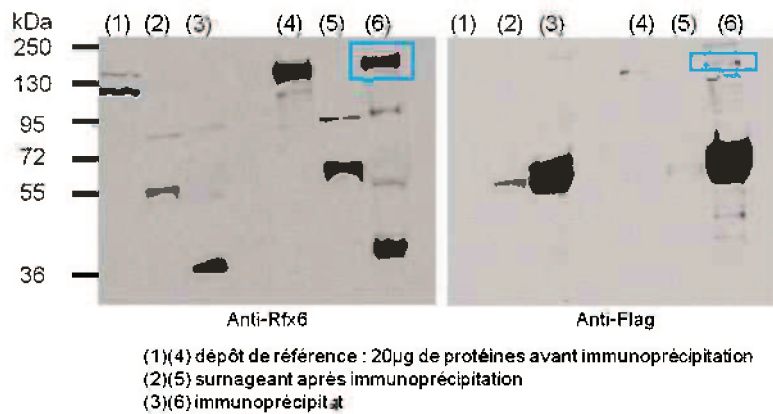


Figure 19 : Immunoprécipitation de la protéine Flag-3HA-Rfx6. Analyse par western blot de l'immunoprécipitation de la protéine de fusion Flag-3HA-RFX6 (cadres bleus) d'extraits protéiques totaux de cellules Min6b1 non transfectées ((1)(2)(3)) ou transfectées par pCMVTag2A-Flag-3HA-Rfx6 ((4)(5)(6)). L'immunoprécipitation est réalisée par l'anticorps anti-HA et les révélations sont réalisées avec les anticorps anti-RFX6 ou anti-Flag. Les bandes à 55 et 36kDa correspondent aux billes de G-Sépharose utilisées pour l'immunoprécipitation.

En résumé, après production d'un vecteur pCMVTag2A-Flag-3HA-Rfx6, l'expression de la protéine de fusion Flag-3HA-RFX6 dans la lignée cellulaire Min6b1 a été réalisée. La capacité de l'anticorps anti-HA à immunoprécipiter cette protéine a également été démontrée. Cette immunoprécipitation validée, la mise au point des expériences de CHIP a pu être mise en place dans le but de déterminer les gènes cibles de RFX6 dans la lignée cellulaire Min6b1.

A.2. Développement d'une souris transgénique 3HA-RFX6-mCherry

A.2.a. Stratégie de clonage

La détermination des cibles de RFX6 dans les cellules Min6b1 nous permettra d'avoir une première idée de son rôle dans les cellules beta sécrétrices d'insuline, cependant les cibles de RFX6 dans ce modèle cellulaire ne reflètent pas l'ensemble des cibles de ce facteur de transcription *in vivo*. Ainsi, dans le but d'obtenir un modèle *in vivo* que l'on pourra utiliser en CHIP Seq à l'aide de l'anticorps anti-HA pour déterminer les cibles de RFX6, une stratégie de production d'une souris transgénique exprimant la protéine de fusion Flag-3HA-RFX6-2A-mCherry a été développée. L'expression de la protéine fluorescente mCherry permettra de vérifier que le transgène est bien exprimé dans les cellules exprimant Rfx6 de façon endogène, mais également de purifier ces cellules par tri cellulaire. Pour obtenir cette

souris, un BAC (Bacterial Artificial Chromosome) contenant les hypothétiques éléments régulateurs de Rfx6 et codant pour la protéine de fusion Flag-3HA-RFX6-2A-mCherry a été produit. Après ajout de la séquence codante de la mCherry en 3' de la séquence codante de Rfx6, puis d'une cassette de sélection à la néomycine permettant par la suite la sélection des BAC recombinants, ce sont des bras homologues au BAC Rfx6 RP23-415D6 (BAC PAC Resources Center, Children's Hospital Oakland Research Institute) contenant les éléments régulateurs de Rfx6 ainsi que sa séquence codante qui sont ajoutés avant recombinaison pour obtenir le BAC Rfx6 RP23-415 D6-Flag-3HA-Rfx6-2A-mCherry-pA-Néo^R d'intérêt. Ce dernier est enfin utilisé pour l'obtention de souris transgéniques exprimant la protéine de fusion Flag-3HA-RFX6 ainsi que la mCherry sous le contrôle des éléments régulateurs de Rfx6. La stratégie de clonage est présentée ci-dessous.

Ajout de la séquence codante de la mCherry en 3' de Rfx6

Après obtention du plasmide pCMVTag2A-Flag-3HA-Rfx6 (voir paragraphe Résultats A.1) la séquence codante pour la mCherry (Shaner et al, 2004) ainsi que le peptide auto-clivable 2A (Kim et al, 2011) sont ajoutés en 3' de Rfx6. Pour cela, une stratégie de PCR en 3 temps, présentée dans la **Figure 20**, est mise en place et le plasmide pCMVTag2A-Flag-3HA-Rfx6-2A-mCherry est produit.

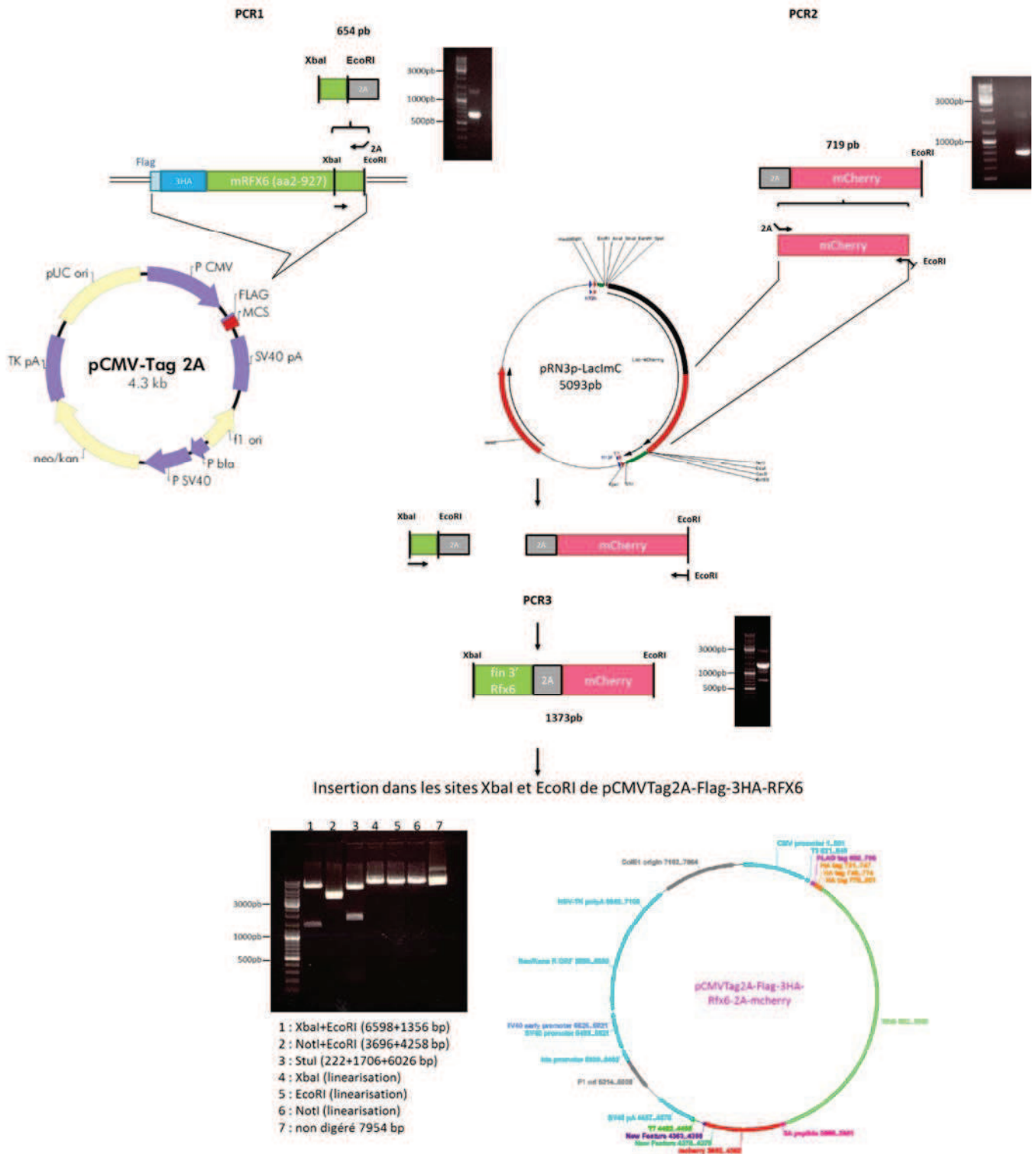


Figure 20 : Production du plasmide pCMVTag2A-Flag-3HA-Rfx6-2A-mCherry. Les trois PCR présentées sont réalisées à l'aide des amorces indiquées dans le **Tableau 6**. Dans un premier temps, le peptide 2A est ajouté par PCR (PCR1) en 3' de Rfx6 puis il est également ajouté en 5', de manière conjointe avec un site EcoRI, de la séquence codante pour la mCherry par PCR (PCR2) sur le plasmide pRN3P-LacI-mCherry (don de Yusuke Miyanari, équipe Maria Elena Torres Padilla). Enfin une troisième PCR (PCR3) permettant la fusion des deux produits de PCR précédents est réalisée.

Après digestion de restriction du produit de PCR final, l'insertion de ce dernier est réalisée dans les sites XbaI et EcoRI du plasmide pCMVTag2A-Flag-3HA-Rfx6. Le produit final de cette première étape est vérifié par restriction et séquençage.

Expression de la protéine de fusion Flag-3HA-RFX6-2A-mCherry dans les cellules Min6b1

Pour s'assurer de la production de la protéine de fusion clonée sous la dépendance du promoteur ubiquitaire CMV et de la possibilité de son utilisation en ChIP Seq HA-RFX6 sur cellules transfectées, son expression a par la suite été testée par WB et IF dans les cellules Min6b1. La protéine endogène RFX6 est observée par WB anti-Rfx6 dans le puits 1 de la **Figure 21**. La protéine de fusion 3HA-RFX6 est détectée dans les Min6b1 transfectées. Des bandes aspécifiques sont présentes pour le WB anti-HA et des protéines dégradées sont également détectées par WB anti-HA ou anti-Rfx6. Par contre, aucune protéine mCherry n'a pu être détectée par WB dans les Min6 transfectées quel que soit l'anticorps testé (Anti-DSred Clonotech, Anti-DSred IGBMC et Anti-RFP AbCys) (**Figure 21**).

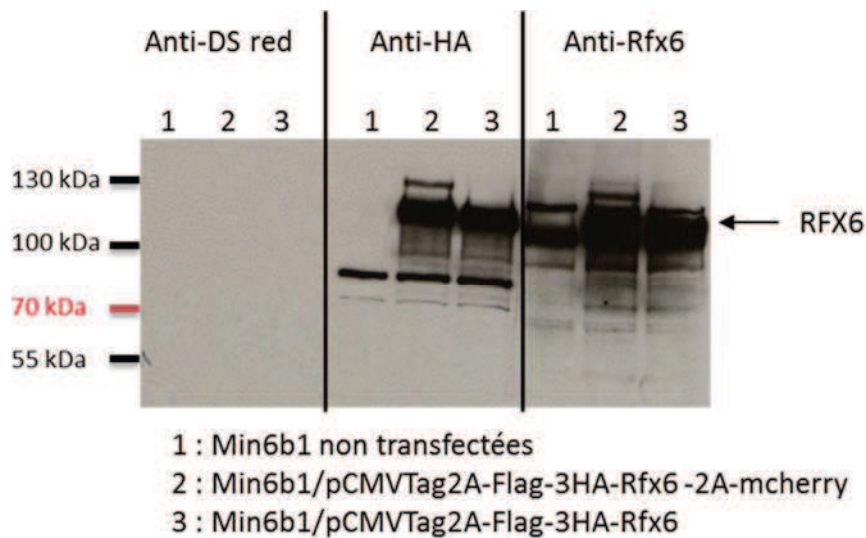


Figure 21 : Expression de la protéine de fusion Flag-3HA-RFX6-2A-mCherry dans les cellules Min6b1. Western blot sur extraits protéiques totaux de cellules Min6b1 transfectées ou non. La protéine de fusion Flag-3HA-RFX6 est produite dans les cellules Min6b1 (105kDa).

Par IF ou fluorescence directe de la mCherry (**Figure 22**), l'expression de la mCherry mais également de la protéine de fusion 3HA-RFX6 est détectée dans les cellules Min6b1 transfectées (flèches blanches).

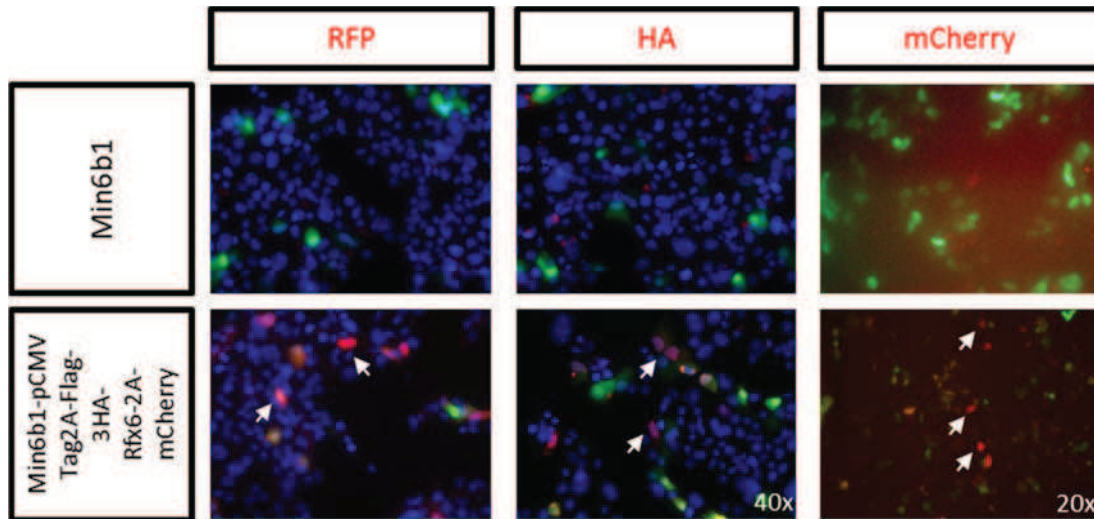


Figure 22 : Expression des protéines 3HA-RFX6 et mCherry dans les cellules Min6b1. Immunofluorescence contre la mCherry (RFP) et la protéine de fusion HA-RFX6 (HA) et auto-fluorescence de la mCherry (mCherry) (en rouge). Les noyaux cellulaires sont marqués au DAPI en bleu. Les cellules sont co-transfectées avec le plasmide pEGFP-N1 qui exprime la GFP (vert) pour estimer l'efficacité de transfection. Les protéines 3HA-RFX6 et mCherry sont exprimées dans les cellules Min6b1 (flèches blanches).

La production de protéine mCherry n'a pas pu être observée par WB à l'aide des anticorps testés mais pourra probablement être révélée par d'autres mises au point cependant, le résultat observé en IF indique que la protéine de fusion 3HA-RFX6 ainsi que la mCherry sont exprimées dans les cellules Min6b1 transfectées par le plasmide pCMVTag2A-Flag-3HA-Rfx6-2A-mCherry. L'expression du transgène 3HA-Rfx6-2A-mCherry a ainsi été validée et la stratégie de clonage a pu être poursuivie.

Ajout de la cassette de sélection pour la Néomycine en 3' de la construction Flag-3HA-RFX6-2A-mCherry

Une cassette de sélection à la Néomycine étant nécessaire pour la suite de ce projet, l'ajout de la séquence pA-FRT-Néo est réalisée. Cette séquence est constituée d'une queue

poly-A permettant de stabiliser l'ARN qui sera produit, et de la séquence de la néomycine entourée de sites FRT (flippase recognition target) qui pourront être utilisés pour éliminer la séquence de sélection à la Néomycine par la recombinase Flp si besoin, dans la souris produite (Zhu et al, 1995). Pour cela, la stratégie d'obtention du vecteur pFlag-3HA-Rfx6-2A-mCherry-Néo à partir de la construction obtenue précédemment, et du vecteur MyogeninVenusA-FRTNéo (don de Jean Marie Garnier, équipe Olivier Pourquié) est présentée en **Figure 23**. Le vecteur d'intérêt pFlag-3HA-Rfx6-2A-mCherry-Néo est ensuite vérifié par restriction et séquençage.

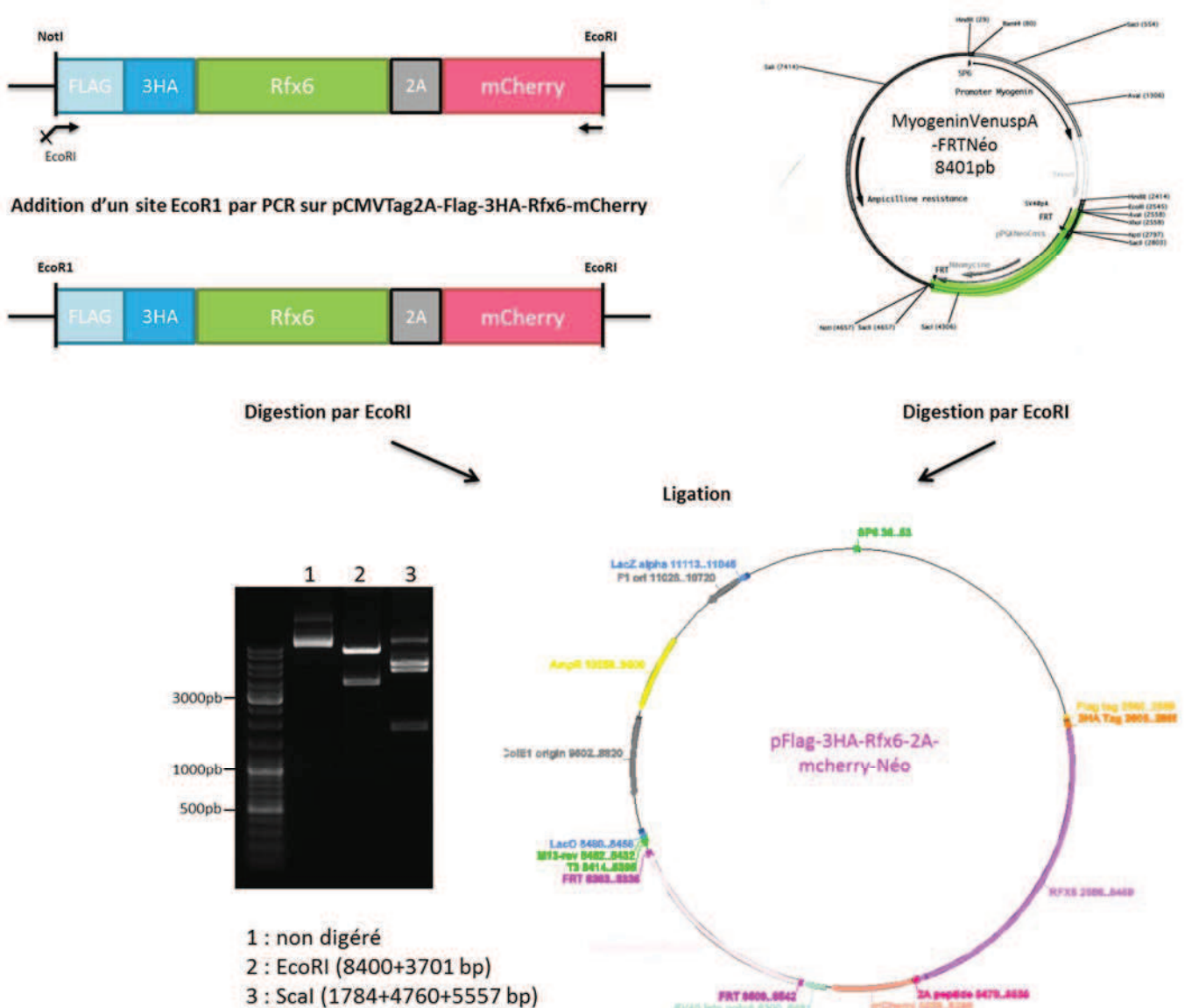


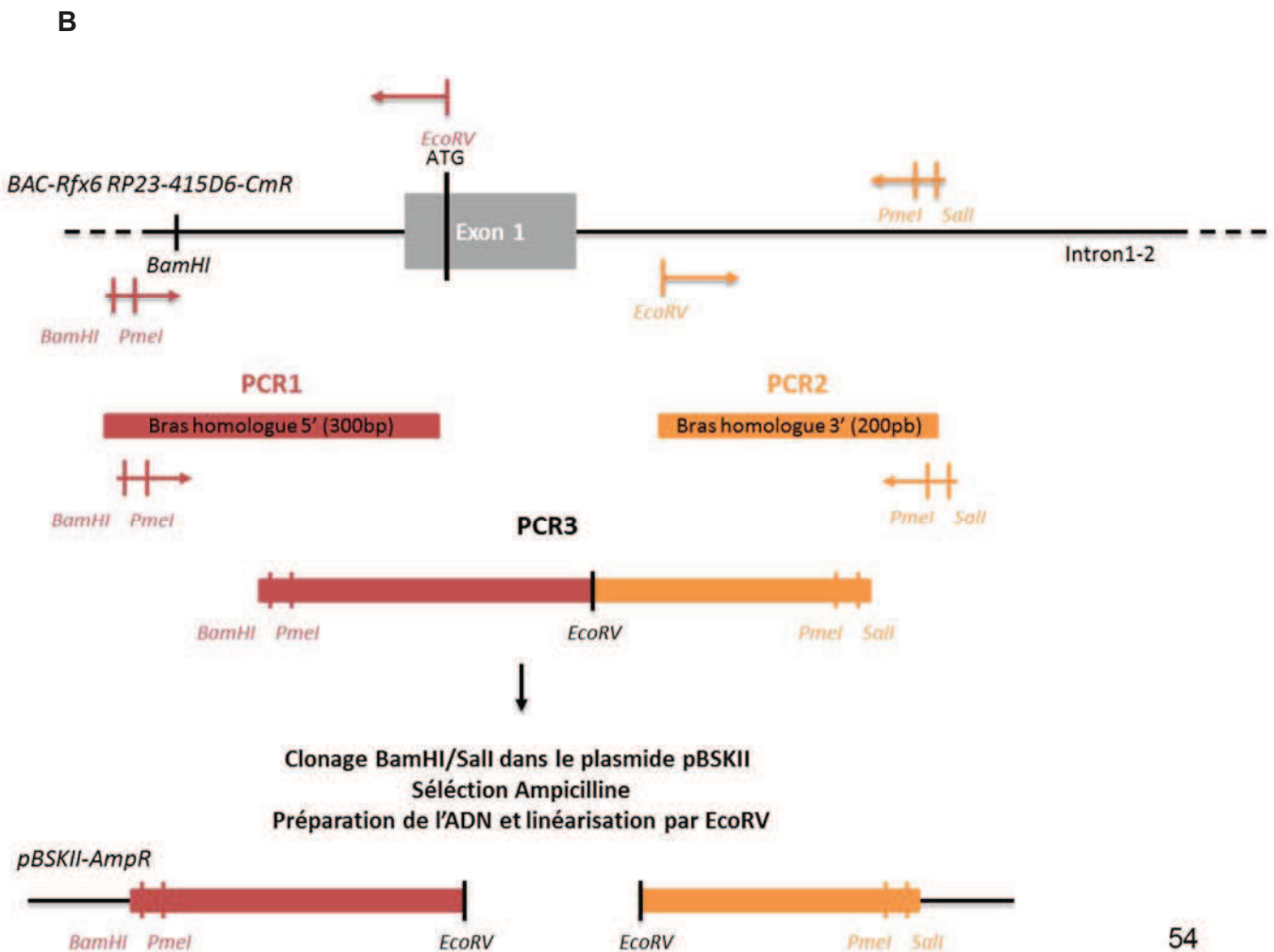
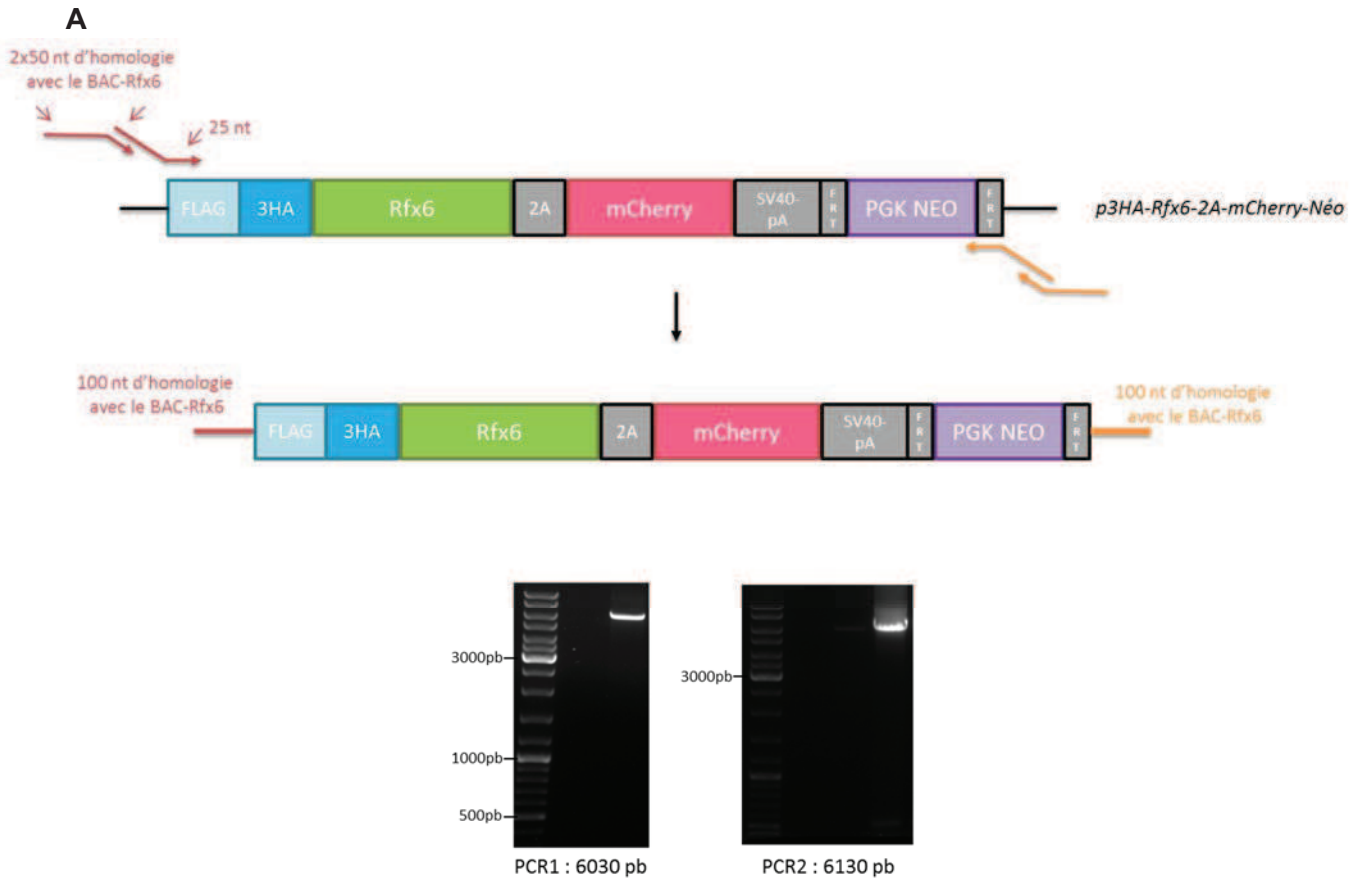
Figure 23 : Production du plasmide pFlag-3HA-Rfx6-2A-mCherry-Néo. Un site EcoRI est ajouté par PCR sur le vecteur pCMVtag2A-Flag-3HA-Rfx6-2A-mCherry en 3' de la construction Flag-3HA-

Rfx6-2A-mCherry avant restriction par EcoRI. Le vecteur MyogeninVenusPA-FRTNéo est digéré par EcoRI. Une ligation entre ces deux produits permet d'obtenir le plasmide d'intérêt pFlag-3HA-Rfx6-2A-mCherry-Néo.

Ajout de bras homologues au gène Rfx6 en 5' et 3' de la construction Flag-3HA-Rfx6-2A-mCherry-Néo

L'étape suivante consiste en l'ajout de bras homologues au gène Rfx6 permettant l'intégration de la construction Flag-3HA-Rfx6-2A-mcherry-Néo par recombinaison homologue dans le BAC Rfx6 RP23-415D6 (BAC PAC Resources Center, Children's Hospital Oakland Research Institute). Celui-ci a été choisi car il contient la séquence codante ainsi que d'hypothétiques éléments régulateurs du gène Rfx6. Ces bras homologues permettront l'intégration du transgène sous le contrôle de ces éléments régulateurs de Rfx6 dans le génome de la souris. Pour cela, deux étapes clés sont accomplies.

Dans un premier temps, 100 nucléotides d'homologie sont ajoutés en 5' et 3' de la construction produite à l'étape précédente par deux PCR successives (**Figure 24A**). En parallèle, par PCR en trois temps, des bras homologues de respectivement 300nt en 5' et 200nt en 3' de la région d'intérêt présente dans le BAC Rfx6 RP23-415D6 sont clonés dans le plasmide pBSKII. (**Figure 24B**). Puis, une recombinaison homologue entre la construction comportant les 100nt d'homologie et le plasmide linéarisé permet d'obtenir l'ADN d'intérêt (**Figure 24C**).



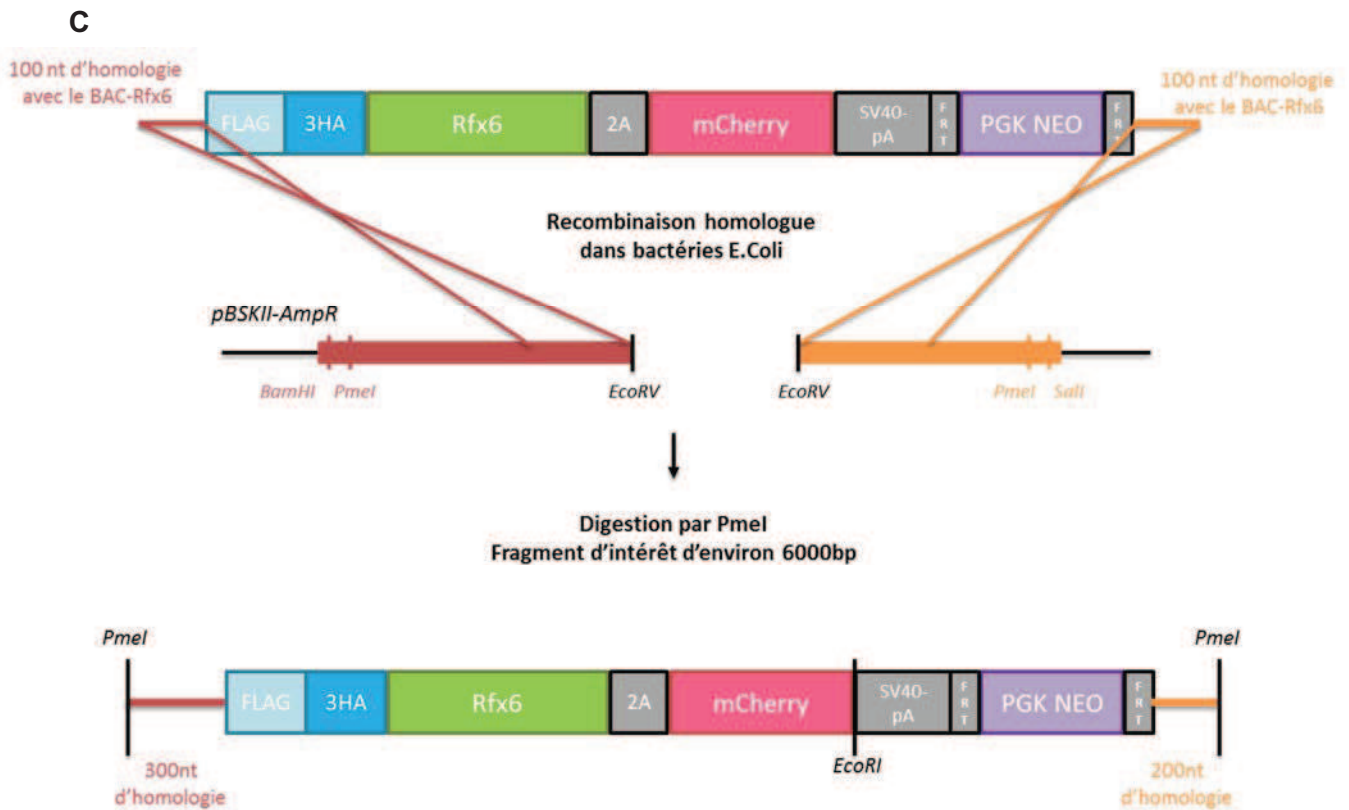


Figure 24 : Recombinaison homologue entre le BAC Rfx6 RP23-415D6 et la construction Flag-3HA-Rfx6-2A-mCherry-Néo. (A) Ajout de 100 nucléotides d'homologie au gène Rfx6 de part et d'autre de la construction par PCR en deux temps à l'aide des amorces indiquées dans le **Tableau 6**. A chaque étape, les produits de PCR obtenus de 6030 puis 6130 pb sont vérifiés sur gel d'agarose. (B) Clonage des bras homologues de respectivement 300nt en 5' et 200nt en 3' de la région d'intérêt du BAC Rfx6 RP23-415D6 dans les sites BamHI et Sall du plasmide pBSKII qui sera par la suite linéarisé par restriction au niveau du site EcoRV (les sites de restriction d'intérêt et nécessaires à cette stratégie sont à chaque fois ajoutés par les oligonucléotides utilisés pour les différentes PCR, de même pour le site PmeI, qui sera utilisé en (C)). (C) Recombinaison homologue entre les produits obtenus en (A) et (B) et obtention du fragment d'intérêt après restriction par PmeI.

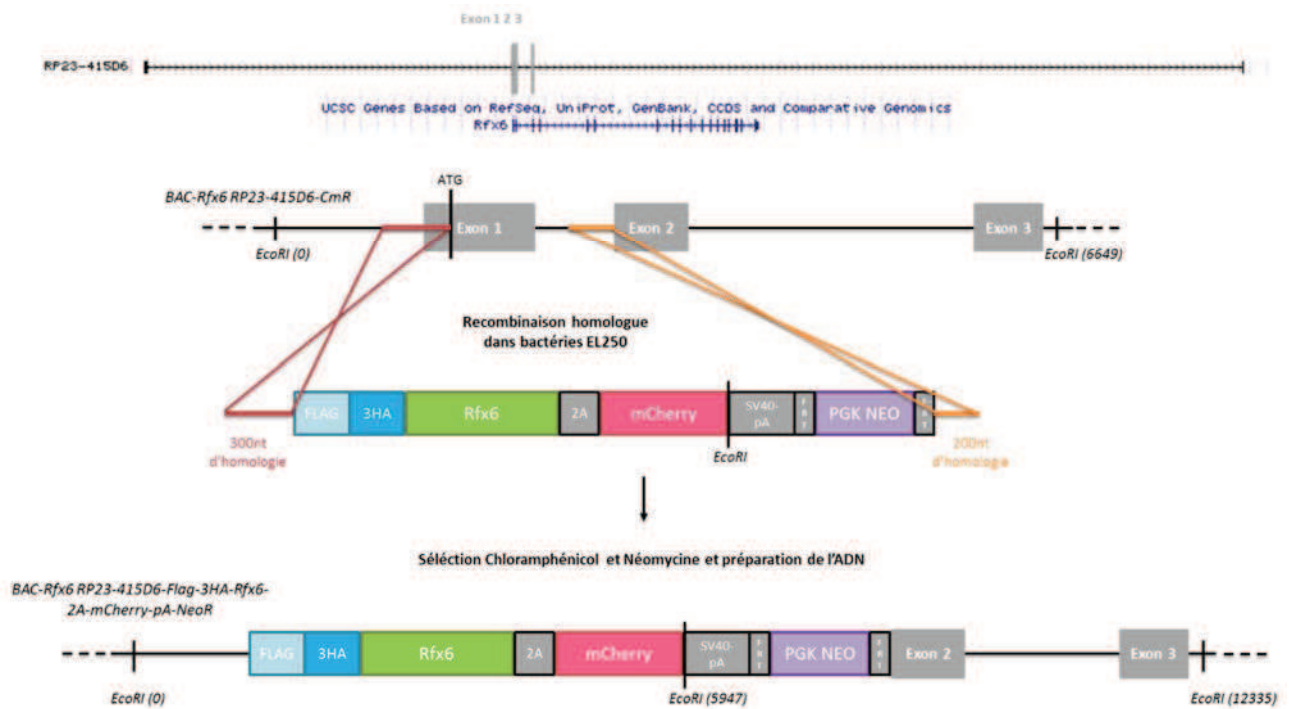
Obtention de l'ADN d'intérêt BAC Rfx6 RP23-415D6-Flag-3HA-Rfx6-2A-mCherry-pA-Neo^R par recombinaison homologue

La dernière étape de cette stratégie de clonage consiste en une recombinaison homologue entre l'ADN Flag-3HA-Rfx6-2A-mCherry-pA-Néo possédant les 300 et 200nt homologues au BAC Rfx6 RP23-415D6 de part et d'autre et le BAC Rfx6 RP23-415D6 lui-même (**Figure 24A**). Ceci est réalisé dans les bactéries génétiquement modifiées EL250 (Lee et al, 2001) possédant une efficace fonction de recombinaison homologue inducible par la température (42°C). Les clones recombinants sont sélectionnés par le chloramphénicol pour

l'insertion du BAC Rfx6 RP23-415D6 et par la Néomycine pour l'insertion du BAC recombinant BAC Rfx6 RP23-415D6-Flag-3HA-Rfx6-2A-mCherry-pA-Néo^R. L'ADN obtenu pour les clones résistants à ces deux sélections est vérifié par séquençage et restriction EcoRI où une bande supplémentaire à 6388pb est observée (**Figure 24B**).

Ces deux dernières étapes essentielles (ajout des bras homologues et recombinaison) ont été réalisées avec l'aide de Jean-Marie Garnier (Equipe Olivier Pourquoi).

A



B

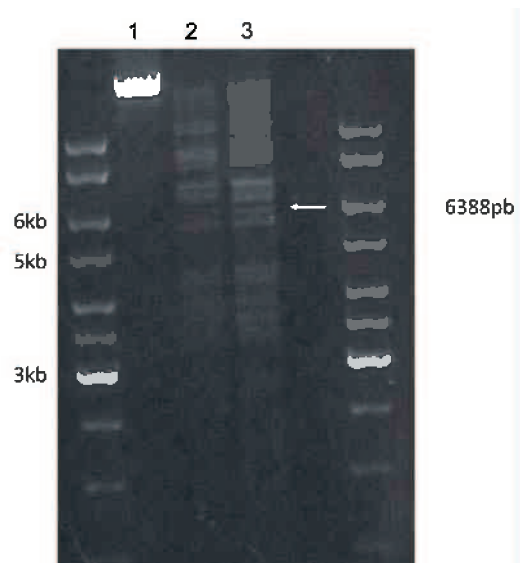


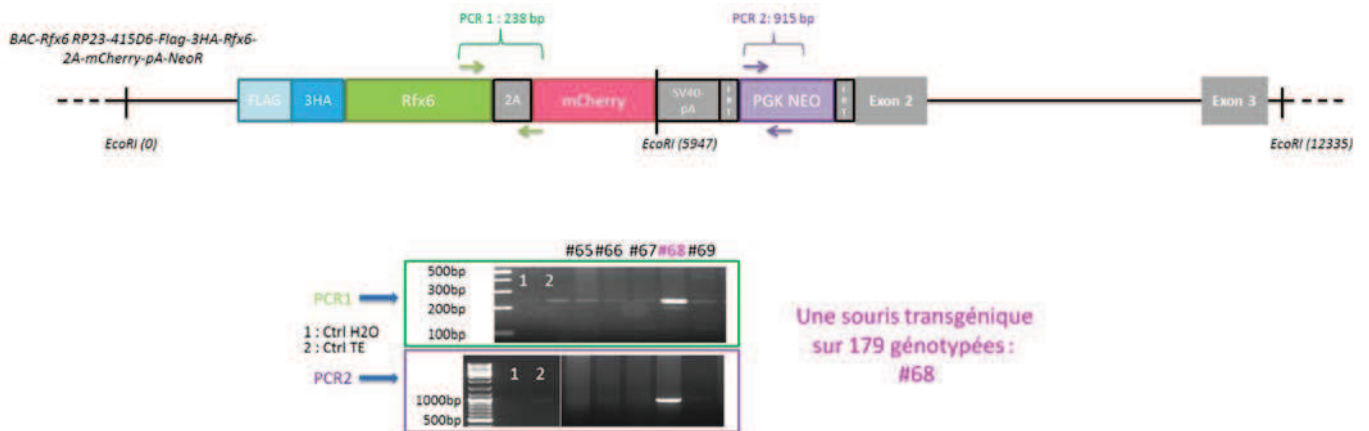
Figure 25 : Obtention de l'ADN d'intérêt BAC Rfx6 RP23-415D6-Flag-3HA-Rfx6-2A-mCherry-pA-Neo^R par recombinaison homologue. (A) Stratégie de recombinaison homologue. (B) Digestion de restriction EcoRI du BAC recombinant BAC Rfx6 RP23-415D6-Flag-3HA-Rfx6-2A-mCherry-pA-FRT-Neo^R (1 : BAC Rfx6 RP23-415D6 non digéré ; 2 : BAC Rfx6 RP23-415D6 digéré par EcoRI ; BAC Rfx6 RP23-415D6-Flag-3HA-Rfx6-2A-mCherry-pA-Neo^R digéré par EcoRI). Le BAC recombinant présente une bande supplémentaire à 6388 pb indiquée par la flèche blanche.

A.2.b. Obtention d'une souris transgénique et expression du transgène 3HA-Rfx6-mCherry

Après préparation et purification de l'ADN du BAC d'intérêt, le projet a été confié à l'Institut Clinique de la Souris, où une injection de cet ADN dans des pro noyaux d'œufs de souris fertilisées puis implantation de ceux-ci chez une souris pseudo-gestante ont été réalisées. L'ADN a été injecté à 2ng/μL et l'implantation de 49 pro noyaux a été réalisée. Parmi les 179 souris obtenues, dont le génotype a été déterminé par PCR (**Figure 26A**), une unique souris transgénique a été identifiée.

Cette souris femelle a ensuite été croisée avec des mâles et la caractérisation de la lignée a été réalisée (**Figure 26B**). Quel que soit le stade testé (stade embryonnaire E15,5 et Adulte 3 semaines ou 11mois (souris fondatrice), aucune expression de la mCherry n'a pu être mise en évidence ni par fluorescence directe ni par IF. Cela peut être dû à un problème au niveau du site d'intégration du transgène dans le génome de la souris où un effet de position ne permet peut-être pas son expression, ou encore à une intégration partielle du transgène dans le génome. Une autre raison pourrait être que le BAC choisi ne contient pas les éléments régulateurs de l'expression spatio-temporelle de Rfx6. Cette lignée a donc été sacrifiée début juin 2014.

A



B

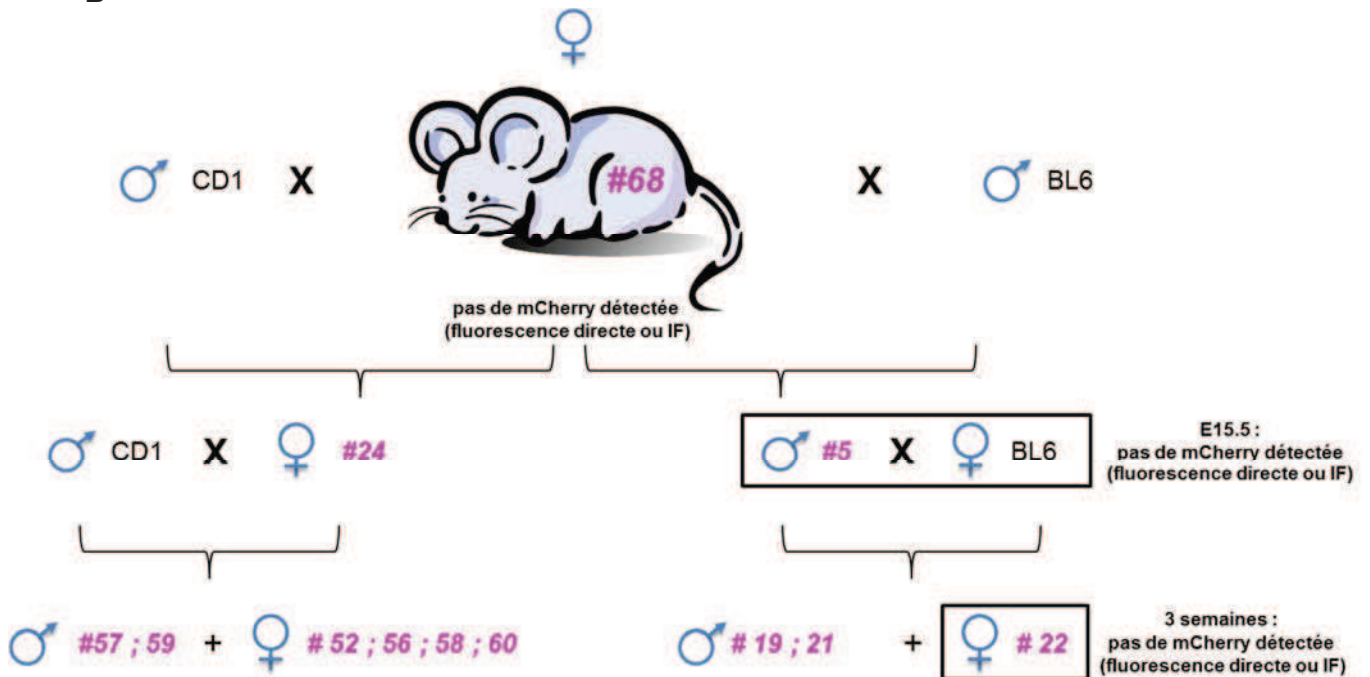


Figure 26 : Obtention d'une souris transgénique et expression du transgène Flag-3HA-Rfx6-2A-mCherry. (A) Stratégie de génotypage de la lignée de souris transgénique. Les amorces utilisées pour les PCR 1 et 2 sont indiquées dans le **Tableau 6**. (Contrôles négatifs : TE = Tris 10mM EDTA 1mM et eau) (B) Caractérisation de la lignée transgénique. Les souris numérotées en rose sont positives pour les deux PCR présentées en (A). Cinq souris transgéniques E15,5, une adulte de 3 semaines et la souris fondatrice de 11 mois ont été testées pour l'expression de la mCherry en fluorescence directe ou par IF.

Dans cette partie, le but était d'obtenir un modèle de souris transgénique pouvant être utilisé en ChIP Seq à l'aide de l'anticorps anti-HA pour déterminer les cibles de RFX6 *in vivo*. Parallèlement, la mise au point d'expériences de ChIP Seq testée dans les cellules Min6b1, qui sont décrites dans le paragraphe Résultats A.3, a permis de valider l'anticorps anti-Rfx6 produit dans l'équipe pour ces expériences. Le développement de cette souris transgénique n'a donc pas été poursuivi.

A.3. Recherche des gènes cibles de RFX6 dans la lignée cellulaire beta Min6b1

A.3.a. Stratégie expérimentale

Le projet de ce travail est de déterminer les gènes cibles de RFX6 afin de définir le rôle de ce facteur de transcription. L'identification des gènes cibles de RFX6 a été tout d'abord recherchée dans un modèle cellulaire beta, les cellules Min6b1, puis dans les îlots pancréatiques de souris purifiés.

Pour cela, j'ai tout d'abord mis au point dans le laboratoire la méthode de ChIP Seq dans la lignée cellulaire beta Min6b1 qui sécrète de l'insuline en réponse au glucose et exprime RFX6 de manière endogène. D'autre part, les cibles d'un membre de la même famille de facteur de transcription, RFX3, avaient déjà été déterminées dans ce modèle en 2010 et sa fonction de régulateur de la différenciation et de la fonction de la cellule beta a pu être démontrée (Ait-Lounis et al, 2010), suggérant que les cellules Min6b1 constituent un modèle adéquat pour identifier les cibles de RFX6. Deux obstacles se sont alors présentés : 1) aucun anticorps dirigé contre RFX6 efficace en ChIP n'était connu et 2) à ce stade, aucune information concernant les cibles de ce facteur de transcription n'étant publiée, nous n'avions aucun contrôle positif nous permettant de valider nos résultats.

Pour surmonter ce problème, l'identification des sites de liaisons de Rfx6 communs entre deux expériences de ChIP réalisées en parallèle, qui sont décrites en **Figure 27**, a été accomplie. D'une part, un ChIP anti-RFX6 à l'aide d'un anticorps de lapin produit précédemment dans l'équipe (Rfx6 2767) sur de la chromatine de cellules Min6b1. D'autre part, à l'aide d'un anticorps anti-HA monoclonal (HA 12CA5) dont l'efficacité en ChIP avait déjà été démontrée (Strub et al, 2011), j'ai réalisé un ChIP anti-HA. Celui-ci a été effectué sur de la chromatine de cellules Min6b1 transfectées avec le plasmide pCMVTag2A-Flag-3HA-Rfx6-2AmCherry et exprimant la protéine de fusion Flag-3HA-RFX6 (**voir paragraphe Résultats A.1**).

Après mise au point du temps de sonication à 4 min pour obtenir des fragments de chromatine entre 100 et 800 pb, les bibliothèques d'ADN ont été préparées (**comme indiqué**

dans le paragraphe Matériels et méthodes C.1) puis séquencées en collaboration avec la plateforme de Biopuces et Séquençage de l'IGBMC. Pour les deux expériences, des bibliothèques d'ADN de bonne qualité (taille exigée entre 100 et 800pb) et en quantité suffisante (10ng d'ADN nécessaire au minimum par expérience pour le séquençage) ont été obtenues et le séquençage a pu être réalisé (Figure 27).

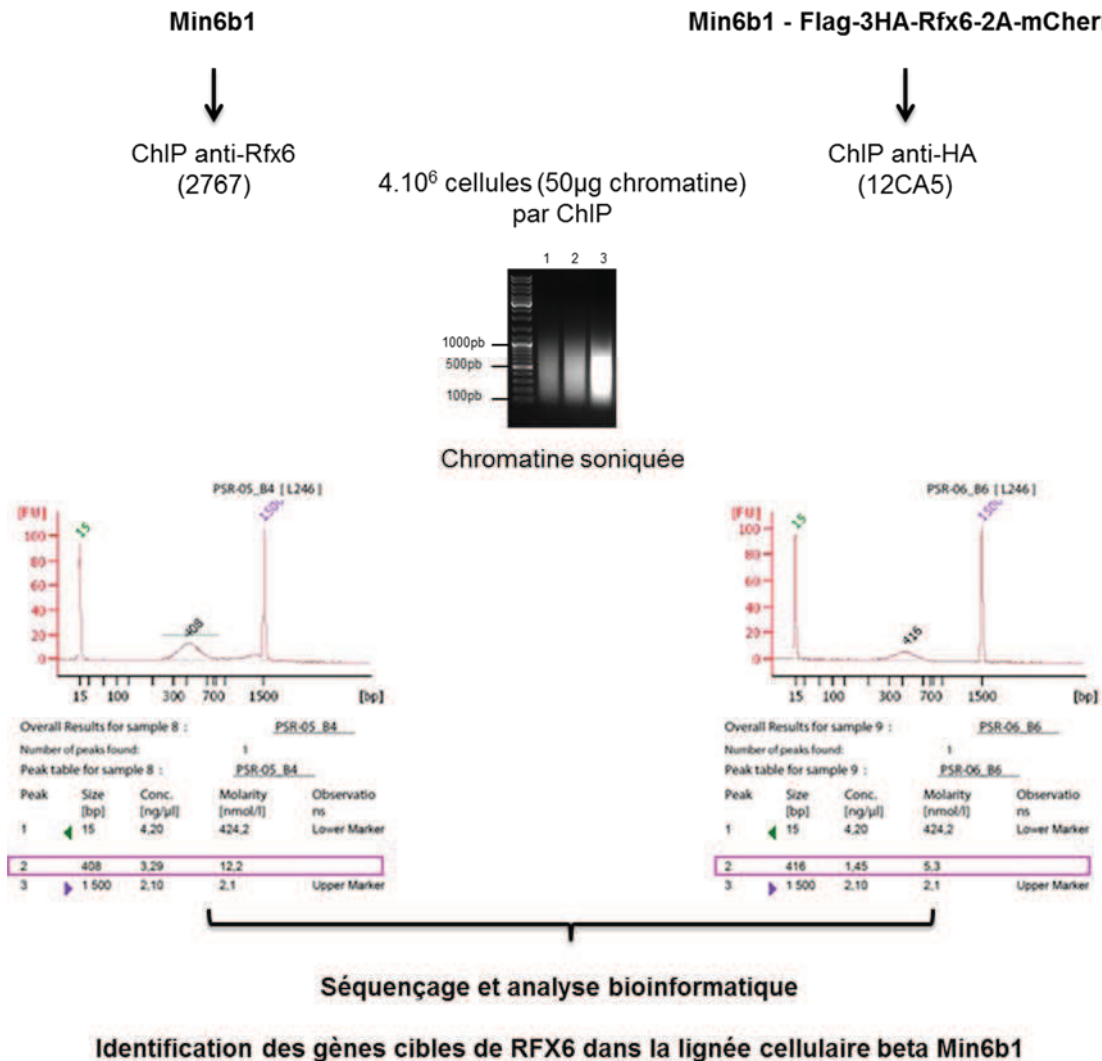


Figure 27 : Stratégie d'identification des gènes cibles de RFX6 dans les cellules Min6b1. Un premier ChIP anti-RFX6 sur de la chromatine de cellules Min6b1 est réalisé (à gauche). Une seconde expérience de ChIP anti-HA-RFX6 est réalisé sur de la chromatine de cellules Min6b1 transfectées (à droite) avec le plasmide pCMVtag2A-Flag-3HA-Rfx6-2A-mCherry. Dans les deux cas, la chromatine est soniquée pour obtenir des fragments d'ADN de 100 à 800pb (au centre) dont la taille est évaluée par dépôt sur gel d'agarose de 500ng (puits 1), 1µg (puits 2), ou 2µg (puits 3) d'échantillon. Enfin, après préparation des bibliothèques d'ADN (encadrements roses), le

séquençage et l'analyse bioinformatique permettront l'identification des cibles de RFX6 dans les cellules Min6b1.

A.3.b. Identification et analyse des sites de liaison de RFX6 dans les cellules Min6b1

Pour chaque jeu de données, différentes analyses bio-informatiques ont été mises en places avec l'aide de Tao YE de la plateforme de Biopuces et Séquençage de l'IGBMC et sont décrites dans le paragraphe Matériels et méthodes C. Pour permettre la comparaison des différentes expériences dans le navigateur génomique UCSC (<https://genome.ucsc.edu/>) les données obtenues ont été normalisées à dix millions de séquences. Le résumé des expériences de ChIP Seq menées dans la lignée Min6b1 à l'aide des anticorps anti-HA et anti-RFX6 est présenté en **Tableau 8**.

Jeu de données	ChIP RFX6	ChIP HA-RFX6	Pics communs ChIP RFX6 et HA-RFX6	Contrôle ChIP HA		ChIP RFX6	ChIP HA-RFX6	Pics communs ChIP RFX6 et HA-RFX6
Matériel biologique	Min6b1	Min6b1-pCMVTag2A-Flag-3HA-Rfx6-2A-mcherry		Min6b1		Min6b1	Min6b1-pCMVTag2A-Flag-3HA-Rfx6-2A-mcherry	
Contrôle	aucun	aucun		/		input	Contrôle ChIP HA - Min6b1 non transfectées	
Taille moyenne de l'ADN (pb)	408	416		655		408	416	
Quantité totale d'ADN après préparation de la librairie (en ng)	164,5	72,5		26		164,5	72,5	
Nombre de séquences	72621350	92195546		25072430		72621350	92195546	
Nombre de séquences alignées	49443919	61039380		16911298		49443919	61039380	
Séquences uniques	33949758	48773606		3567375		33949758	48773606	
Nombre de pics identifiés	3303	22365	1671	/		1552	23308 / 5104	863

Tableau 8 : Résumé des différentes expériences de ChIP Seq réalisées dans les cellules Min6b1. Les séquences totales sont alignées sur le génome de souris mm9/Ensembl et le nombre de séquences alignées est indiqué pour chaque expérience. Les séquences uniques correspondent au nombre de positions différentes dans le génome sur lesquelles des séquences sont alignées. Le nombre de sites de fixation identifiés est indiqué dans la dernière ligne pour chaque expérience. Le ChIP RFX6 est normalisé à l'aide de son input (chromatine avant immunoprécipitation) et permet l'identification de 1552 pics. Les données du ChIP HA-RFX6 sont normalisées par rapport au contrôle (ChIP anti-HA sur cellules Min6b1 non transfectées) et permettent l'identification de 5104 pics, ou par rapport à l'échantillon (ChIP anti-HA sur cellules Min6b1 transfectées avec le plasmide pCMVTag2A-Flag-3HA-Rfx6-2A-mCherry) et permettent l'identification de 23308 pics. Les jeux de données ChIP

HA-RFX6 et Contrôle CHIP HA sont déposés sur la base de données GEO (Gene Expression Omnibus) sous le numéro d'accès GSE62844.

Dans un premier temps, les deux jeux de données anti-RFX6 et anti-HA ont été analysés séparément. Les modèles de pics obtenus pour les CHIP RFX6 et HA-RFX6 sont présentés en **Figure 28**. Dans chaque cas, le modèle de pic définitif représenté en noir présente la morphologie d'un pic attendu pour un facteur de transcription puisqu'il est fin et peu étendu (Kidder et al, 2011) et présente un sommet qui peut être considéré comme le site de fixation exact du facteur de transcription RFX6.

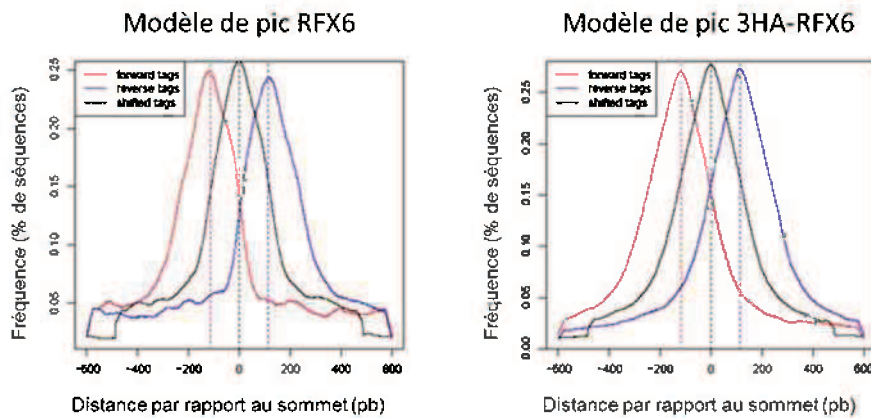


Figure 28 : Modèles de pics obtenus pour les CHIP RFX6 (gauche) et HA-RFX6 (droite) sur cellules Min6b1. Les modèles de pic pour les séquences sens (rouge) et anti-sens (bleu) sont compilés pour obtenir le modèle de pic définitif pour chaque expérience (noir) dont le sommet est considéré comme le site de fixation exact du facteur de transcription RFX6.

Les résultats obtenus pour chaque expérience sont résumés **Figure 29** et **Figure 30**. Après normalisation des données pour le CHIP anti-RFX6, 1552 pics ont été identifiés, ils sont regroupés autour des sites d'initiation de la transcription des gènes (TSS) (**Figure 29A**) et, pour 43 % d'entre eux, situés dans les régions promotrices et au niveau des TSS. 24 et 27% des pics sont situés respectivement dans les régions intergéniques ou introniques du génome (**Figure 29B**). Un motif de type « RFX » est retrouvé de manière centralement enrichie sous 33% des pics de CHIP Seq analysés par MEME-CHIP (**Matériels et méthodes C.3**). Un second motif de type NRF1 (Nuclear Respiratory factor 1) est également retrouvé dans 28% des pics analysés (**Figure 29C**).

Cette protéine NRF1 régule un ensemble de gènes clés impliqués dans diverses voies cellulaires dont la réponse au stress oxydatif, le développement des os ou encore l'inflammation (Biswas & Chan, 2010). Ce gène a également été identifié comme régulateur clé de nombreux gènes dérégulés spécifiquement chez des patients diabétiques et pré-diabétiques (Patti, 2004). Ceci suggère que RFX6 et NRF1 pourraient ainsi cibler et co-réguler des gènes impliqués dans la fonction beta.

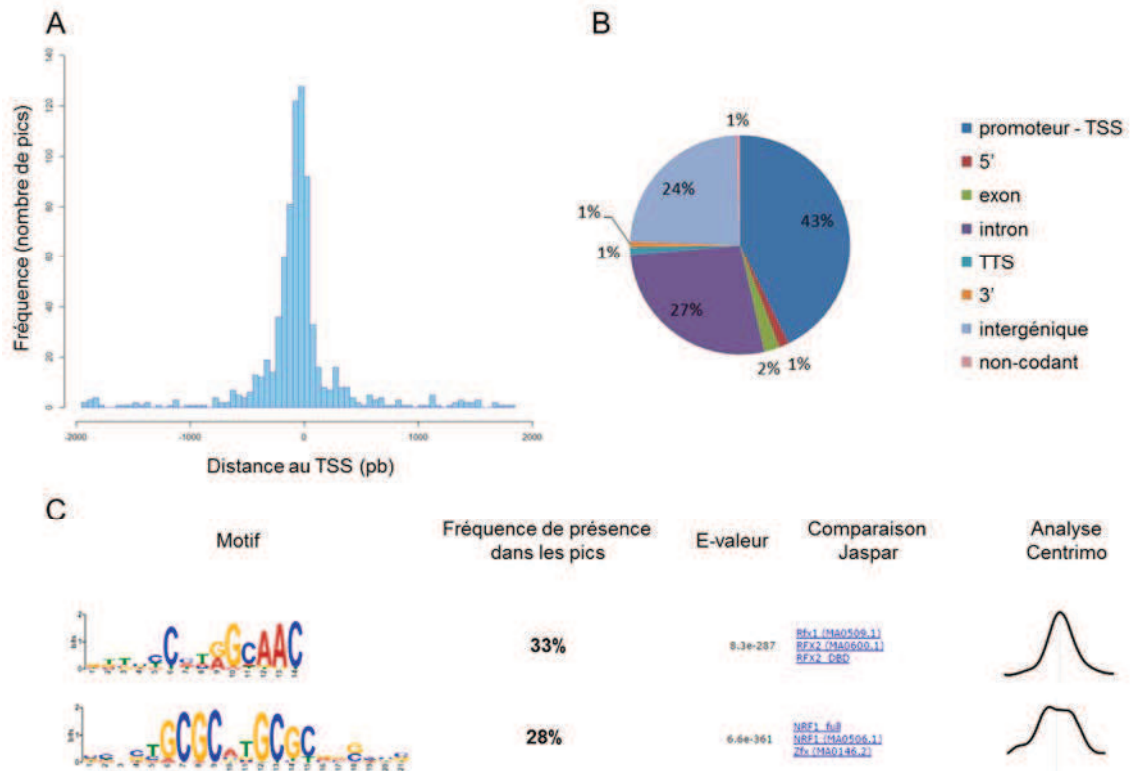


Figure 29 : 1552 sites de liaison de RFX6 sont identifiés dans les cellules Min6b1. (A) Distribution des sites de fixation de RFX6 autour des sites d'initiation de la transcription (TSS) des gènes. **(B)** Distribution des sites de fixation de RFX6 au sein du génome (TSS = site d'initiation de la transcription, TTS = site de terminaison de la transcription, 5' = région 5'UTR, 3' = région 3'UTR, non-codant = gènes non-codant). **(C)** L'analyse de motif *de novo* sous les sites de fixation de RFX6 (analyse de 100 pb centrées sous le sommet de 1000 pics choisis aléatoirement parmi les pics 1552 totaux) révèle un motif consensus de fixation de RFX6 centralement enrichi sous 33% des pics de CHIP Seq.

Concernant les résultats du CHIP HA-RFX6, 5104 pics ont été identifiés avec une normalisation à l'aide du contrôle anti-HA dans les cellules Min6b1 non transfectées. Dans cette analyse globale de CHIP, les 5104 pics HA-RFX6 sont répartis de manière plus large autour des TSS (**Figure 30A**) et une proportion plus faible de pics, soit 11%, sont situés dans les régions promotrices et au niveau des TSS en comparaison avec le CHIP RFX6 dans le même modèle cellulaire. *A contrario*, une proportion plus importante de pics est située dans les régions intergéniques (32%) ou introniques (44%) du génome (**Figure 30B**). De même que pour les sites de liaison de RFX6 dans les cellules Min6b1, un motif de type « RFX » est retrouvé sous 59% des pics de CHIP Seq analysés (**Figure 30C**). Cependant l'analyse Centrimo, qui détermine la position du motif dans la séquence analysée, démontre que la présence de ce motif est moins centrale par rapport au sommet du pic.

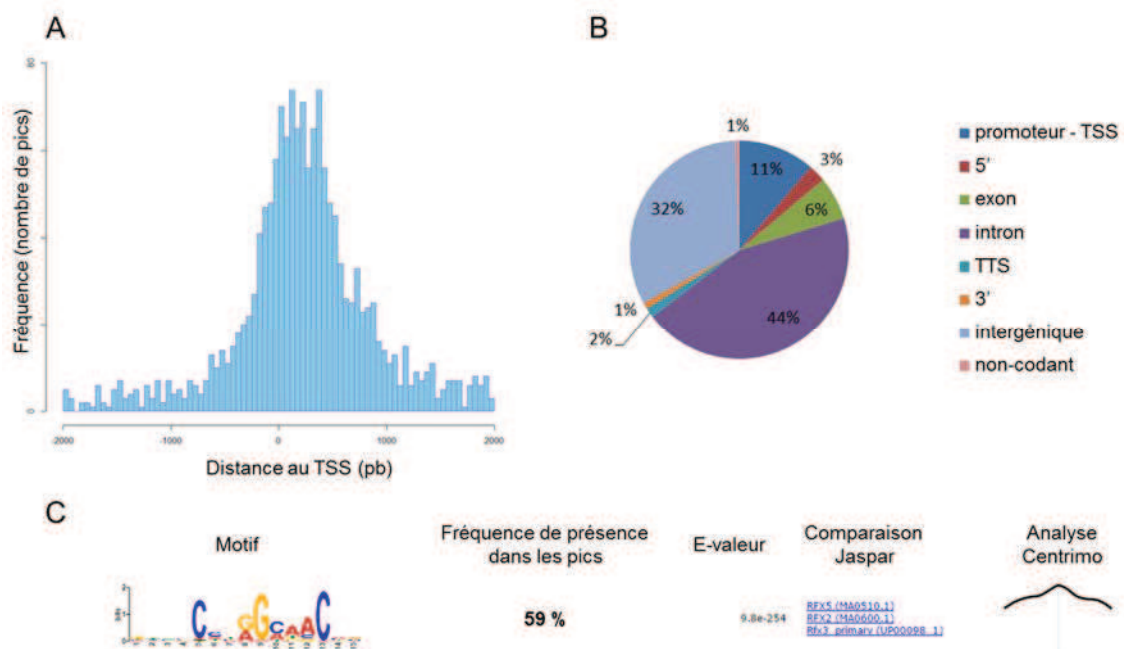


Figure 30 : 5104 sites de liaison de 3HA-RFX6 sont identifiés dans les cellules Min6b1. 5104 pics identifiés dans cette expérience sont analysés ici. **(A)** Distribution des sites de fixation de 3HA-RFX6 autour des sites d'initiation de la transcription (TSS) des gènes. **(B)** Distribution des sites de fixation de 3HA-RFX6 au sein du génome (TSS = site d'initiation de la transcription, TTS = site de terminaison de la transcription, 5' = région 5'UTR, 3' = région 3'UTR, non-codant = gènes non-codant). **(C)** L'analyse de motif *de novo* sous les sites de fixation de 3HA-RFX6 (analyse de 100 pb centrées sous le sommet de 1000 pics choisis aléatoirement parmi les pics 5104 totaux) révèle un motif consensus de fixation de 3HA-RFX6 enrichi sous 59% des pics de CHIP Seq.

Un nombre de pic différent est ainsi observé entre les deux expériences de ChIP RFX6 (1552 pics) et HA-RFX6 (5104 pics) dans les cellules Min6b1. On remarque également une différence de répartition autour des TSS des gènes et de localisation génomiques des sites de liaison identifiés. Les différences observées entre les deux expériences de ChIP RFX6 et HA-RFX6 dans les cellules Min6b1 peuvent être expliquées par l'utilisation de deux anticorps différents (HA 12CA5 et Rfx6 2767) dont la spécificité et l'efficacité en ChIP est inévitablement inégale. D'autre part, le ChIP HA-RFX6 est réalisé sur des cellules Min6b1 transfectées qui surexpriment une protéine de fusion Flag-3HA-RFX6, ce qui pourrait impliquer que des cibles, qui ne sont pas liées par la protéine endogène RFX6 dans la cellule, peuvent être révélées dans cette expérience. Ceci suggère que le ChIP HA-RFX6 est moins spécifique que le ChIP RFX6. Cependant le plus grand nombre de pic identifiés dans le ChIP HA-RFX6 pourrait également indiquer une plus grande sensibilité de cette expérience. En effet, l'exemple du gène de la glucokinase, identifié comme cible dans le ChIP HA-RFX6 mais pas dans le ChIP RFX6 peut être cité (**Figure 46A**). Ce gène cible a effectivement été confirmé en ChIP quantitatif anti-RFX6 dans les cellules Min6b1 (voir papier Cell Reports, paragraphe A.3.e). D'autre part, l'étude des gènes ciblés et régulés dans le modèle de souris n'exprimant pas Rfx6 spécifiquement dans les cellules beta (paragraphe A.3.c) atténue le problème du bruit de fond plus important de l'anticorps anti-HA et du modèle de surexpression utilisé pour cette étude.

Pour nous assurer d'identifier les cibles de RFX6 de manière juste dans les cellules Min6b1, nous avons ensuite choisi de nous concentrer sur les pics communs entre les deux jeux de données. Un pic est considéré comme commun entre deux ChIP si la séquence sous les pics comporte au minimum une base en commun.

Le premier jeu de données choisi est celui du ChIP HA-RFX6 normalisé par rapport à l'échantillon qui permet l'identification d'un beaucoup plus grand nombre de pics (23308) par rapport à une normalisation à l'aide du contrôle (5104). En effet, le choix de la normalisation qui est lié au nombre de séquences de chaque échantillon permet ainsi d'identifier plus ou moins de cibles en fonction de ce qui est recherché. Afin de ne perdre aucune information et identifier un maximum de cibles de RFX6, les 23308 pics du ChIP HA-RFX6 ont été comparés aux 1552 pics identifiés dans le ChIP RFX6 et ce sont 863 pics communs qui ont été identifiés (**Figure 31**). L'analyse de motif au niveau de ces sites de liaison de RFX6 montre qu'un motif de type « RFX » est centralement enrichi sous le sommet du pic dans 55% des sites de liaison identifiés. La présence du motif de liaison de type « NRF1 » est également retrouvé dans 17% des pics analysés pour ce jeu de données. Enfin, les pics

communs sont à 33% présents dans les régions promotrices et au niveau des TSS des gènes, alors que 32% et 29% sont respectivement présents dans les régions introniques et intergéniques du génome.

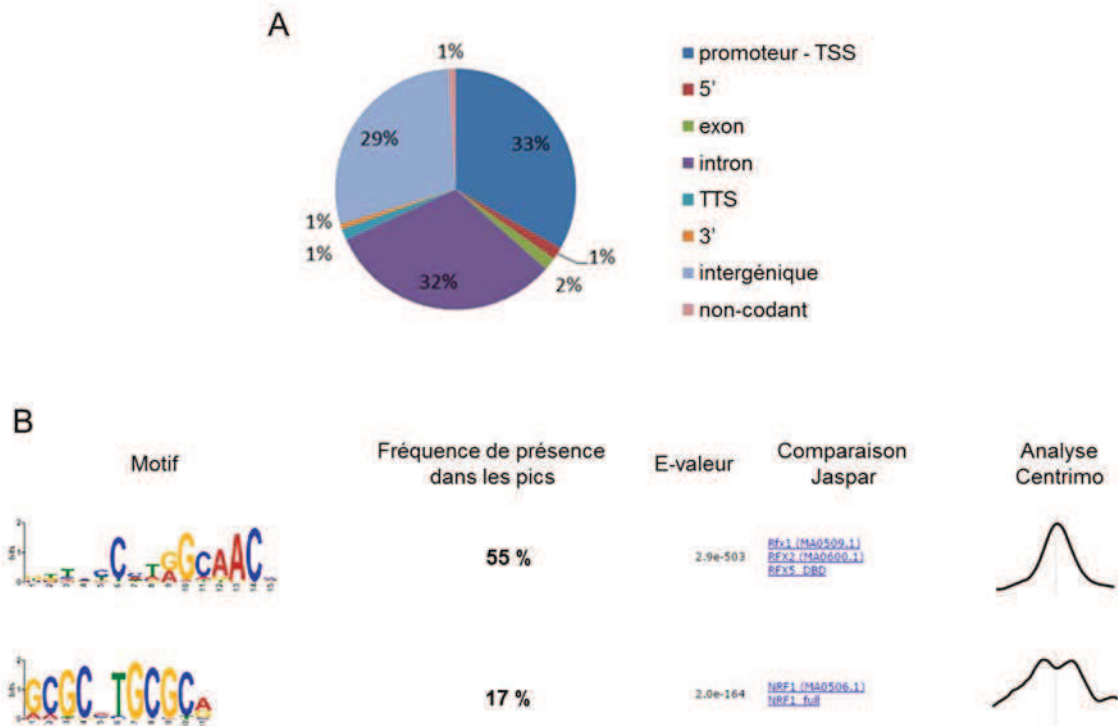


Figure 31 : 863 sites de liaison de 3HA-RFX6 et RFX6 communs sont identifiés dans les cellules Min6b1. 863 pics communs identifiés dans cette expérience sont analysés ici. **(A)** Distribution des sites de fixation communs de RFX6 au sein du génome (TSS = site d'initiation de la transcription, TTS = site de terminaison de la transcription, 5' = région 5'UTR, 3' = région 3'UTR, non-codant = gènes non-codant). **(B)** L'analyse de motif *de novo* sous les sites de fixation de RFX6 révèle un motif consensus commun de fixation (analyse de 100 pb centrées sous le sommet des 863 pics) centralement enrichi sous 55% des pics de ChIP Seq.

Ces premières expériences de ChIP RFX6 et HA-RFX6 ont ainsi permis d'identifier les sites de fixation de RFX6 au sein du génome des cellules Min6b1. Les motifs consensus de fixation de RFX6 ont été définis dans les cellules Min6b1 et corrélés aux motifs de type RFX déjà connus suggérant que la stratégie de recherche mise en place permet l'identification des gènes cibles spécifiques de RFX6 dans la lignée cellulaire Min6b1.

A.3.c. Identification des gènes ciblés dans les cellules Min6b1 et régulés dans les cellules beta pancréatiques par RFX6

Pour identifier les gènes directement régulés par RFX6, il est nécessaire de connaître ceux dont l'expression est affectée lorsque Rfx6 n'est plus exprimé. Pour cela, des études transcriptomiques réalisées par Julie Piccand, ancienne membre de l'équipe, à partir d'ilots pancréatiques de souris n'exprimant pas Rfx6 spécifiquement dans les cellules beta ont été réalisées. L'expression de Rfx6 est ainsi inactivée spécifiquement dans les cellules beta après traitement des souris Rfx6 fl/fl ; Ins1-CrERT au tamoxifène et le RNA Seq est réalisé par la suite sur les ilots pancréatiques purifiés.

Une comparaison des données de ChIP Seq décrites précédemment avec ces données de RNA Seq chez les souris Rfx6 fl/fl ; Ins1-CrERT (KO induit) appelées souris delta-beta m'ont permis de déterminer un répertoire de gènes directement régulés par RFX6.

Cette étude a été réalisée avant séquençage de l'input et du contrôle et donc avant normalisation des données de ChIP Seq sur les gènes associés aux 22365 pics HA-RFX6 et et aux 1671 pics communs RFX6 et HA-RFX6 (**Tableau 8**). L'attribution d'un pic de ChIP Seq avec un gène est réalisée par annotation des données de ChIP Seq comme indiqué dans le paragraphe Matériels et méthodes C.2. Par la suite, cette analyse ne sera pas reconduite avec les données normalisées au niveau du modèle cellulaire Min6b1 puisque nous avons pu identifier et nous sommes concentrés sur les cibles de RFX6 *in vivo* dans les ilots pancréatiques de souris (**paragraphe Résultats A.4**).

La comparaison des gènes ciblés, identifiés dans le ChIP Seq HA-RFX6, et des gènes régulés dans la cellule beta, identifiés dans le RNA Seq, a permis de mettre en évidence 1540 gènes potentiellement directement régulés par Rfx6 dont 377 gènes sont activés et 1163 gènes sont réprimés par RFX6 (**Figure 32**).

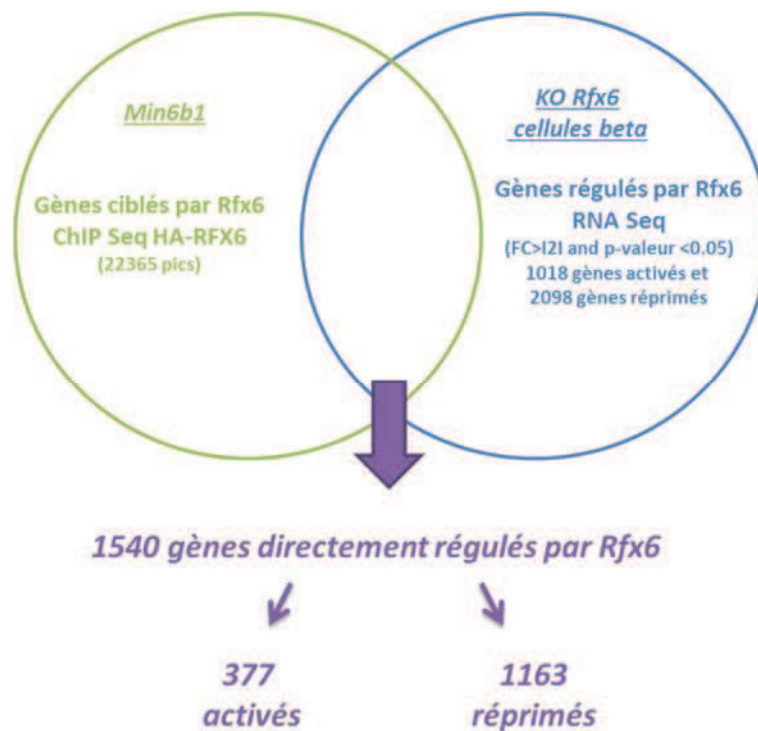


Figure 32 : Identification des gènes ciblés dans la cellule Min6b1 et régulés dans les cellules beta pancréatiques de souris par RFX6. La compilation des données de transcriptomique (gènes dont l'expression est dérégulée (RNA Seq) dans les cellules beta pancréatiques chez la souris delta beta, $FC > 121$ et $p\text{-valeur} < 0,05$) et des sites de liaisons de RFX6 (gènes associés aux 22365 pics identifiés pour le ChIP Seq HA-RFX6 dans les cellules Min6b1 avant séquençage de l'input) permet l'identification de 1540 gènes potentiellement directement régulés par RFX6 dont 377 sont activés et 1163 sont réprimés.

Parmi ceux-ci, les 40 gènes dont les variations d'expression positives ou négatives les plus importantes observées par RNA Seq dans les cellules beta lorsque Rfx6 est inactivé sont présentés en **Figure 33**.

Gène	Identifiant Ensembl	Description	Moyenne Expression CTRL	Moyenne Expression MUT	FC	p-valeur	Position génomique du pic de ChIP Seq anti-HA-RFX6	Pic commun ChIP Seq RFX6 et HA-Rfx6
Cnn1	ENSMUSG0000001349	calponin 1	1.80	78.28	16.0	1.16E-17	Intergénique	Oui
Cyt11	ENSMUSG00000062329	cytokine-like 1	3.04	113.11	12.9	1.37E-14	5'UTR	Non
Eln	ENSMUSG00000029675	elastin	19.37	308.59	8.6	7.73E-12	Intron et intergénique	Non
Tagln	ENSMUSG00000032085	transgelin	76.41	889.51	8.3	9.81E-16	Exon et intron	Non
Vwf	ENSMUSG0000001930	Von Willebrand factor homolog	93.75	1156.24	8.1	2.03E-13	Intron	Non
Nnmt	ENSMUSG00000032271	nicotinamide N-methyltransferase	6.47	155.95	7.1	6.43E-08	Intron	Non
Plin4	ENSMUSG00000002831	perilipin 4	0.92	17.25	7.0	3.45E-08	Intron et exon	Non
Igfbbp6	ENSMUSG00000023046	insulin-like growth factor binding protein 6	26.84	736.78	7.0	1.46E-07	Promoteur-TSS et exon	Non
Mpo	ENSMUSG00000009350	myeloperoxidase	7.49	257.21	7.0	2.22E-07	Intron	Non
Lrrn4	ENSMUSG00000043110	leucine rich repeat neuronal 4	2.44	88.62	6.8	3.23E-07	Exon	Non
Camp	ENSMUSG00000038357	cathelicidin antimicrobial peptide	1.77	102.71	6.8	4.30E-07	Intergénique	Non
Padi4	ENSMUSG00000025330	peptidyl arginine deiminase, type IV	1.69	50.71	6.7	3.52E-07	Intron et intergénique	Non
Pdpn	ENSMUSG00000028583	podoplanin	13.10	204.62	6.6	5.59E-08	Intron	Non
Wisp2	ENSMUSG00000027656	WNT1 inducible signaling pathway protein 2	3.82	40.99	6.5	1.07E-09	Promoteur-TS ; intron et intergénique	Non
Fbln2	ENSMUSG00000064080	fibulin 2	250.30	2858.46	6.5	5.45E-09	Exon ; intergénique et intron	Non
Lgals1	ENSMUSG00000068220	lectin, galactose binding, soluble 1	249.92	3040.68	6.5	1.38E-08	Promoteur-TSS et intron	Non
Upk3b	ENSMUSG00000042985	uropodin 3B	3.73	108.52	6.5	6.24E-07	Intergénique	Non
C1qtnf9	ENSMUSG00000071347	C1q and tumor necrosis factor related protein 9	8.23	104.01	6.4	2.19E-08	Promoteur-TSS	Non
Col18a1	ENSMUSG00000001435	collagen, type XVIII, alpha 1	31.05	338.65	6.4	3.13E-09	Intron et TTS	Non
A2m	ENSMUSG00000030111	alpha-2-macroglobulin	0.17	7.11	6.4	8.84E-07	Exon	Non

Gène	Identifiant Ensembl	Description	Moyenne Expression CTRL	Moyenne Expression MUT	FC	p-valeur	Position génomique du pic de ChIP Seq anti-HA-RFX6	Pic commun ChIP Seq RFX6 et HA-Rfx6
Apoa4	ENSMUSG00000032080	apolipoprotein A-IV	50.61	0.00	-31.9	5.48E-26	Intergénique	Non
Clca4	ENSMUSG00000037033	chloride channel calcium activated 4	24.47	0.00	-25.0	1.89E-21	Intron	Non
Alpi	ENSMUSG00000079440	alkaline phosphatase, intestinal	135.15	0.51	-20.9	3.58E-17	Exon	Non
2210415F13Rik	ENSMUSG00000013643	RIKEN cDNA 2210415F13 gene	59.02	0.70	-18.3	5.53E-19	Intron	Non
Mlxipl	ENSMUSG00000005373	MLX interacting protein-like	1595.32	84.51	-17.7	5.35E-148	Intron	Oui
Prap1	ENSMUSG00000025467	proline-rich acidic protein 1	214.03	0.00	-17.6	5.18E-15	Promoteur-TSS et intron	Oui
Rbp2	ENSMUSG00000032454	retinol binding protein 2, cellular	183.18	4.84	-16.9	4.85E-21	Intergénique et intron	Non
Cdh17	ENSMUSG00000028217	cadherin 17	48.43	0.87	-16.0	9.53E-17	Intron	Non
Muc2	ENSMUSG00000025515	mucin 2	22.21	0.28	-13.2	2.83E-13	Intergénique et exon	Non
Nlrp6	ENSMUSG00000038745	NLR family, pyrin domain containing 6	14.75	0.43	-10.9	6.37E-13	Exon	Non
Arm4	ENSMUSG00000061802	armadillo repeat containing 4	30.64	2.36	-9.7	9.06E-25	Promoteur-TSS et intron	Non
Grin2c	ENSMUSG00000020734	glutamate receptor, ionotropic, NMDA2C (epsilon 3)	92.57	8.04	-8.9	5.06E-21	Intron	Non
H2-T3	ENSMUSG00000054128	histocompatibility 2, T region locus 3	20.73	0.30	-8.8	1.78E-09	Intergénique	Non
Gm1123	ENSMUSG00000044860	predicted gene 1123	23.46	1.41	-7.9	3.90E-11	Promoteur-TSS	Non
Myo1a	ENSMUSG00000025401	myosin IA	47.55	2.98	-7.8	3.15E-10	Intron	Non
Cdx2	ENSMUSG00000029646	caudal type homeobox 2	7.31	0.00	-7.6	4.92E-08	Exon et intergénique	Oui
Dcdc5	ENSMUSG00000074981	doublecortin domain containing 5	81.94	9.17	-7.5	1.12E-23	Intergénique	Non
Pck1	ENSMUSG00000027513	phosphoenolpyruvate carboxykinase 1, cytosolic	23.55	1.19	-6.9	4.07E-08	Intergénique	Non
Cdh7	ENSMUSG00000026312	cadherin 7, type 2	445.52	54.85	-6.7	7.77E-19	Promoteur-TSS	Non
Cdx1	ENSMUSG00000024619	caudal type homeobox 1	7.08	0.00	-6.3	1.04E-06	Intron	Non

Figure 33 : Gènes ciblés dans la cellule Min6b1 et régulés dans les cellules beta pancréatiques par RFX6. Les 20 gènes les plus surexprimés (en haut) ou sous-exprimés (en bas) (RNA Seq) dans les cellules beta pancréatiques chez la souris n'exprimant pas Rfx6 dans les cellules beta (MUT) par rapport à la souris sauvage (WT) sont listés avec leur variation d'expression (FC) et la p-valeur associée. La position du site de fixation de RFX6 identifié lors du ChIP HA-RFX6 dans les cellules Min6b1 et sa présence commune dans le ChIP RFX6 dans les Min6b1 est indiquée pour chaque gène.

Enfin, un classement des gènes ciblés dans la cellule Min6b1 et régulés dans les cellules beta pancréatiques par RFX6 est effectué par voie biologique (Gene Ontology) (**Figure 34**). Les 10 gènes les plus dérégulés dans chaque liste sont présentés en **Figure 34A** et la totalité des gènes impliqués dans la sécrétion d'insuline et l'homéostasie du glucose sont présentés en **Figure 34B**.

A

Voies biologiques	% des gènes ciblés et régulés par RFX6	Gènes ciblés et régulés par RFX6
Homéostasie du glucose	1.7	Aspscr1 ; Bad ; Met ; Foxa3 ; Fgfr4 ; Bhlha15 ; Vgf ; Sh2b2 ; Slc18a2 ; Cartpt
		Mlxipl ; Pck1 ; Cacna1e ; Ffar1 ; Cacna1c ; Pax6 ; Neurod1 ; Gck ; Alms1 ; Crtc2
Sécrétion d'insuline	3.4	Kcnq1 ; Il1rn ; Blk ; Arrb1 ; Rest ; Slc18a2 ; Cartpt ; Cckar ; Cplx1 ; Ildr2
		Nnat ; Cacna1e ; Ffar1 ; Gpr ; Fam3b ; Cacna1c ; Ucn3 ; Pde4c ; Oit1 ; Gip
Régulation de la transcription	36.9	Barx1 ; Wt1 ; Col1a1 ; Aebp1 ; Ifi205 ; Sfrp2 ; Trf ; Esr1 ; Padi4 ; Cyt11
		Mlxipl ; Cdx2 ; Cdx1 ; Nr1i3 ; Foxn1 ; Mesp1 ; Sap25 ; Barhl1 ; Per3 ; Creb3l3
Cycle cellulaire	16	Kif2c ; Ckap2 ; F630043A04Rik ; Dhh ; Tpx2 ; Gas1 ; Kif18b ; Cdh13 ; Rprm ; Pdpn
		Mlxipl ; Crocc ; Csp1 ; Inha ; 2410004A20Rik ; Nanog ; Hmga2 ; Apex2 ; Prkcq ; Appl2

SECRETION D'INSULINE

Gène	Identifiant Ensembl	Description	Moyenne Expression CTRL	Moyenne Expression MUT	FC	p-valeur	Position génomique du pic de ChIP Seq anti-HA-RFX6	Pic commun ChIP Seq RFX6 et HA-Rfx6
Nnat	ENSMUSG00000067786	neuronatin	723.43	130.72	-5.0	1.98E-18	5'UTR	Non
Cacna1e	ENSMUSG00000004110	calcium channel, voltage-dependent, R type, alpha 1E subunit	46.15	7.09	-4.9	4.89E-09	Intron	Non
Ffar1	ENSMUSG00000044453	free fatty acid receptor 1	975.95	232.17	-4.0	9.24E-34	Promoteur-TSS	Non
Gipr	ENSMUSG00000030406	gastric inhibitory polypeptide receptor	1531.66	498.03	-3.0	8.22E-20	Promoteur-TSS, TTS, intron et intergénique	Oui
Fam3b	ENSMUSG00000022938	family with sequence similarity 3, member B	5.94	0.43	-2.5	0.016000179	Intron	Non
Cacna1c	ENSMUSG00000051331	calcium channel, voltage-dependent, L type, alpha 1C subunit	337.79	148.65	-2.2	1.54E-05	Promoteur-TSS, intron et intergénique	Oui
Ucn3	ENSMUSG00000044988	urocortin 3	3889.29	1769.32	-2.1	2.45E-06	Intergénique	Non
Pde4c	ENSMUSG00000031842	phosphodiesterase 4C, cAMP specific	158.42	73.93	-2.1	8.07E-11	Promoteur-TSS	Non
Oit1	ENSMUSG00000021749	oncoprotein induced transcript 1	149.37	70.47	-2.0	0.000169382	Intergénique	Non
Gip	ENSMUSG00000014351	gastric inhibitory polypeptide	22.28	0.00	-2.0	0.042455392	Intergénique et intron	Non
Hmgn3	ENSMUSG00000066456	high mobility group nucleosomal binding domain 3	7401.27	3914.30	-1.9	5.82E-17	Intergénique et intron	Non
Neurod1	ENSMUSG00000034701	neurogenic differentiation 1	1764.47	971.99	-1.8	7.27E-09	Exon	Non
Ghr	ENSMUSG00000051136	growth hormone secretagogue receptor	44.13	25.18	-1.7	0.000499342	Intergénique et intron	Non
Syt9	ENSMUSG00000062542	synaptotagmin IX	156.39	87.70	-1.7	0.014305292	Intergénique	Non
Pfkfb2	ENSMUSG00000026409	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 2	2962.69	1739.12	-1.7	9.87E-07	Intron	Non
Glp1r	ENSMUSG00000024027	glucagon-like peptide 1 receptor	5989.60	3404.97	-1.7	0.010632697	Intergénique et intron	Non
Gck	ENSMUSG00000041798	glucokinase	2130.97	1344.81	-1.6	8.11E-05	Promoteur-TSS, TTS, intron et intergénique	Non
Irs2	ENSMUSG00000038894	insulin receptor substrate 2	471.33	315.47	-1.5	0.000936225	Promoteur-TSS, intron et intergénique	Non
Htt	ENSMUSG00000029104	huntingtin	1263.38	848.02	-1.5	0.005451179	3'UTR, intergénique, exon, intron	Non
Rapgef4	ENSMUSG00000049044	Rap guanine nucleotide exchange factor (GEF) 4	6952.18	4735.31	-1.5	9.86E-05	Intron et intergénique	Oui
Sirt4	ENSMUSG00000029524	sirtuin 4 (silent mating type information regulation 2 homolog) 4 (S. cerevisiae)	374.51	261.68	-1.4	0.000374994	Intron et intergénique	Oui
Abcc8	ENSMUSG00000040136	ATP-binding cassette, sub-family C (CFTR/MRP), member 8	6540.14	4579.93	-1.4	0.013125843	Intron, 5'UTR et TTS	Oui
Mgea5	ENSMUSG00000025220	meningioma expressed antigen 5 (hyaluronidase)	1086.36	774.29	-1.4	1.18E-06	Intron	Non
Dgat1	ENSMUSG00000022555	diacylglycerol O-acyltransferase 1	760.11	545.88	-1.4	0.000198633	Promoteur-TSS et intron	Non
Rims2	ENSMUSG00000037386	regulating synaptic membrane exocytosis 2	565.95	410.57	-1.4	0.032127359	Intron	Non
Snap25	ENSMUSG00000027273	synaptosomal-associated protein 25	1976.84	1485.16	-1.3	0.045847988	Intron	Oui
Rbp4	ENSMUSG00000024990	retinol binding protein 4, plasma	1972.41	1536.61	-1.3	0.032673226	Exon et promoteur-TSS	Non
Capn10	ENSMUSG00000026270	calpain 10	808.65	700.32	-1.2	0.006094375	Intron et gène non-codant	Non
Glud1	ENSMUSG00000021794	glutamate dehydrogenase 1	5609.92	6657.22	1.2	0.015041828	Promoteur-TSS, TTS, et intergénique	Non
Oxct1	ENSMUSG00000022186	3-oxoacid CoA transferase 1	5552.09	6691.36	1.2	0.00382301	Gène non codant	Non
Lrp5	ENSMUSG00000024913	low density lipoprotein receptor-related protein 5	252.83	324.63	1.3	0.039717232	Intron	Non
Bad	ENSMUSG00000024959	BCL2-associated agonist of cell death	487.89	635.72	1.3	0.000308632	Promoteur-TSS et intron	Oui
Ppp3ca	ENSMUSG00000028161	protein phosphatase 3, catalytic subunit, alpha isoform	1499.58	2172.78	1.4	0.000211182	Intergénique et intron	Non
Serp1	ENSMUSG00000027808	stress-associated endoplasmic reticulum protein 1	6684.80	10255.67	1.5	5.92E-06	Promoteur-TSS	Non
Cpt1a	ENSMUSG00000024900	carnitine palmitoyltransferase 1a, liver	266.92	439.29	1.6	0.014039923	Promoteur-TSS et intergénique	Non
Pim3	ENSMUSG00000035828	proviral integration site 3	610.11	1082.10	1.7	0.012859894	Intron et intergénique	Oui
Cplx3	ENSMUSG00000039714	complexin 3	11.96	23.36	1.8	0.033898869	Promoteur-TSS	Non
Inhbb	ENSMUSG00000037035	inhibin beta-B	29.03	70.58	2.1	0.008277884	Intron et 5'UTR	Non
Cd38	ENSMUSG00000029084	CD38 antigen	52.46	130.16	2.2	0.00536328	Intron	Non
Pde8b	ENSMUSG00000021684	phosphodiesterase 8B	36.60	82.69	2.2	2.84E-06	Intron et 5'UTR	Non
Nr0b2	ENSMUSG00000037583	nuclear receptor subfamily 0, group B, member 2	18.01	47.49	2.2	0.005468259	Intron et exon	Oui
Vgf	ENSMUSG00000037428	VGF nerve growth factor inducible	92.06	228.37	2.4	8.02E-09	Intron	Oui
Kcnq1	ENSMUSG00000009545	potassium voltage-gated channel, subfamily Q, member 1	69.65	180.15	2.4	5.38E-06	Promoteur-TSS, gène non codant, intron et intergénique	Oui
Il1rn	ENSMUSG00000026981	interleukin 1 receptor antagonist	23.37	79.94	2.4	0.009157639	Intergénique	Non
Blk	ENSMUSG00000014453	B lymphoid kinase	1.54	10.15	2.5	0.018155648	Intergénique, exon, intron, et promoteur-TSS	Oui
Arrb1	ENSMUSG00000018909	arrestin, beta 1	753.76	1948.76	2.5	1.35E-08	Intron et exon	Oui
Rest	ENSMUSG00000029249	RE1-silencing transcription factor	18.88	58.54	2.6	0.000489274	Intron	Non
Slc18a2	ENSMUSG00000025094	solute carrier family 18 (vesicular monoamine), member 2	1.91	10.71	3.0	0.002831966	Exon	Non
Cartpt	ENSMUSG00000021647	CART prepropeptide	10.06	43.55	3.8	2.95E-09	Intergénique et promoteur-TSS	Non
Cckar	ENSMUSG00000029193	cholecystokinin A receptor	160.40	931.47	3.8	3.32E-05	Intergénique	Non
Cplx1	ENSMUSG00000033615	complexin 1	3.56	21.81	4.3	1.53E-06	Promoteur-TSS	Non
Ildr2	ENSMUSG00000040612	immunoglobulin-like domain containing receptor 2	40.11	321.18	4.4	1.26E-05	Exon	Non

B

HOMEOSTASIE DU GLUCOSE								
Gène	Identifiant Ensembl	Description	Moyenne Expression CTRL	Moyenne Expression MUT	FC	p-valeur	Position génomique du pic de ChIP Seq anti-HA-RFX6	Pic commun ChIP Seq RFX6 et HA-Rfx6
Mlxipl	ENSMUSG00000005373	MLX interacting protein-like	1595.32	84.51	-17.7	5.35E-148	Intron	Oui
Pck1	ENSMUSG00000027513	phosphoenolpyruvate carboxykinase 1, cytosolic	23.55	1.19	-6.9	4.07E-08	Intergénique	Non
Cacna1e	ENSMUSG00000004110	calcium channel, voltage-dependent, R type, alpha 1E subunit	46.15	7.09	-4.9	4.89E-09	Intron	Non
Ffar1	ENSMUSG00000004453	free fatty acid receptor 1	975.95	232.17	-4.0	9.24E-34	Promoteur-TSS	Non
Cacna1c	ENSMUSG00000051331	calcium channel, voltage-dependent, L type, alpha 1C subunit	337.79	148.65	-2.2	1.54E-05	Promoteur-TSS, intron et intergénique	Oui
Pax6	ENSMUSG00000027168	paired box gene 6	6600.11	3210.75	-2.0	7.73E-14	Intron	Non
Neurod1	ENSMUSG00000034701	neurogenic differentiation 1	1764.47	971.99	-1.8	7.27E-09	Exon	Non
Gck	ENSMUSG000000041798	glucokinase	2130.97	1344.81	-1.6	8.11E-05	Promoteur-TSS, TTS, intron et intergénique	Non
Alms1	ENSMUSG000000063810	Alstrom syndrome 1 homolog (human)	149.27	94.62	-1.6	0.000153237	Exon	Non
Crtc2	ENSMUSG000000027936	CREB regulated transcription coactivator 2	113.83	80.42	-1.4	0.014315956	Intron	Non
Rph3al	ENSMUSG000000020847	rabphilin 3A-like (without C2 domains)	2893.11	2180.00	-1.3	8.90E-03	Intron et intergénique	Oui
Rbp4	ENSMUSG000000024990	retinol binding protein 4, plasma	1972.41	1536.61	-1.3	3.27E-02	Exon et promoteur-TSS	Non
Stk11	ENSMUSG000000003068	serine/threonine kinase 11	1780.29	1450.70	-1.2	2.59837E-08	Intron	Non
Inpp5k	ENSMUSG000000006127	inositol polyphosphate 5-phosphatase K	375.92	322.35	-1.2	3.61E-02	Promoteur-TSS et intron	Non
Sirt1	ENSMUSG000000020063	sirtuin 1 (silent mating type information regulation 2, homolog) 1 (S. cerevisiae)	265.22	229.93	-1.2	4.44E-02	Promoteur-TSS	Non
Aspscr1	ENSMUSG000000025142	alveolar soft part sarcoma chromosome region, candidate 1 (human)	1032.65	907.14	-1.1	9.46E-03	Exon et intron	Non
Bad	ENSMUSG000000024959	BCL2-associated agonist of cell death	487.89	635.72	1.3	3.09E-04	Promoteur-TSS et intron	Oui
Met	ENSMUSG000000009376	met proto-oncogene	282.31	381.08	1.3	0.000117821	5'UTR	Non
Foxa3	ENSMUSG000000040891	forkhead box A3	183.05	253.14	1.4	1.47E-05	Promoteur-TSS	Oui
Fgfr4	ENSMUSG00000005320	fibroblast growth factor receptor 4	72.39	132.91	1.8	6.41E-05	Intron et intergénique	Non
Bhlha15	ENSMUSG000000052271	basic helix-loop-helix family, member a15	260.44	490.27	1.8	9.31E-07	Intron	Non
Vgf	ENSMUSG000000037428	VEGF nerve growth factor inducible	92.06	228.37	2.4	8.02E-09	Intron	Oui
Sh2b2	ENSMUSG000000005057	SH2B adaptor protein 2	4.53	17.06	2.6	0.004483866	Intron	Non
Slc18a2	ENSMUSG000000025094	solute carrier family 18 (vesicular monoamine), member 2	1.91	10.71	3.0	0.002831966	Exon	Non
Cartpt	ENSMUSG000000021647	CART prepropeptide	10.06	43.55	3.8	2.95E-09	Intergénique et promoteur-TSS	Non

Figure 34 : Classement par voies biologiques des gènes ciblés dans la cellule Min6b1 et régulés dans les cellules beta pancréatiques par RFX6. (A) Les gènes ciblés par RFX6 (ChIP HA-RFX6 dans les cellules Min6b1) et surexprimés (en bleu) ou sous-exprimés (en orange) en absence de RFX6 (RNA Seq) dans les cellules beta pancréatiques sont listés et classés par voies biologiques (Gene Ontology). Le pourcentage du nombre total de gènes ciblés et régulés par RFX6 (1540 gènes au total) et les 10 gènes les plus dérégulés dans chaque liste sont présentés. Les gènes en gras présentent une variation d'expression significative supérieure à deux (RNA Seq IFCI>2 et p-valeur<0,05). **(B)** Listes complètes des gènes impliqués dans la sécrétion d'insuline et l'homéostasie du glucose. Les niveaux d'expression dans le RNA Seq dans les îlots pancréatiques chez la souris n'exprimant pas RFX6 dans les cellules beta (MUT) par rapport à la souris sauvage (WT) sont listés avec leur variation d'expression (FC) et la p-valeur associée. La position génomique du site de fixation de RFX6 identifié lors du ChIP HA-RFX6 dans les cellules Min6b1 et sa présence commune dans le ChIP RFX6 dans les Min6b1 est indiquée pour chaque gène.

Parmi les gènes mis en avant lors de cette analyse, on en compte 53 (3,4%) ayant un rôle connu dans la sécrétion d'insuline et 26 (1,7%) ayant un rôle connu dans l'homéostasie du glucose. De nombreux gènes ont également une fonction connue dans la régulation de la transcription (36,9%) ou le cycle cellulaire (16%).

Des gènes clés nécessaires à la fonction de sécrétion d'insuline par les cellules beta tels que les canaux calciques *Cacnab2* et *Cacna1c*, la *Gck* (glucokinase), ou *Abcc8*, le canal potassium sensible à l'ATP, constituent des cibles directement régulées par RFX6 dans cette première analyse. Ils sont décrits plus précisément dans le paragraphe Résultats A.3.e (Piccand et al, 2014).

De même, les gènes décrits ci-dessous impliqués spécifiquement dans la sécrétion d'insuline et l'homéostasie du glucose, sont ciblés dans les cellules Min6b1 et régulés dans les cellules beta pancréatiques par RFX6 et pourraient expliquer le rôle de RFX6 dans ces cellules.

Ffar1 (Free fatty acid receptor 1), un récepteur aux acides gras, est exprimé dans les cellules beta où il régule la sécrétion d'insuline (Ichimura et al, 2014; Itoh et al, 2003). Par étude de souris KO pour *Ffar1*, l'activation de la sécrétion d'insuline par les d'acides gras via le récepteur FFAR1 est démontrée (Steneberg et al, 2005). De même, la surexpression du gène humain *Ffar1* spécifiquement dans les cellules beta chez la souris empêche le développement d'une hyperglycémie sous régime alimentaire riche en graisse. Des agonistes de *Ffar1* sont ainsi proposés pour traiter le diabète de type 2 par leur action

positive sur l'augmentation de la sensibilité à l'insuline et le métabolisme du glucose (Tanaka et al, 2014).

GIP (Gastric inhibitory polypeptide) est l'une des hormones incrétine avec GLP-1 (glucagon like peptide1) sécrétée par l'intestin qui stimule la sécrétion d'insuline par les cellules beta en réponse au glucose via sa liaison à son récepteur GIPR (gastric inhibitory polypeptide receptor). Il a également été montré que GIP et GLP-1 préservent la masse des cellules beta par inhibition de leur apoptose et stimulation de leur prolifération, faisant de ces hormones un traitement de choix pour le diabète de type 2 (Yabe & Seino, 2011).

Irs2 (Insulin receptor substrate 2) est une molécule signal cytoplasmique des cellules beta. Les souris KO pour *Irs2* spécifiquement dans les cellules beta présentent une intolérance au glucose due à un défaut de sécrétion d'insuline et de mobilisation des ions calcium dans ce processus. Il est également observé une réduction de la masse de cellules beta et alpha chez ces souris. Cette publication suggère ainsi que la signalisation dépendante d'IRS2 dans les îlots pancréatiques est nécessaire non seulement pour le maintien de la masse des cellules alpha et beta, mais est également impliquée dans la régulation de la sécrétion d'insuline (Cantley et al, 2007). D'autre part, *Irs1* (Insulin receptor substrate 1) est impliqué dans la régulation de la durée de vie des mammifères puisque des souris KO pour ce gène vivent plus longtemps et ne présentent qu'un phénotype diabétique modéré contrairement aux souris n'exprimant pas *Irs2* (Selman et al, 2008).

Hmgn3 (High mobility group nucleosomal binding protein 3) est une protéine qui lie les nucléosomes et régule la fonction de la chromatine, elle est exprimée dans toutes les cellules endocrines du pancréas. Hmgn3 est fortement exprimée dans les cellules beta et sa perte chez la souris entraîne une sécrétion d'insuline perturbée et un phénotype diabétique. La transcription de gènes clés de la sécrétion d'insuline tels que le transporteur de glucose *Glut2* est en effet affectée (Ueda et al, 2009). Son rôle dans les cellules alpha a également été mis en avant mais son action directe sur la sécrétion de glucagon n'a pas pu être démontrée (Kurahashi et al, 2010).

L'ensemble de ces gènes à la fois régulés dans les cellules beta pancréatiques et ciblés dans le modèle cellulaire Min6b1 par RFX6 place ainsi ce facteur de transcription comme un régulateur clé de la sécrétion d'insuline et du maintien de l'homéostasie du glucose dans la cellule beta.

Par la même stratégie, la comparaison des gènes uniquement ciblés de manière commune entre les deux expériences CHIP HA-RFX6 et CHIP RFX6 (avant normalisation)

réalisées dans les Min6b1 et des gènes régulés dans les cellules beta pancréatiques met en avant 266 gènes directement régulés dont 88 sont activés et 138 réprimés par RFX6 (**Figure 35**).

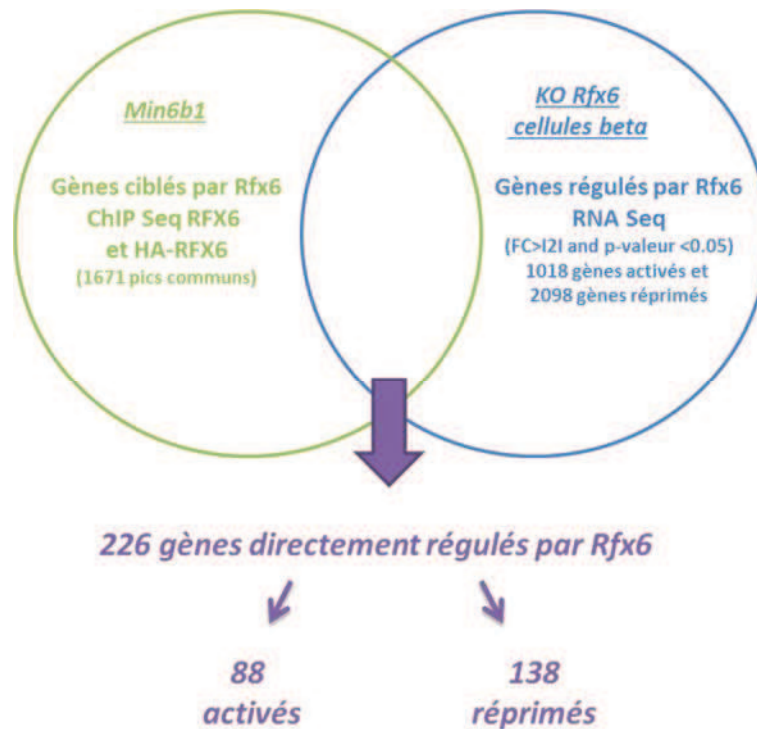


Figure 35 : Identification des gènes ciblés dans la cellule Min6b1 et régulés dans les cellules beta pancréatiques de souris par RFX6. La compilation des données de transcriptomique (gènes dont l'expression est dérégulée (RNA Seq) dans les cellules beta pancréatiques chez la souris delta beta, IFCI>2 et p-valeur<0,05) et des sites de liaisons de RFX6 (gènes associés aux 1671 pics identifiés en commun pour les ChIP Seq HA-RFX6 et RFX6 dans les cellules Min6b1 avant séquençage de l'input) permet l'identification de 226 gènes potentiellement directement régulés par RFX6 dont 88 sont activés et 138 sont réprimés.

Parmi ceux-ci, les 40 gènes dont les variations d'expression positives ou négatives les plus importantes observées par RNA Seq dans les cellules beta lorsque Rfx6 est inactivé sont présentés en **Figure 36**.

Gene name	Identifiant Ensembl	Description	Moyenne Expression CTRL	Moyenne Expression MUT	FC	p-valeur	Position Génomique du pic de ChIP Seq
Mlxipl	ENSMUSG00000005373	MLX interacting protein-like	1595.32	84.51	-17.7	5.35E-148	Intron
Prap1	ENSMUSG00000025467	proline-rich acidic protein 1	214.03	0.00	-17.6	5.18E-15	Intron et promoteur-TSS
Cdx2	ENSMUSG00000029646	caudal type homeobox 2	7.31	0.00	-7.6	4.92E-08	Intergénique
Adh4	ENSMUSG00000037797	alcohol dehydrogenase 4 (class II), pi polypeptide	24.73	0.00	-5.5	7.33E-06	Intron
Otop3	ENSMUSG00000018862	otopetrin 3	13.26	0.75	-5.2	7.40E-06	Exon
Baiap3	ENSMUSG00000047507	BAI1-associated protein 3	1366.19	245.47	-5.1	6.15E-23	Promoteur-TSS et intron
2810442I21Rik	ENSMUSG00000087060	RIKEN cDNA 2810442I21 gene	80.56	14.11	-4.9	8.10E-14	Intron
Ptprh	ENSMUSG00000035429	protein tyrosine phosphatase, receptor type, H	4.21	0.19	-4.8	2.50E-05	Promoteur-TSS
Fam179a	ENSMUSG00000045761	family with sequence similarity 179, member A	22.63	3.64	-4.7	3.60E-09	Intron
4930555F03Rik	ENSMUSG00000031559	RIKEN cDNA 4930555F03 gene	76.12	13.68	-4.4	1.09E-08	Intergénique
Ttc16	ENSMUSG00000039021	tetratricopeptide repeat domain 16	44.38	9.22	-4.3	7.14E-14	Promoteur-TSS
Gm5577	ENSMUSG00000084950	predicted gene 5577	137.69	31.20	-4.1	2.50E-19	3'UTR
Tcp11	ENSMUSG00000062859	t-complex protein 11	138.98	34.24	-3.9	1.66E-28	Promoteur-TSS
9330101J02Rik	ENSMUSG00000049571	RIKEN cDNA 9330101J02 gene	24.43	5.39	-3.9	2.73E-09	Intergénique
Acta1	ENSMUSG00000031972	actin, alpha 1, skeletal muscle	28.03	5.98	-3.8	6.65E-08	Intron
Tnr	ENSMUSG00000015829	tenascin R	531.84	134.97	-3.6	1.45E-10	Intron
Lgals4	ENSMUSG00000053964	lectin, galactose binding, soluble 4	685.46	164.95	-3.5	1.03E-06	Exon
Glb1l2	ENSMUSG00000036395	galactosidase, beta 1-like 2	1574.52	439.42	-3.5	3.93E-22	Intron
Oprm1	ENSMUSG00000000766	opioid receptor, mu 1	17.42	3.65	-3.4	6.89E-05	Intron
Ng23	ENSMUSG00000036185	Ng23 protein	19.54	2.73	-3.4	0.000644524	Intergénique

Gene name	Identifiant Ensembl	Description	Moyenne Expression CTRL	Moyenne Expression MUT	FC	p-valeur	Position Génomique du pic de ChIP Seq
Cnn1	ENSMUSG00000001349	calponin 1	1.80	78.28	16.0	1.16E-17	Intergénique
Gzmk	ENSMUSG00000042385	granzyme K	3.06	31.85	6.1	7.84E-09	Intergénique
Esr1	ENSMUSG00000019768	estrogen receptor 1 (alpha)	56.31	554.35	6.1	5.24E-09	Intron
Mybpc2	ENSMUSG00000038670	myosin binding protein C, fast-type	0.34	7.32	6.0	1.08E-06	Intron
Ptgis	ENSMUSG00000017969	prostaglandin I2 (prostacyclin) synthase	20.05	404.22	6.0	1.36E-06	Intron
Lgals3	ENSMUSG00000050335	lectin, galactose binding, soluble 3	93.63	938.65	4.9	5.60E-06	Intergénique
Madcam1	ENSMUSG00000020310	mucoasal vascular addressin cell adhesion molecule 1	3.75	54.90	4.9	1.91E-05	Intergénique et promoteur-TSS
Mfap4	ENSMUSG00000042436	microfibrillar-associated protein 4	3.73	32.68	4.8	2.99E-06	Promoteur-TSS
Pcdh7	ENSMUSG00000029108	protocadherin 7	8.24	68.26	4.8	3.02E-06	Intron
Tpx2	ENSMUSG00000027469	TPX2, microtubule-associated protein homolog (Xenopus laevis)	22.13	134.82	4.7	3.64E-09	Intron
Cldn10	ENSMUSG00000022132	claudin 10	34.34	257.21	4.5	5.22E-06	Intron
Rsad2	ENSMUSG00000020641	radical S-adenosyl methionine domain containing 2	20.87	102.26	4.5	1.31E-17	Exon
Ckap2	ENSMUSG00000037725	cytoskeleton associated protein 2	10.77	61.51	4.4	5.67E-08	Intron
Chi3l1	ENSMUSG00000064246	chitinase 3-like 1	19.77	186.26	4.3	5.22E-05	Intron
Antxr1	ENSMUSG00000033420	anthrax toxin receptor 1	12.99	68.95	4.3	1.23E-08	Intron
Chl1	ENSMUSG00000030077	cell adhesion molecule with homology to L1CAM	16.59	96.29	4.3	5.57E-07	Intron
Ifitm3	ENSMUSG00000025492	interferon induced transmembrane protein 3	166.87	1058.44	4.2	5.39E-06	Intergénique
Anxa2	ENSMUSG00000032231	annexin A2	532.37	3431.67	4.2	7.86E-06	Intron
Cenpf	ENSMUSG00000026605	centromere protein F	30.21	147.34	4.0	2.08E-08	Promoteur-TSS
Ncaph	ENSMUSG00000034906	non-SMC condensin I complex, subunit H	14.42	68.72	3.9	4.07E-08	Promoteur-TSS

Figure 36 : Gènes ciblés dans la cellule Min6b1 et régulés dans les cellules beta pancréatiques par RFX6. Les 20 gènes les plus surexprimés (en haut) ou sous-exprimés (en bas) (RNA Seq) dans les cellules beta pancréatiques chez la souris n'exprimant pas RFX6 dans les cellules beta (MUT) par rapport à la souris sauvage (WT) sont listés avec leur variation d'expression

(FC) et la p-valeur associée. La position du site de fixation commun de RFX6 identifié lors du ChIP HA-RFX6 et RFX6 dans les cellules Min6b1 est indiquée pour chaque gène.

Enfin, un classement des gènes ciblés dans la cellule Min6b1 et régulés dans les cellules beta pancréatiques par RFX6 est effectué par voies biologiques (Gene Ontology) (**Figure 37**). Les 10 gènes les plus dérégulés dans chaque liste sont présentés en **Figure 37A** et la totalité des gènes impliqués dans la sécrétion d'insuline et l'homéostasie du glucose sont présentés en **Figure 37B**.

A

Voies biologiques	% des gènes ciblés et régulés par RFX6	Gènes ciblés et régulés par RFX6
Homéostasie du glucose	3.1	<i>Bad</i> ; <i>Foxa3</i> ; <i>Vgf</i> <i>Mlxipl</i> ; <i>Cacna1c</i> ; <i>Rph3al</i>
Sécrétion d'insuline	6.6	<i>Bad</i> ; <i>Ppp3ca</i> ; <i>Pim3</i> ; <i>Nr0b2</i> ; <i>Vgf</i> ; <i>Kcnq1</i> ; <i>Blk</i> ; <i>Arrb1</i> <i>Gipr</i> ; <i>Cacna1c</i> ; <i>Rapgef4</i> ; <i>Sirt4</i> ; <i>Abcc8</i> ; <i>Snap25</i>
Régulation de la transcription	49.1	<i>Smo</i> ; <i>Tlx2</i> ; <i>Atoh8</i> ; <i>Sp110</i> ; <i>Litaf</i> ; <i>Scx</i> ; <i>Spib</i> ; <i>Fgfr2</i> ; <i>Ckap2</i> ; <i>Esr1</i> <i>Mlxipl</i> ; <i>Cdx2</i> ; <i>Zfp174</i> ; <i>Zfp541</i> ; <i>Fosb</i> ; <i>Tcea2</i> ; <i>Bcl6b</i> ; <i>Hdac6</i> ; <i>Kdm2b</i> ; <i>Myt1</i>
Cycle cellulaire	22.6	<i>6030408B16Rik</i> ; <i>Esco2</i> ; <i>Ccnf</i> ; <i>Dlgap5</i> ; <i>Cks2</i> ; <i>Cdc20</i> ; <i>Nek2</i> ; <i>Ncaph</i> ; <i>Ckap2</i> ; <i>Tpx2</i> <i>Mlxipl</i> ; <i>Crocc</i> ; <i>BC024139</i> ; <i>Sycp2</i> ; <i>Sept3</i> ; <i>Cntrob</i> ; <i>Mapre3</i> ; <i>Plk5</i> ; <i>Spice1</i> ; <i>4922501C03Rik</i>

B

SECRETION D'INSULINE							
Gène	Identifiant Ensembl	Description	Moyenne Expression CTRL	Moyenne Expression MUT	FC	p-valeur	Position génomique du pic de ChIP Seq anti-HA-RFX6
<i>Gipr</i>	ENSMUSG00000030406	gastric inhibitory polypeptide receptor	1531.66	498.03	-3.0	8.22E-20	Promoteur-TSS et intron
<i>Cacna1c</i>	ENSMUSG00000051331	calcium channel, voltage-dependent, L type, alpha 1C subunit	337.79	148.65	-2.2	1.54E-05	Intron
<i>Rapgef4</i>	ENSMUSG00000049044	Rap guanine nucleotide exchange factor (GEF) 4	6952.18	4735.31	-1.5	9.86E-05	Intron
<i>Sirt4</i>	ENSMUSG00000029524	sirtuin 4 (silent mating type information regulation 2 homolog) 4 (S. cerevisiae)	374.51	261.68	-1.4	3.75E-04	Intergénique
<i>Abcc8</i>	ENSMUSG00000040136	ATP-binding cassette, sub-family C (CFTR/MRP), member 8	6540.14	4579.93	-1.4	1.31E-02	Intron
<i>Snap25</i>	ENSMUSG00000027273	synaptosomal-associated protein 25	1976.84	1485.16	-1.3	4.58E-02	Intron
<i>Bad</i>	ENSMUSG00000024959	BCL2-associated agonist of cell death	487.89	635.72	1.3	3.09E-04	Promoteur-TSS
<i>Ppp3ca</i>	ENSMUSG00000028161	protein phosphatase 3, catalytic subunit, alpha isoform	1499.58	2172.78	1.4	2.11E-04	Intron
<i>Pim3</i>	ENSMUSG00000035828	proviral integration site 3	610.11	1082.10	1.7	1.29E-02	Intron
<i>Nr0b2</i>	ENSMUSG00000037583	nuclear receptor subfamily 0, group B, member 2	18.01	47.49	2.2	5.47E-03	Exon
<i>Vgf</i>	ENSMUSG00000037428	VGF nerve growth factor inducible	92.06	228.37	2.4	8.02E-09	Intergénique et promoteur-TSS
<i>Kcnq1</i>	ENSMUSG00000009545	potassium voltage-gated channel, subfamily Q, member 1	69.65	180.15	2.4	5.38E-06	Intergénique
<i>Blk</i>	ENSMUSG00000014453	B lymphoid kinase	1.54	10.15	2.5	1.82E-02	Promoteur-TSS
<i>Arrb1</i>	ENSMUSG00000018909	arrestin, beta 1	753.76	1948.76	2.5	1.35E-08	Intron

HOMEOSTASIE DU GLUCOSE							
Gène	Identifiant Ensembl	Description	Moyenne Expression CTRL	Moyenne Expression MUT	FC	p-valeur	Position génomique du pic de ChIP Seq anti-HA-RFX6
<i>Mlxipl</i>	ENSMUSG00000005373	MLX interacting protein-like	1595.32	84.51	-17.7	5.35E-148	Intron
<i>Cacna1c</i>	ENSMUSG00000051331	calcium channel, voltage-dependent, L type, alpha 1C subunit	337.79	148.65	-2.2	1.54E-05	Intron
<i>Rph3al</i>	ENSMUSG00000020847	rabphilin 3A-like (without C2 domains)	2893.11	2180.00	-1.3	8.90E-03	Intergénique
<i>Bad</i>	ENSMUSG00000024959	BCL2-associated agonist of cell death	487.89	635.72	1.3	3.09E-04	Promoteur-TSS
<i>Foxa3</i>	ENSMUSG00000040891	forkhead box A3	183.05	253.14	1.4	1.47E-05	Promoteur-TSS
<i>Vgf</i>	ENSMUSG00000037428	VGF nerve growth factor inducible	92.06	228.37	2.4	8.02E-09	Intergénique et promoteur-TSS

Figure 37 : Classement par voies biologiques des gènes ciblés dans la cellule Min6b1 et régulés dans les cellules beta pancréatiques par RFX6. Les gènes ciblés par RFX6 (cibles communes ChIP HA-RFX6 et RFX6 dans les cellules Min6b1) et surexprimés (en bleu) ou sous-exprimés (en orange) en absence de RFX6 (RNA Seq) dans les cellules beta pancréatiques sont listés et classés par voies biologiques (Gene Ontology). **(A)** Le pourcentage du nombre total de gènes ciblés et régulés par RFX6 (226 gènes au total) et les 10 gènes les plus dérégulés dans chaque liste sont présentés. Les gènes en gras présentent une variation d'expression significative supérieure à deux (RNA Seq IFCI>2 et p-valeur<0,05). **(B)** Listes complètes des gènes impliqués dans la sécrétion d'insuline et l'homéostasie du glucose. Les niveaux d'expression dans le RNA Seq dans les cellules beta pancréatiques chez la souris n'exprimant pas RFX6 dans les cellules beta (MUT) par rapport à la souris sauvage (WT) sont listés avec leur variation d'expression (FC) et la p-valeur associée. La position génomique du site de fixation commun de RFX6 identifié lors du ChIP HA-RFX6 et RFX6 dans les cellules Min6b1 est indiquée pour chaque gène.

Parmi les gènes mis en avant lors de cette analyse, on en compte 15 (6,6%) ayant un rôle connu dans la sécrétion d'insuline et 7 (3,1%) ayant un rôle connu dans l'homéostasie du glucose. De nombreux gènes ont également une fonction connue dans la régulation de la transcription (49,1%) ou le cycle cellulaire (22,6%). On retrouve notamment certains gènes exposés précédemment tels que *Cacna1c*, *Abcc8* ou *GipR*, confirmant le rôle fondamental de *Rfx6* dans la sécrétion d'insuline et l'homéostasie du glucose dans la lignée cellulaire *Min6b1*.

A.3.d. Analyse d'occurrence et de co-occurrence de motifs dans les cibles de RFX6 dans les cellules *Min6b1*

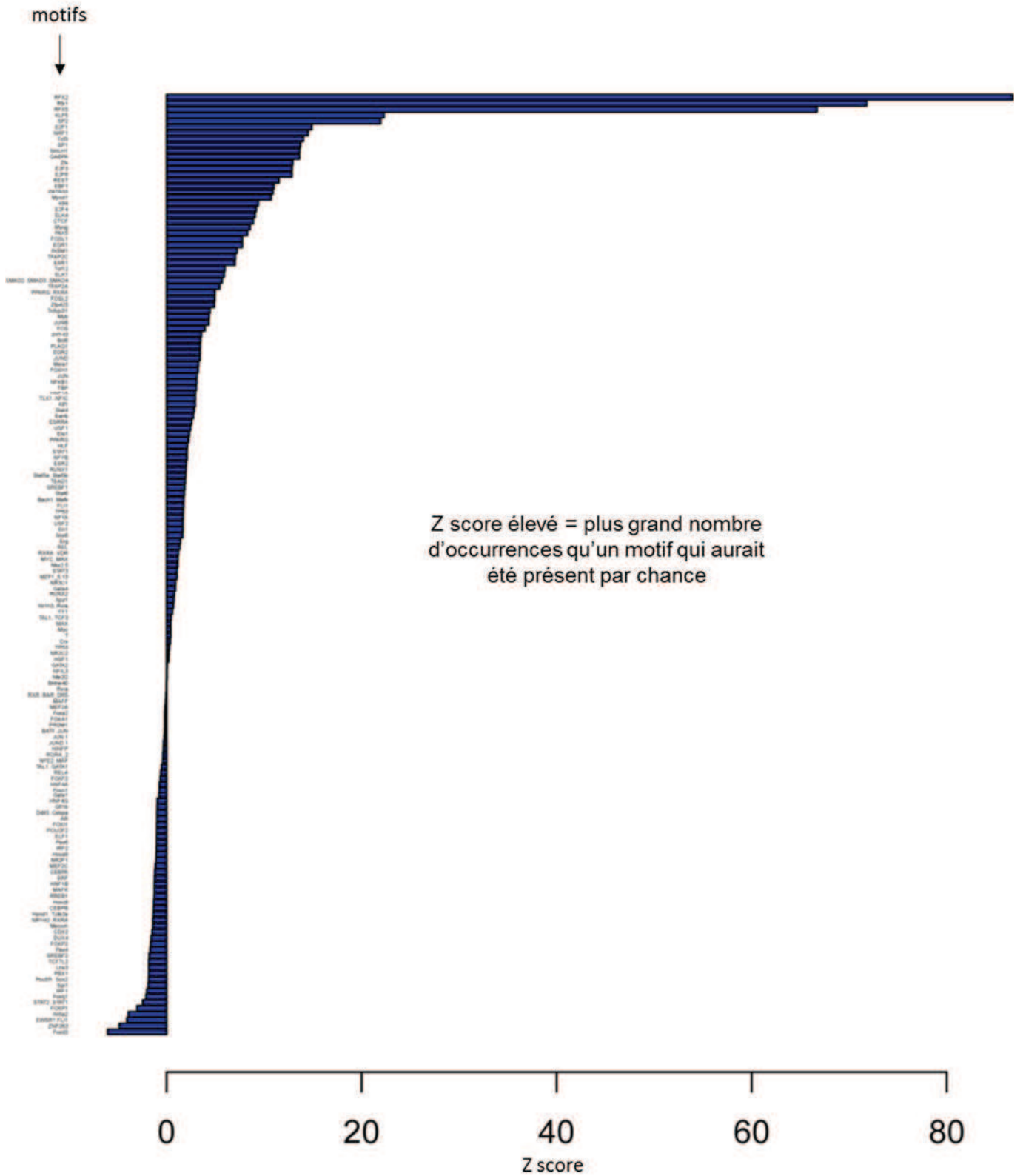
Enfin, une recherche de motifs connus (base de données Jaspar 2014) dans les séquences ciblées par RFX6 dans les cellules *Min6b1* a été mise en place. Pour cela, ce sont les séquences sous les 3300 pics de ChIP Seq anti-HA-RFX6 (500 pb centrées sous les sommets des 3300 pics correspondants aux 1540 gènes directement régulés dans les cellules beta pancréatiques (**Figure 32**), qui ont été utilisées pour une analyse d'occurrence de motif. Celle-ci a été mise en place (**paragraphe Matériels et méthodes C.4**) en association avec Stéphanie Legras, doctorante, de la plateforme de bio-informatique et séquençage de l'IGBMC. Des protéines ou cofacteurs agissant de concert avec RFX6 pourront ainsi être identifiés, s'ils existent. Les résultats obtenus sont présentés en **Figure 38**. La présence ou occurrence des motifs est dénombrée et un Z score représentant la présence du motif plus ou moins élevée par rapport à ce qui aurait été obtenu par chance est calculé (**Figure 38A**). Les motifs occurrents de manière significative au niveau des cibles de RFX6 dans les îlots pancréatiques sont listés en **Figure 38B** et un histogramme est présenté pour les 10 motifs les plus occurrents (**Figure 38C**).

Comme attendu et suggéré par l'analyse de motif *de novo* qui avait été réalisée sur le même jeu de données (**paragraphe Résultats A.3.b Figure 30**), les motifs les plus occurrents sont les motifs de type « RFX ». On retrouve également le motif « NRF1 » dans cette analyse d'occurrence. D'autres motifs de type « SP2 », « SP1 » ou encore « E2F1 » sont également enrichis dans ces séquences.

Le facteur de transcription **SP2** (specificity protein 2) est essentiel au développement des souris et poissons zèbre (Baur et al, 2010; Xie et al, 2010) et a également été identifié comme régulateur clé de la neurogenèse au niveau embryonnaire et postnatal chez la souris (Liang et al, 2013). **SP1** (specificity protein 1) qui fait partie de la même famille régule

l'expression de gènes impliqués dans la prolifération cellulaire et l'apoptose et joue un rôle dans le développement de cancers chez l'homme (Vizcaino et al, 2015). Une occurrence légèrement moins importante des motifs de type « E2F1 » est également observée. Ce facteur de transcription est directement impliqué dans la fonction et la prolifération de la cellule beta sécrétrice d'insuline (Annicotte et al, 2009; Fajas et al, 2010). Lorsqu'**E2F1** est surexprimé, la prolifération des cellules beta est stimulée et leur contenu en insuline augmente (Grouwels et al, 2010). Les souris KO pour E2F1 présentent une homéostasie au glucose perturbée ainsi qu'un pancréas de taille réduite mais ne sont pas diabétiques (Fajas et al, 2004). Enfin, E2F1 régule directement l'expression de la sous-unité Kir6.2 du canal potassium ATP dépendant, impliqué dans la synthèse d'insuline, mais aussi de Ngn3 de manière Cdk4 dépendante (Kim & Rane, 2011).

A



B

Motifs	Zscore	Nombre d'occurrence
RFX2	86.71	331
Rfx1	71.78	277
RFX5	66.69	164
KLF5	22.30	132
SP2	21.95	232
E2F1	14.89	34
NRF1	14.50	46
Tcf3	14.03	62
SP1	13.73	168
NHLH1	13.64	26
GABPA	13.62	17
Zfx	12.93	40
E2F3	12.86	54
E2F6	12.85	33
REST	11.52	45
EBF1	11.01	45
ZBTB33	10.91	8
Myod1	10.71	43
Klf4	9.39	45
E2F4	9.18	23
ELK4	9.07	11
CTCF	8.89	63
Myog	8.57	27
PAX5	8.29	15
FOSL1	7.78	20
EGR1	7.76	160
INSM1	7.21	17
TFAP2C	7.05	39
ESR1	7.01	34

Motifs	Zscore	Nombre d'occurrence
Tcf12	5.98	26
ELK1	5.90	6
SMAD2..SMAD3..SMAD4	5.69	25
TFAP2A	5.43	30
PPARG..RXRA	4.95	35
FOSL2	4.93	14
Zfp423	4.89	17
Tcfcp2l1	4.45	16
Myb	4.36	13
JUNB	4.33	14
FOS	3.93	11
znf143	3.57	22
Bcl6	3.49	20
PLAG1	3.45	18
EGR2	3.44	78
JUND	3.41	10
Meis1	3.26	20
FOXH1	3.23	5
JUN	3.09	21
NFKB1	3.08	11
TBP	3.05	7
HNF1A	2.96	7
TLX1..NFIC	2.94	10
Klf1	2.93	25
Stat4	2.82	15
Esrrb	2.75	16
ESRRA	2.56	17
USF1	2.45	8
Ets1	2.34	25

C

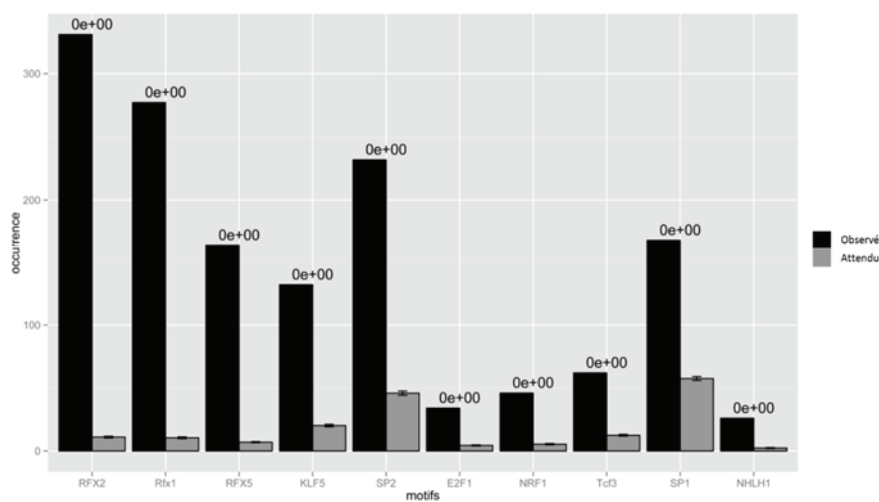


Figure 38 : Analyse d'occurrence de motifs dans les 1540 gènes ciblés dans les Min6b1 et régulés dans les cellules beta pancréatiques par RFX6. Les 500pb centrées autour du sommet

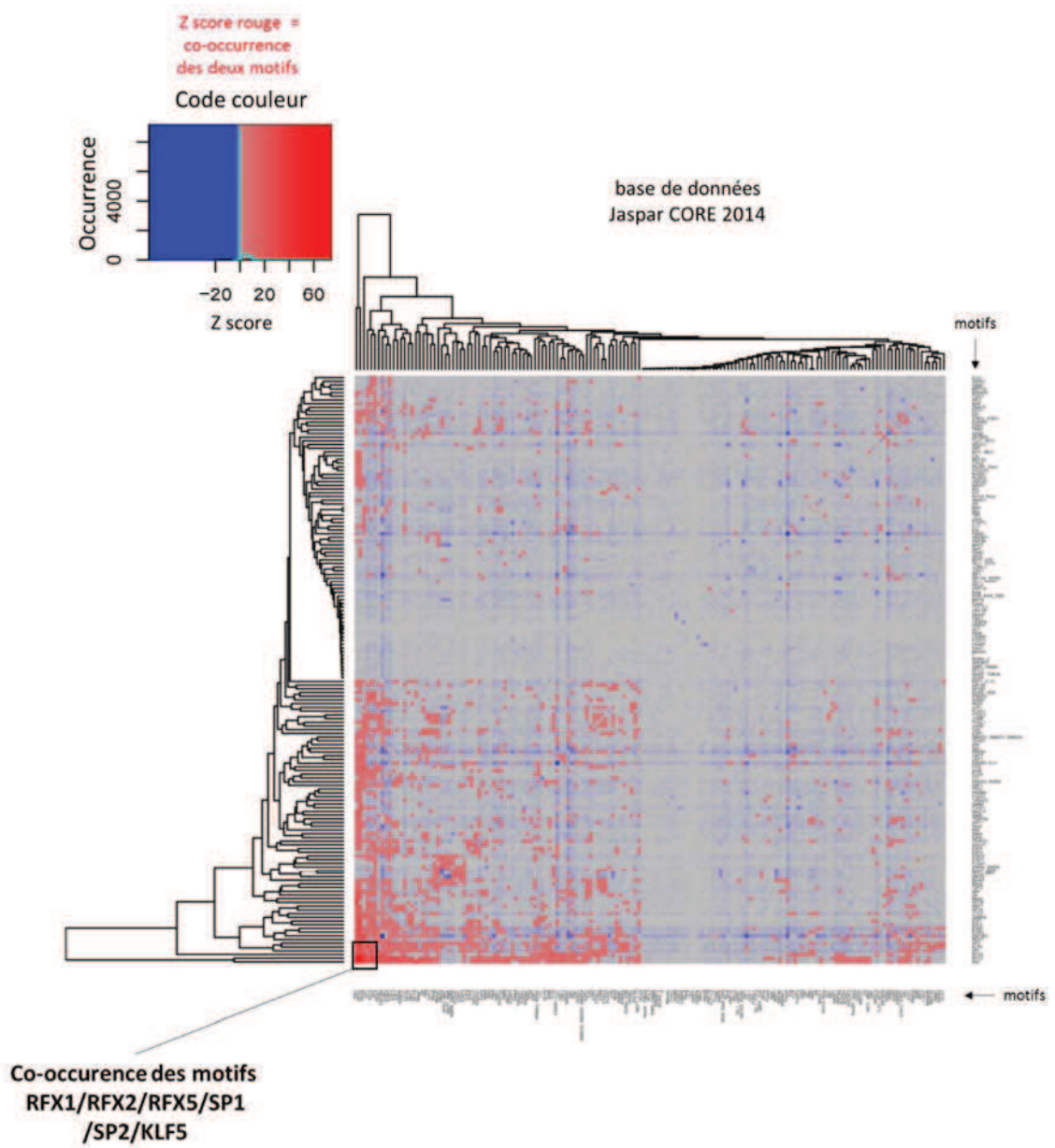
des pics dans les séquences sous les 3300 pics de ChIP Seq HA-RFX6 correspondants aux 1540 gènes affectés sont analysées avec la base de données Jaspar 2014. **(A)** Histogramme où un Z score élevé signifie que l'occurrence du motif est plus fréquente que ce qui aurait été obtenu par chance, inversement un Z score peu élevé signifie que l'occurrence du motif est moins fréquente que ce qui aurait été obtenu par chance. **(B)** Les motifs sont classés par ordre décroissant de Z score et sont associés à leur nombre d'occurrence. Les motifs dont l'occurrence est significative sont présentés. **(C)** Histogramme d'occurrence des motifs. Le nombre d'occurrences attendues et observées est représenté pour les 10 motifs les plus fréquents.

D'autre part, une analyse de co-occurrence de motifs a été réalisée et consiste à vérifier combien de fois deux motifs différents sont présents dans la même séquence. Les motifs dont la co-occurrence est la plus élevée avec les motifs de type « RFX » sont listés en **Figure 39A**. On retrouve dans cette liste les motifs les plus fréquents identifiés dans l'analyse précédente. Le nombre de co-occurrence est particulièrement élevé pour les motifs « RFX » entre eux, et correspond à des « homotypic clusters » (Ezer et al, 2014) qui permettent une régulation transcriptionnelle fine des gènes. Enfin, l'analyse de co-occurrence est représentée par une heatmap où un Z score élevé représenté en rouge signifie que la co-occurrence des deux motifs apparaît plus fréquemment que ce qui aurait été obtenu par chance (**Figure 39B**). Un zoom de cette représentation comprenant la co-occurrence avec les motifs RFX (encadrés en noir) est présentée en **Figure 39C**.

A

	SP1	SP2	NRF1	KLF5	E2F1	RFX2	RFX5	RFX1
RFX1	10	13	4	7	3	182	53	76
RFX2	13	20	6	10	5	98	57	182
RFX5	6	3	1	4	0	57	8	53

B



C

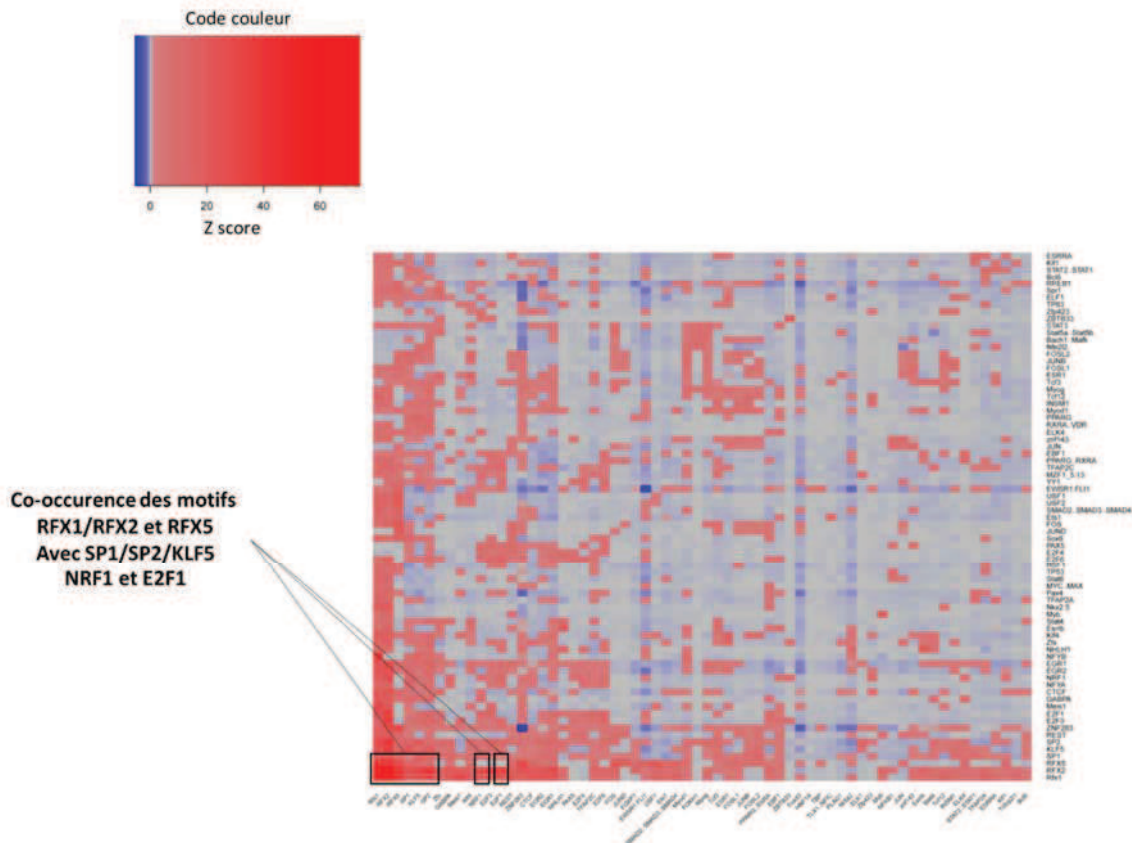


Figure 39 : Analyse de co-occurrence de motifs dans les 1540 gènes ciblés dans les Min6b1 et régulés dans les cellules beta pancréatiques par RFX6. Les séquences sous les 3300 pics de ChIP Seq HA-RFX6 correspondants aux 1540 gènes affectés sont analysées avec la base de données de motifs Jaspar 2014. **(A)** Les motifs dont la co-occurrence est la plus élevée avec les motifs de type RFX sont présentés. **(B)** Heatmap où un Z score élevé représenté en rouge signifie que la co-occurrence des deux motifs apparaît plus fréquemment que ce qui aurait été obtenu par chance, inversement un Z score peu élevé représenté en bleu signifie que la co-occurrence des deux motifs apparaît moins fréquemment que ce qui aurait été obtenu par chance. **(C)** Zoom sur la région de la heatmap comprenant les motifs RFX.

Ces analyses d'occurrence et de co-occurrence ont ainsi permis de confirmer la présence enrichie de motifs « RFX » au niveau des sites de fixation de RFX6 dans les cellules Min6b1 et de mettre en avant d'éventuels cofacteurs tels que NRF1 et les autres membres de la famille RFX entre autres. Cependant, d'autres analyses pourraient encore être développées telles que la recherche de co-régulation de certaines cibles par test de transactivation par exemple ou une recherche des voies biologiques liées aux motifs

hautement occurs permettant l'identification des fonctions encore non dévoilées de RFX6 dans les cellules Min6b1.

Enfin, une première idée du rôle de RFX6 dans la régulation de la sécrétion d'insuline et donc de l'homéostasie du glucose dans la cellule beta a été dévoilée par l'identification des gènes ciblés dans les cellules Min6b1 et régulés dans les cellules beta pancréatiques. Ces premiers résultats ont mené à la publication d'un manuscrit décrit dans le paragraphe suivant.

A.3.e. Rfx6 maintains the functional identity of adult pancreatic beta cells / Piccand et al. 2014, Cell Reports

Dans la publication de 2014 dans Cell Reports présentée dans ce paragraphe, notre groupe montre que RFX6 est essentielle au maintien des principales caractéristiques des cellules beta fonctionnellement matures chez la souris (Piccand et al, 2014). L'inactivation conditionnelle de Rfx6 spécifiquement dans les cellules beta chez l'adulte mène à l'intolérance au glucose, mais à la différence de l'embryon, l'insuline est encore produite dans ces cellules dépourvues de RFX6. Toutefois, la sécrétion d'insuline est altérée.

Pour déterminer quels gènes sont impliqués dans ce phénotype, l'étude des gènes régulés et ciblés par RFX6 a été menée. Pour cela, des expériences de RNA Seq dans la souris Rfx6 fl/fl ; Ins1-CrERT où Rfx6 est spécifiquement délété dans les cellules beta 5 jours après injection de tamoxifène ont été réalisées d'une part pour identifier tout le répertoire des gènes régulés par ce facteur de transcription dans les cellules beta pancréatiques.

D'autre part, j'ai pu identifier l'ensemble des gènes cibles de RFX6 par ChIP Seq HA-RFX6 dans les cellules Min6b1 transfectées avec le plasmide pCMVTag2A-3HA-RFX6-2A-mCherry dont les résultats détaillés sont décrits dans le paragraphe Résultats A.3.b **Figure 30 et Figure 32 à Figure 34.**

C'est par la suite, la combinaison des gènes régulés dans les cellules beta pancréatiques et ciblés dans les cellules Min6b1 par RFX6, qui a permis l'identification des gènes potentiellement directement régulés par RFX6.

Cette étude révèle ainsi que la sécrétion d'insuline déficiente chez la souris n'exprimant pas Rfx6 résulte d'une réduction combinée des expressions de gènes nécessaires au mécanisme de sécrétion d'insuline dont la glucokinase, la sous-unité Abcc8 du canal potassium sensible à l'ATP ainsi que plusieurs canaux calciques voltage dépendant qui sont tous des cibles directes de RFX6 identifiés par ChIP Seq et confirmés par ChIP qPCR.

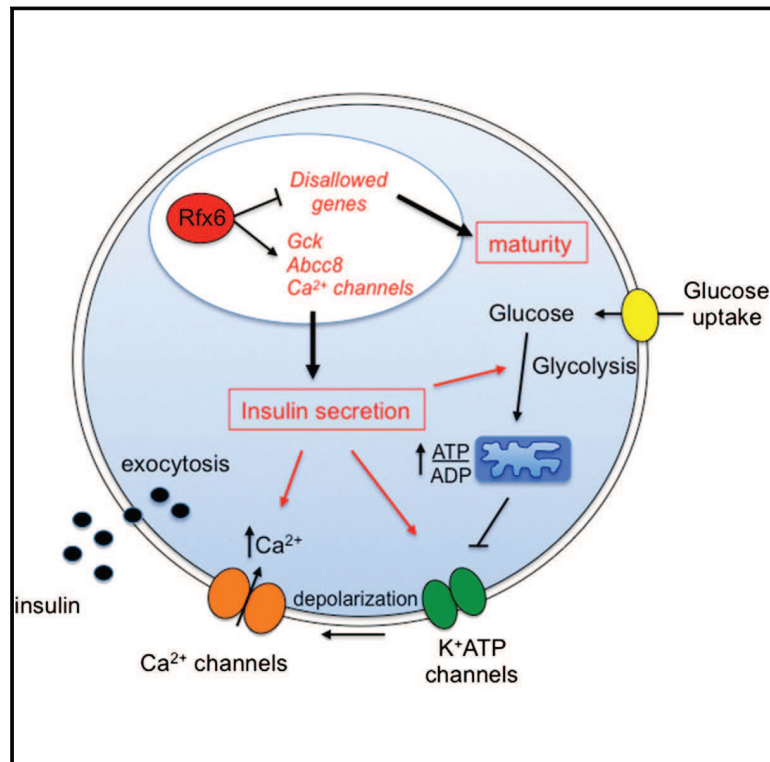
En outre, nous avons découvert que RFX6 contribue à la répression de la grande majorité des gènes dit «disallowed» (Pullen et al, 2010; Quintens et al, 2008; Thorrez et al, 2011), habituellement spécifiquement réprimés dans les cellules beta adultes, et donc au maintien de l'identité de la cellule beta. En effet, ce sont plus de 58% de la totalité de ces gènes qui sont ciblés par RFX6 dans les cellules Min6b1 dans mes études de ChIP Seq.

Mes analyses de ChIP Seq ont ainsi permis d'élucider le rôle fondamental de RFX6 dans la fonction et le maintien de l'identité de la cellule beta sécrétrice d'insuline.

Cell Reports

Rfx6 Maintains the Functional Identity of Adult Pancreatic β Cells

Graphical Abstract



Authors

Julie Piccand, Perrine Strasser, ...,
Guy A. Rutter, Gérard Gradwohl

Correspondence

gradwohl@igbmc.fr

In Brief

Rfx6 transcription factor is critical for the development of pancreatic β cells, and Rfx6 mutations cause neonatal diabetes in humans. Using adult β cell deletion of Rfx6, Piccand et al. show that Rfx6 controls glucose homeostasis and maintains the differentiated state by activation of the insulin secretory pathway and repression of disallowed genes, a group of genes selectively repressed in mature β cells.

Highlights

- Inactivation of Rfx6 in adult β cells leads to impaired insulin secretion
- Glucose-induced Ca^{2+} influx is reduced in β cells lacking Rfx6
- Rfx6 activates components of the insulin secretion pathway
- Rfx6 maintains β cell maturity by repressing “disallowed” genes

Accession Numbers

GSE59622
GSE62844



Rfx6 Maintains the Functional Identity of Adult Pancreatic β Cells

Julie Piccand,¹ Perrine Strasser,¹ David J. Hodson,² Aline Meunier,¹ Tao Ye,¹ Céline Keime,¹ Marie-Christine Birling,³ Guy A. Rutter,² and Gérard Gradwohl^{1,*}

¹Institut de Génétique et de Biologie Moléculaire et Cellulaire, Institut National de la Santé et de la Recherche Médicale U964, Centre National de Recherche Scientifique UMR7104, Université de Strasbourg, Illkirch 67404, France

²Section of Cell Biology, Division of Diabetes, Endocrinology and Metabolism, Department of Medicine, Imperial College London, Hammersmith Hospital, du Cane Road, London W12 0NN, UK

³Institut Clinique de la Souris-ICS-MCI, PHENOMIN, Illkirch 67404, France

*Correspondence: gradwohl@igbmc.fr

<http://dx.doi.org/10.1016/j.celrep.2014.11.033>

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/3.0/>).

SUMMARY

Increasing evidence suggests that loss of β cell characteristics may cause insulin secretory deficiency in diabetes, but the underlying mechanisms remain unclear. Here, we show that Rfx6, whose mutation leads to neonatal diabetes in humans, is essential to maintain key features of functionally mature β cells in mice. Rfx6 loss in adult β cells leads to glucose intolerance, impaired β cell glucose sensing, and defective insulin secretion. This is associated with reduced expression of core components of the insulin secretion pathway, including glucokinase, the Abcc8/SUR1 subunit of K_{ATP} channels and voltage-gated Ca^{2+} channels, which are direct targets of Rfx6. Moreover, Rfx6 contributes to the silencing of the vast majority of “disallowed” genes, a group usually specifically repressed in adult β cells, and thus to the maintenance of β cell maturity. These findings raise the possibility that changes in Rfx6 expression or activity may contribute to β cell failure in humans.

INTRODUCTION

The mammalian pancreas comprises an exocrine compartment, secreting digestive enzymes into the intestine, and an endocrine compartment, secreting hormones in the bloodstream. Pancreatic endocrine cells are grouped in small clusters of cells, the islets of Langerhans, containing different cell types secreting distinct hormones. Islet cells include β cells, which secrete insulin, the hormone stimulating glucose uptake in peripheral tissues. Briefly, glucose enters β cells by facilitated diffusion and, after phosphorylation by glucokinase (lynedjian, 1993), is metabolized by aerobic glycolysis (Sekine et al., 1994), producing metabolic signals such as a rise in ATP/ADP concentration (Tarasov et al., 2012). The latter in turn closes ATP-sensitive K^+ channels, causing membrane depolarization and the subsequent opening of voltage-gated Ca^{2+} channels (Yang and Berggren, 2006).

Ca^{2+} influx then stimulates the exocytosis of insulin granules (Rutter, 2004).

Diabetes is a chronic metabolic disease characterized by hyperglycemia due to defective insulin secretion, insulin action, or both. β cells are lacking in type 1 diabetes, while in type 2 diabetic patients, β cells cannot compensate for the increased insulin demand due to their reduced capacity to secrete insulin in response to high blood glucose. Alterations in both β cell mass (Butler et al., 2003; Marselli et al., 2013; Rahier et al., 2008) and function (Rosengren et al., 2012) are likely to contribute to the overall secretory deficiency observed in type 2 diabetes (Rutter, 2014). Recently, it has been proposed that β cell dysfunction in type 2 diabetes might also result from a mechanism of dedifferentiation, which would compromise β cell function (Talchai et al., 2012) and contribute to the development of the disease together with cell death and decreased β cell mass. This hypothesis, which builds on earlier findings (Jonas et al., 1999), has been based on the observation that ablation of FoxO1 transcription factor in adult β cells in mice caused hyperglycemia with a concomitant reversion of β cells to a progenitor- or α -like state. Along the same lines, additional loss-of-function studies in adult β cells revealed that NeuroD1 (Gu et al., 2010), Nkx6.1 (Taylor et al., 2013), or Pdx1 (Gao et al., 2014) transcription factors are important to maintain the maturity and differentiated state as well as the insulin-secretive function of β cells. Thus, it appears that the loss of key β cell transcription factors results in the loss of both β cell identity and function.

Rfx6 is a winged-helix transcription factor that has been shown to be essential for islet cell development in zebrafish (Soyer et al., 2010), *Xenopus* (Pearl et al., 2011), mice (Smith et al., 2010), and humans (Concepcion et al., 2014; Pearl et al., 2011; Smith et al., 2010; Spiegel et al., 2011). Rfx6 null mice lack all endocrine cells (excepting PP cells), including β cells, and die shortly after birth. It was thus concluded that Rfx6 is necessary for insulin production during embryogenesis (Smith et al., 2010). In humans, mutations in RFX6 have been reported to be the cause of the Mitchell-Riley syndrome, an autosomal-recessive syndrome of neonatal diabetes and small bowel atresia, often associated with intestinal malabsorption (Concepcion et al., 2014; Smith et al., 2010; Spiegel et al., 2011). Clusters



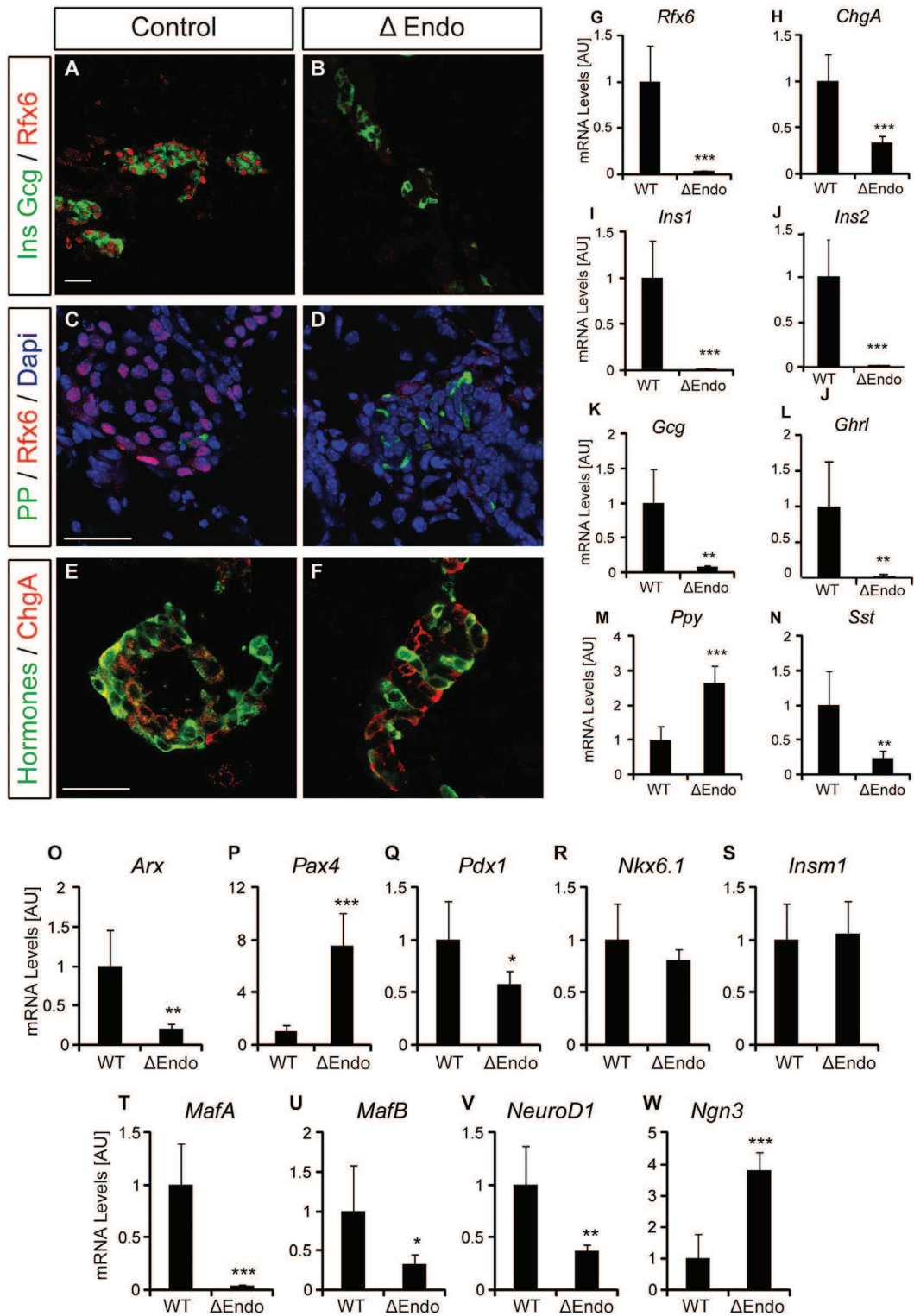


Figure 1. Deletion of *Rfx6* Downstream of *Ngn3* Results in the Loss of Insulin-, Glucagon-, Somatostatin-, and Ghrelin-Producing Islet Cells in Newborn Mice

(A–F) Immunofluorescence experiments on pancreata from controls and *Rfx6* ^{Δ Endo} pups at postnatal day 0 (P0). Staining for insulin and glucagon (green) and *Rfx6* (red) showing efficient deletion of *Rfx6* and strong reduction of insulin- and glucagon-expressing cells in *Rfx6* ^{Δ Endo} mutants (A and B). (B) and (F) show very rare

(legend continued on next page)

of chromogranin A-positive hormone-negative cells have been reported in the pancreas of several patients, suggesting a critical role for RFX6 in the formation of islet and β cells in humans. The complex spatiotemporal expression pattern of Rfx6 in mice, namely its broad expression very early in the gut and pancreas endoderm and then its restriction to developing endocrine cells in the embryo and its maintenance in adult islets, including β cells (Smith et al., 2010; Soyer et al., 2010), suggests multiple functions at different stages and in different organs. Thus, the phenotype of Rfx6 null mice might result from multiple deficiencies during development. Importantly, the postnatal lethality of Rfx6^{-/-} pups precluded the study of Rfx6 function in adult β cells. Therefore, to decipher the multiple functions of Rfx6, we generated a floxed allele. Conditional inactivation of Rfx6 in developing endocrine cells phenocopies the null phenotype, demonstrating that Rfx6 controls islet development downstream of the proendocrine transcription factor Ngn3. Conditional inactivation of Rfx6 in adult β cells led to insulin secretion deficiency and glucose intolerance, although insulin was still produced. Removal of Rfx6 perturbs key molecular traits of functional β cells with the reduction in expression of glucokinase, the ATP-sensitive K⁺ channel *Abcc8/SUR1*, as well as eight voltage-gated Ca²⁺ channel genes, some of which are direct targets of Rfx6. Collectively, lowered levels of these core components of the insulin secretion pathway seem likely to underlie defective insulin secretion. Unexpectedly, we found that inactivation of Rfx6 in adult β cells caused the re-expression of the “disallowed” genes (Pullen et al., 2010; Quintens et al., 2008; Thorrez et al., 2011), a set of genes expressed normally in most (if not all) mammalian tissue but selectively and highly repressed in mouse mature β cells. This finding reveals a common repression mechanism of disallowed gene expression in β cells involving a unique transcription factor. Taken together, our data show that Rfx6 inactivation in β cells causes decreased expression of β cell-specific genes combined with the upregulation of disallowed genes, a feature of immature β cells, demonstrating that Rfx6 can act as an activator or repressor of transcription. Thus, the present study demonstrates that Rfx6 is essential for the maintenance of the differentiated state and functional identity of adult β cells. Therefore, Rfx6 might be an interesting target should “ β cell-identity drugs” be developed in the future as a novel therapeutic strategy in type 2 diabetes.

RESULTS

Rfx6 Is Essential for β Cell Differentiation Downstream of the Proendocrine Transcription Factor Ngn3

To decipher the different roles of Rfx6, we generated a conditional knockout mouse by flanking exon 3, which encodes a part of the DNA binding domain, by two *loxP* sites, generating a null allele (Figure S1). To determine the role of Rfx6 specif-

ically in the endocrine lineage and circumvent any effect of a possible endodermal function, we first generated Rfx6 ^{Δ Endo} mutant animals by crossing Rfx6^{fl/fl} mice with Ngn3-Cre mice (Yoshida et al., 2004). Rfx6 is efficiently deleted in Rfx6 ^{Δ Endo} pancreas (Figure 1G), and like the full-body knockout, Rfx6 ^{Δ Endo} pups are diabetic and die between 2 and 3 days after birth (not shown). *Ins1*, *Ins2*, *Gcg*, *Sst*, and *Ghr* mRNAs are strongly decreased, while the amount of *Ppy* transcripts is increased (Figures 1I–1N). Accordingly, the differentiation of hormone-expressing cells was strongly impaired (except for PP), as insulin- glucagon- or somatostatin-expressing cells are hardly found (Figures 1A–1D). However, although the overall amount of *ChgA* mRNA is decreased (Figure 1H), an endocrine program has been implemented since chromogranin A-positive but hormone-negative cells are found (Figures 1E and 1F). We further characterized the role of Rfx6 downstream of Ngn3 by exploring the expression levels of several important transcription factors controlling islet cell fate and maturation. Interestingly, *Ngn3* transcripts were elevated 4-fold in the pancreas of Rfx6 ^{Δ Endo} mutants at birth, suggesting that after being induced by Ngn3 in endocrine progenitors, Rfx6 would in turn repress *Ngn3* in developing islet cells or that the number of Ngn3 cells increases (Figure 1W). Cell quantification did not validate the later hypothesis (not shown). *Arx* transcripts are low, while those encoding *Pax4* are increased (Figures 1O and 1P). Thus, as proposed for the full knockout at embryonic stages (Smith et al., 2010), Rfx6 might control islet subtype specification by repressing the β fate while promoting α destiny in Rfx6 ^{Δ Endo} mice (see Discussion). This increase in *Pax4* was, however, not sufficient to induce β cell development, because the expression of key regulators of insulin transcription (*Pdx1*, *MafA*, *NeuroD1*) as well as *Ins1* was reduced (Figures 1Q, 1T, and 1V). Together, and as reported in Rfx6 null embryos (Smith et al., 2010), Rfx6 ^{Δ Endo} newborns almost entirely lack insulin-positive cells. These findings demonstrate that it is Rfx6 function, downstream of Ngn3, that is essential for β cell generation and that impaired β cell development does not result from defective endoderm.

Rfx6 Is Not Absolutely Necessary for Insulin Production in the Adult Mouse Pancreas

Rfx6 expression is maintained in adult islets, including β cells, suggesting a role in β cell function (Smith et al., 2010; Soyer et al., 2010). Because of the postnatal lethality of Rfx6^{-/-} and Rfx6 ^{Δ Endo} mutants, we generated Rfx6^{fl/fl}; Ins1-CreER^{T2} mice (called hereafter Rfx6 ^{Δ beta}) to determine the role of Rfx6 in adult β cells. All subsequent experiments, to address the role of Rfx6 in adult β cells, were performed with 8- to 10-week-old mice (age of tamoxifen treatment). Ins1-CreER^{T2} mice specifically and efficiently delete floxed alleles in β cells upon tamoxifen treatment (Figure S2). Similarly, in Rfx6 ^{Δ beta} mice, Rfx6 is efficiently deleted (95%) as early as 5 days after

sections where hormone-positive cells were found. Staining for PP (green) and Rfx6 (red) revealed that PP is not dependent on Rfx6 (C and D). Staining for hormones (insulin, glucagon, PP, somatostatin) in green and the panendocrine marker chromogranin A in red show that chromogranin A-positive endocrine cells, which do not express any of the islet hormones, are found in the pancreas of Rfx6 ^{Δ Endo} pups (E and F).

(G–W) qRT-PCR experiments for Rfx6, hormones, and transcription factors controlling islet cell development in Rfx6 ^{Δ Endo} pups and wild-type controls at P0. Scale bars, 50 μ M. Data are presented as mean \pm SD on n = 4 samples; ***p < 0.001, **p < 0.01, *p < 0.05.

the first tamoxifen injection as shown by quantitative RT-PCR (qRT-PCR) on whole islets (Figures 2G and S2) and specifically in insulin-positive β cells (Figures 2A, 2B, and S2). Although *Ins1* transcripts were reduced by ~54%–65% (Table S1; Figure 2H), but not *Ins2* (Table S1), insulin peptide, as well as c-peptide 1 and 2, are still detected in β cells lacking Rfx6 as late as 3 weeks after tamoxifen injections (Figures 2A–2F). In agreement with these observations, Rfx6 ^{Δ Beta} mice do not develop overt diabetes. Detection of insulin hormone in Rfx6 ^{Δ Beta} islets was surprising, because a previous study suggested that Rfx6 was necessary for insulin production in the embryonic pancreas (Smith et al., 2010). Using quantitative chromatin immunoprecipitation (ChIP) in Min6B1 cells (Lilla et al., 2003), we could not reveal any binding of Rfx6 to any of the eight predicted X-boxes in a 10 kb region upstream of *Ins1* gene, while in a control experiment, MafA mapped to the *Ins1* promoter region (not shown). Thus, induced inactivation of Rfx6 in a mature β cell does not preclude the production of insulin hormone.

Impaired Insulin Secretion and Glucose Intolerance in Rfx6 ^{Δ Beta} Mice

To determine whether glucose homeostasis was perturbed in Rfx6 ^{Δ Beta} mice, we measured the blood glucose concentration in fasted and ad libitum-fed mice and conducted intraperitoneal (i.p.) and oral (O) glucose tolerance tests (GTT) in 3-month-old males, 1 month after tamoxifen treatment. Basal blood glucose concentration was similar in Rfx6 ^{Δ Beta} and control mice when either fasted or fed (Figure 3A). After 16 hr fasting, glucose was administered by either i.p. injection or intragastric gavage (2 g/kg body weight). In both conditions, mutant mice displayed significant glucose intolerance (Figures 3B and 3C). Impaired glucose homeostasis in Rfx6 ^{Δ Beta} mice appears not to be due to defects in β cell mass, α cell mass, or β cell proliferation under the same conditions (Figure S3). To determine whether the reduced glucose clearance resulted from altered insulin levels, we measured plasma insulin in Rfx6 ^{Δ Beta} and control mice 5 days after tamoxifen treatment (Figure 3D). In fasted animals, the plasma insulin level in Rfx6 ^{Δ Beta} and control mice was similar (mean controls = 0.35 ng/ml; mean mutants = 0.37 ng/ml; $n = 7$, $p > 0.05$). However, significant differences were observed following i.p. glucose injection (3 g/kg body weight), as glucose-induced insulin secretion was much less efficient in Rfx6 ^{Δ Beta} mice (Figure 3D). Both the first- and second-phase responses were reduced in mutant mice. To determine whether the defect was islet autonomous, we purified islets from Rfx6 ^{Δ Beta} and control adult mice under the same conditions and performed ex vivo glucose-stimulated insulin secretion (GSIS) and KCl-stimulated insulin secretion (KCl-SIS) studies (Figure 3E). The amount of secreted insulin was normalized to the total insulin content, which was identical in controls and mutants (not shown), confirming that Rfx6 is not key for insulin production in adult β cells. Clearly, the extent of insulin secretion in response to a glucose or KCl stimulus was reduced in mutant islets compared with controls. Notably, the altered KCl-induced response suggests that the defective insulin secretion might result from glucose-independent depolarization defects. These results suggest that impaired glucose clearance results from

insulin secretory failure that we detect as early as 1 week after Rfx6 deletion both in vivo and ex vivo (in islets).

Rfx6 Binds to X-Boxes in the *Gck* and *Abcc8* Genes, and Their Expression Is Strongly Decreased in Rfx6 ^{Δ Beta} Islets

Next, our goal was to decipher the molecular basis underlying defective insulin secretion and glucose intolerance in Rfx6 ^{Δ Beta} mice. Therefore, we performed a series of qRT-PCR experiments to measure the expression of key regulators of glucose simulated insulin secretion in islets isolated from mutant or control mice, 5 days after tamoxifen treatment. Importantly, we found that transcripts encoding the key genes *Gck* and *Abcc8* were dramatically reduced (Figures 4A and 4B; Table S1). *Gck* encodes glucokinase, an enzyme that controls the first step of glycolysis and is considered the rate-limiting step in glucose metabolism (Matschinsky, 2005). *Abcc8/SUR1* encodes the regulatory sulphonylurea-binding subunits of the ATP-sensitive K⁺ channel (K_{ATP} channels) linking glucose metabolism to the electrical activity of β cells (McTaggart et al., 2010). Furthermore, *Ucn3* mRNA, a marker of β cell maturation and also reported to have a positive effect on GSIS (Li et al., 2007), is strongly decreased (Figure 4J). In contrast, the expression of the glucose-facilitated transporter *Slc2a2/Glut2*, the proconvertase *Pcsk1*, and the pore-forming unit of K_{ATP} channels (*Kcnj11/Kir6.2*) was not altered (Figures 4C–4E). Among the transcription factors important for β cell function and insulin transcription tested (*MafA*, *Pdx1*, *Nkx6.1*, *Pax6*, and *NeuroD1*), only *NeuroD1* and *Pax6* transcripts were downregulated (Figures 4F–4I; Table S1).

To determine whether Rfx6-regulated genes might be direct targets of Rfx6, we analyzed chromatin immunoprecipitation sequencing (ChIP-seq) experiments performed in the mouse β cell line Min6B1 (Lilla et al., 2003) transfected with hemagglutinin (HA)-tagged Rfx6. One peak was detected in the β cell-specific *Gck* promoter region (Figure 4K) overlapping with Pal1 and Pal2 regions previously demonstrated to contain X-boxes bound by Rfx3 (Ait-Lounis et al., 2010). Similarly, a peak was observed in a conserved region of intron 10 of the *Abcc8* gene (Figure 4L). Quantitative ChIP experiments in nontransfected Min6B1 cells, using an Rfx6 antibody, confirmed that endogenous Rfx6 binds to the *Gck* and *Abcc8* genes, although less strongly to *Gck* (Figures 4M and 4N). Taken together, our results suggest that Rfx6 is a direct positive transcriptional regulator of *Gck* and *Abcc8* and that the downregulation of these genes could impact glucose-stimulated insulin secretion.

Glucose Fails to Induce Normal Increases in ATP/ADP Ratio and Intracellular Free Ca²⁺ in Rfx6 ^{Δ Beta} Mice

Given the abnormalities in glucose-stimulated insulin secretion and gene expression described above, we next explored the possibility that signal generation by the sugar may be impaired in Rfx6 null β cells. Examined using functional multicellular imaging within intact islets (Hodson et al., 2013, 2014a), both phases of the intracellular calcium ([Ca²⁺]_i) response to a step increase in glucose concentration were sharply (50%–60%) reduced by Rfx6 deletion (Figures 5A and 5D). A smaller (~20%) reduction was also apparent in the overall number of

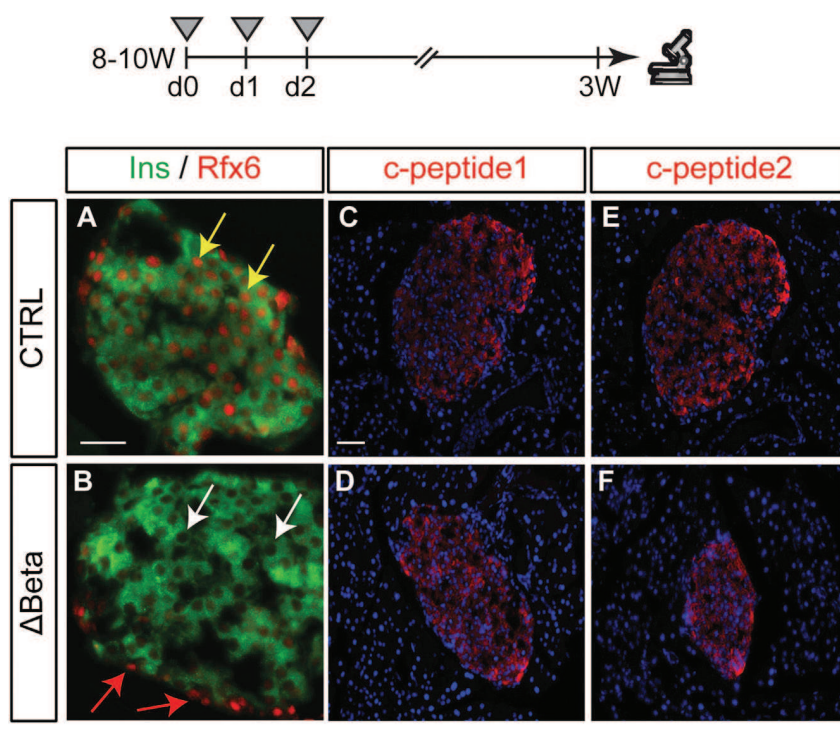
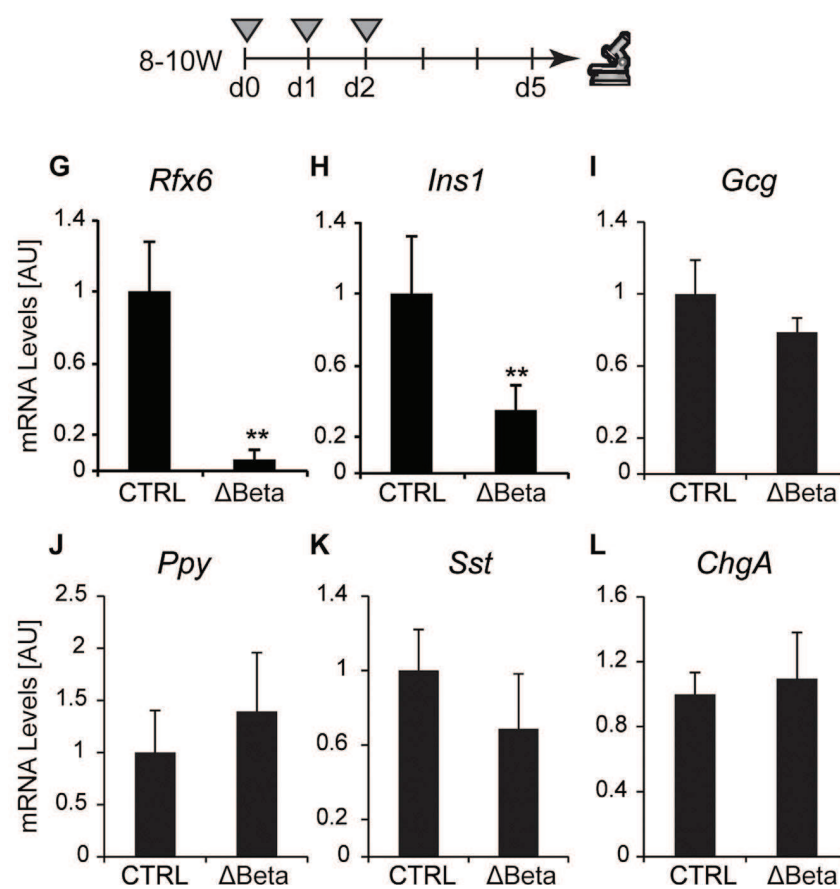


Figure 2. Insulin Is Produced in Adult β Cells Lacking Rfx6

(A–F) Immunofluorescence staining on controls and Rfx6 Δ Beta (Rfx6^{fl/fl}; Ins1-CreERT²) adult mice, 3 weeks after the first day of tamoxifen injections (8- to 10-week-old mice were injected once a day during 3 consecutive days). Staining for insulin (green, A and B) and Rfx6 (red, A and B) reveal insulin expression despite efficient deletion of Rfx6 in β cells of Rfx6 Δ Beta mice (white arrows in B). Staining for c-peptide1 (red, C and D) and c-peptide2 (red, E and F) supports efficient insulin synthesis.

(G–L) qRT-PCR experiments on islets purified from controls and Rfx6 Δ Beta adult (8- to 10-week-old) mice 5 days after tamoxifen injections revealing rapid and specific deletion of Rfx6 (G) in β cells and decreased *Ins1* transcription (H), while the expression of *Gcg*, *Ppy*, *Sst* and *ChgA* is unaltered. Grey triangles indicate the days of tamoxifen injections. Yellow and red arrows point to examples of β cells expressing Rfx6 in controls and insulin-negative/Rfx6-positive cells in Rfx6 Δ Beta mice, respectively.

Scale bars, 50 μ M. Data are presented as mean \pm SD on n = 4 samples; ***p < 0.001, **p < 0.01, *p < 0.05.



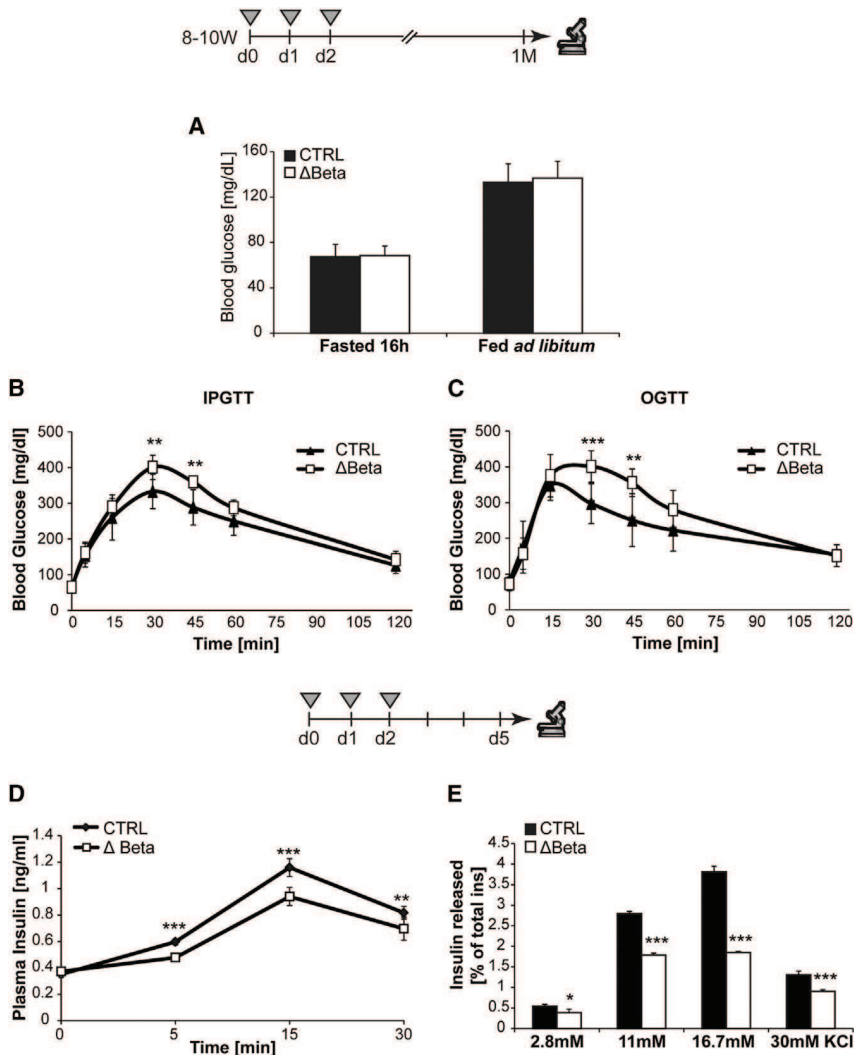


Figure 3. The Deletion of *Rfx6* in β Cells Causes Glucose Intolerance due to Defective Insulin Secretion

(A–C) Exploration of glucose metabolism in adult (12- to 14-week-old) controls ($n = 8$) and $Rfx6^{\Delta\beta}$ ($n = 7$) males under normal diet 4 weeks after tamoxifen injections. Blood glucose levels are measured in overnight (16 hr)-fasted and ad libitum-fed animals 1 month after tamoxifen injections (A). Intraperitoneal glucose tolerance test (IPGTT) after 16 hr fasting in male mice 1 month after tamoxifen injections (B). Oral glucose tolerance test (OGTT) after 16 hr fasting in male mice 1 month after tamoxifen injections (C).

(D) Histogram representing the plasma insulin levels of control and $Rfx6^{\Delta\beta}$ mice (9–11 weeks old) during an in vivo glucose-stimulated insulin secretion test ($n = 6$) performed 5 days after tamoxifen injections.

(E) Histogram representing the insulin released during an ex vivo glucose- and KCl-stimulated insulin secretion tests on islets purified from controls and $Rfx6^{\Delta\beta}$ ($n = 4$) mice (9–11 weeks old), 5 days after tamoxifen injections.

Data are presented as mean \pm SD; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

responsive cells (Figure 5B). These changes were accompanied by more modest, but significant, defects in the response to depolarization with KCl (Figures 5C and 5E), suggesting a defect downstream of metabolic signal generation by glucose. Defects in the latter process were, nonetheless, highlighted by a substantial (~60%) impairment in glucose-induced cytosolic ATP/ADP increases (Hodson et al., 2014b; Tarasov et al., 2012) (Figures 5F and 5G), consistent with lowered *Gck* expression.

In line with the above defects in calcium influx, RNA sequencing (RNA-seq) performed on $Rfx6^{\Delta\beta}$ and control whole islets, 5 days after tamoxifen treatment, revealed that the expression of eight voltage-dependent calcium channels (VDCCs) transcripts was downregulated by the deletion of *Rfx6* in adult β cells (Table S2). These genes included *Cacna1a*, *Cacnb2*, *Cacna1d*, and *Cacna1c*, which are among the most abundantly expressed VDCCs in rodent islets (Table S2). Downregulation of *Cacna1d*, *Cacna1c*, and *Cacnb2* was confirmed by RT-PCR (Figures 5H–5J). ChIP sequencing (ChIP-seq) in Min6B1 cells revealed that *Rfx6* binds to two of the *Rfx6*-dependent

VDCC genes, at X-boxes in intron 1 of *Cacna1c* and *Cacnb2* genes, respectively (Figures 5K and 5L), suggesting that *Rfx6* can directly regulate these genes, a result confirmed by quantitative ChIP (Figures 6M–6O). Collectively, these data suggest that compromised insulin secretion in *Rfx6*-deficient β cells is the consequence, at least in large part, of the combined alteration of key steps in glucose signaling due to a failure of *Rfx6* to directly promote the expression of key genes. These include both glycolysis, impaired by downregulation of *Gck*, and Ca^{2+} influx, affected by the repression of voltage-gated calcium channels.

Derepression of the Disallowed Genes in β Cells Lacking *Rfx6*

We next sought to determine whether additional mechanisms may also play a role in defective glucose-induced insulin secretion in *Rfx6* null β cells. *Rfx6* transcription factors have been reported to be either activators or repressors of transcription (Aftab et al., 2008). Accordingly, RNA-seq revealed a series of genes with a higher expression in β cells after *Rfx6* deletion. Surprisingly, we found a specific upregulation of the so-called disallowed or forbidden genes. Disallowed genes are genes that are abundantly expressed in most tissues but selectively repressed in adult mouse β cells. Two different studies identified a total of 68 of these genes by microarray analysis (Pullen et al., 2010; Thorrez et al., 2011). We found that 54 out of the 68 disallowed genes reported were upregulated in $Rfx6^{\Delta\beta}$ islets, including the 11 common genes identified in both studies (Pullen

and Rutter, 2013), namely *Slc16a1*, *Ldha*, *Pdgfra*, *Igfbp4*, *Cxcl12*, *Oat*, *Smad3*, *Lmo4*, *C1qbp*, *Maf*, and *Cd302* (Table S3). We confirmed the upregulation of some of the selected genes by qRT-PCR: *Ldha*, *Slc16a1*, *Pdgfra*, and *Igfbp4* (Figures 6A–6D). ChIP-seq and quantitative ChIP in Min6B1 cells revealed that Rfx6 binds to a conserved region approximately 10 kb upstream of *Ldha* gene (Figures 6E and 6F), suggesting that Rfx6 could directly regulate the repression of this gene in β cells. Some of the disallowed genes have been reported to modulate insulin secretion (see Discussion and Pullen and Rutter, 2013). Thus, by repressing the disallowed genes, Rfx6 contributes to the maintenance of β cell identity and function.

DISCUSSION

Previous studies have revealed that Rfx6 is required during development in rodents and in humans for the generation of β cells (Smith et al., 2010). The present work now demonstrates that this factor is also essential for the maintenance of the functional identity of the adult β cell.

Rfx6 and Insulin Production

To determine the role of Rfx6 in β cell function, we conditionally and specifically inactivated this factor in mature adult β cells. The resulting Rfx6 ^{Δ Beta} mice are glucose intolerant in contrast to the phenotype of the *Rfx6* null mice, which die shortly after birth with severe and sustained hyperglycemia (>600 mg/dl). This difference indicates that Rfx6 has a distinct function in adult β cells that is different from its earlier function in the embryonic pancreas. In the embryo, our data in Rfx6 ^{Δ Endo} mouse pancreas clearly show that Rfx6 is required downstream of Ngn3 for proper β cell development, when cells are already committed to an islet fate. It is unclear whether the strongly impaired expression of insulin transcripts and peptide in the Rfx6 ^{Δ Endo} embryonic pancreas reflects a blockade in the β cell differentiation program or a direct regulation of the transcription of *Insulin* genes by Rfx6. However, it is likely that Rfx6 has a very early function downstream of Ngn3 and implements genetic programs essential for islet subtype specification to proceed. This hypothesis is supported by the up- and downregulation of *Pax4* and *Arx*, respectively (also observed in the constitutive Rfx6 knockout; our own data and Smith et al., 2010), known to control β - δ versus α destiny (Collombat et al., 2003). One possibility would be that in a wild-type situation, Rfx6 would favor the α destiny by promoting *Arx* expression and simultaneously blocking the β fate by repressing *Pax4*. However, Rfx6 must have an additional role; otherwise, we would have observed an increased number of β cells in Rfx6 ^{Δ Endo} embryos, which is not the case. The maintenance of generic markers of endocrine differentiation such as chromogranin A testifies that endocrine differentiation proceeds in the absence of Rfx6. On the other hand, the strong downregulation of *Ins1*, *Ins2*, *Gcg*, *Ghr*, and *Sst* mRNAs suggests that Rfx6 implements a genetic program required for hormone production in embryonic islet cells. In adult mice, the overall production of insulin is not altered in Rfx6 ^{Δ Beta} islets and thus Rfx6 is not indispensable for insulin production in β cells. However, *Ins1* transcripts are decreased when *Rfx6* is deleted in β cells, while *Ins2* transcripts do not vary. A similar result has been reported in mice

lacking *NeuroD1* in β cells (Gu et al., 2010), although *Ins1* was reduced by 95% in this model, in contrast to 65% in Rfx6 ^{Δ Beta} islets. ChIP-seq and quantitative ChIP in Min6B1 cells did not reveal any binding of Rfx6 to putative X-boxes, suggesting that Rfx6 does not directly regulate *Ins1* gene, although we cannot exclude that it is the case in bona fide β cells. Furthermore, like in mice lacking *NeuroD1* in β cells, the expression of other regulators of *Insulin* transcription such as *MafA*, *Nkx6.1*, and *Pdx1* is unchanged in Rfx6 ^{Δ Beta} islets, suggesting that the reduction in *Ins1* transcripts results from decreased levels of *NeuroD1*. Decreased *Ins1* transcript levels could also result from low *Pax6*, which is necessary for insulin synthesis in adult β cells (Hart et al., 2013). Importantly, we observed a downregulation, and not a complete loss, of *NeuroD1* and *Pax6* and suspect that levels of these transcription factors are sufficient to account for the insulin mRNA levels measured in Rfx6 ^{Δ Beta} islets. However, we did not observe any signal reduction in anti-insulin immunofluorescence experiments, and total insulin content was not affected either, probably because the reduction of *Ins1* mRNA was less severe compared to mice lacking *NeuroD1* in β cells. Such a decrease in *Ins1* transcripts without any significant effect on insulin content has been described previously, such as in *MafA*-deficient mice (Zhang et al., 2005), and may reflect alterations in mRNA translation.

Rfx6 and Insulin Secretion

We found that without the transcription factor Rfx6, insulin secretion is impaired in β cells and mice become glucose intolerant. Defective insulin secretion resulted from a combined effect to reduce expression of three core components of the glucose-stimulated signaling pathway in β cells, which we show to be direct targets of Rfx6. First, mRNAs encoding the glucose sensor Glucokinase were decreased (~75%) in Rfx6 ^{Δ Beta} islets. Glucokinase is the flux-generating enzyme of oxidative glycolysis, coupling blood glucose concentration to metabolic signals in β cells, ultimately leading to rises in cytosolic ATP/ADP ratios triggering the closure of K_{ATP} channels and subsequent membrane depolarization. Gene-deletion experiments in mice have demonstrated that *Gck* is essential for insulin secretion and maintenance of glucose homeostasis and that reduction in *Gck* gene dosage (*Gck*^{+/-} mice) is sufficient to induce mild hyperglycemia (Terauchi et al., 1995). We thus believe that decreased *Gck* levels are the principal cause of the failure of the cytosolic ATP/ADP ratio to increase appropriately upon glucose stimuli in Rfx6 ^{Δ Beta} islets. Interestingly, insulin secretion reaches a plateau at 11 mM glucose in Rfx6 ^{Δ Beta} islets, and further increases in glucose concentrations have no effect on insulin secretion, demonstrating that the response to glucose is saturated. Though other steps in the pathway may also be affected, this change in glucose dose response is consistent with a shift in the control of glycolytic flux toward other hexokinase family members (HKI-III) with lower Michaelis constants for the sugar. Second, we found that *Abcc8/SUR1* transcripts encoding the regulatory subunit of the ATP-sensitive K^+ channel (K_{ATP} channels) are downregulated as well in Rfx6 ^{Δ Beta} islets. Thus, we propose that lowered *Abcc8*, which would tend to raise the resting plasma membrane potential at low glucose while impairing stimulation by the sugar (Nenquin et al., 2004), combined with suboptimal increases in

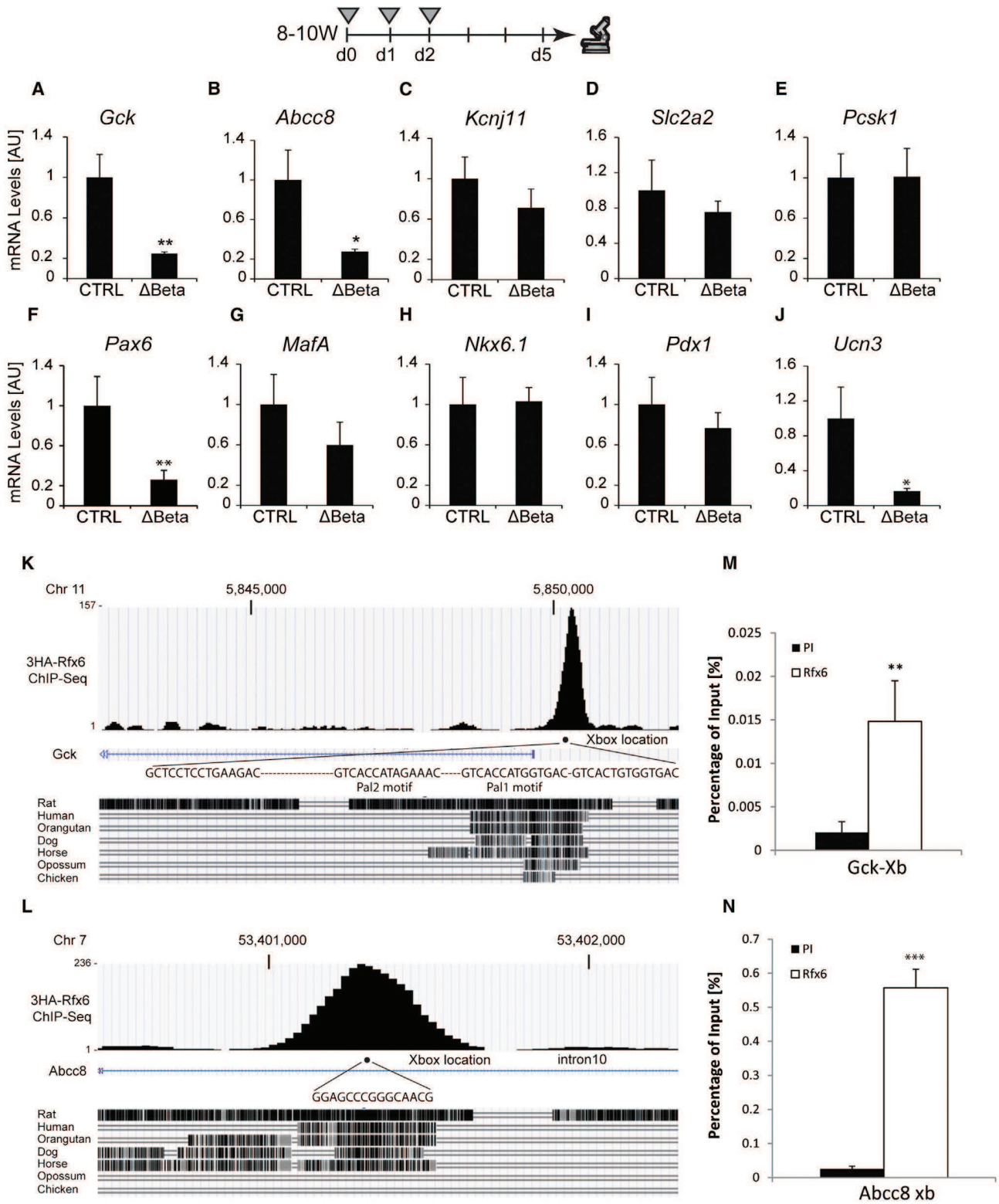


Figure 4. Rfx6 Regulates a Subset of Genes Controlling β Cell Function and Directly Targets *Gck* and *Abcc8*
(A–J) qRT-PCR experiments for *Gck*, *Kcnj11*, *Abcc8*, *Slc2a2*, *Pcsk1*, *Pax6*, *MafA*, *Nkx6.1*, *Pdx1*, and *Ucn3* on islets purified from controls and Rfx6 Δ Beta adult mice (9–11 weeks old), 5 days after tamoxifen injections. Grey triangles indicate the days of tamoxifen injections.

(legend continued on next page)

ATP/ADP ratios, alters β cell membrane excitability. This, in turn, will affect the opening of voltage-dependent Ca^{2+} channels, at least partially explaining impaired glucose-induced Ca^{2+} influx. Third, we found that Rfx6 is also critical for the expression of several voltage-dependent calcium channels mRNAs of the L, R, and P/Q type, which are downregulated in Rfx6 ^{Δ Beta} islets. These include central players such as the L-type Ca^{2+} channel subunit alpha 1C (*Cacna1c*, also named as *Ca_v1.2*) and the R-type Ca^{2+} channel alpha 1E (*Cacna1e*, also named *Cav2.3*) subunits, which have been shown to be essential for insulin secretion in mice by regulating first and second phases of secretion, respectively (Jing et al., 2005; Schulla et al., 2003). Thus, in addition to decreased glycolysis and reduced levels of K_{ATP} channel Abcc8/SUR1 subunits, lowered voltage-gated Ca^{2+} channels likely contributes to the impaired elevation of intracellular Ca^{2+} and accounts for defective KCl-stimulated insulin secretion. Finally, we noted that the zinc transporter *Slc30a8* (*ZnT8*), required for normal insulin crystallization and secretion (Rutter, 2010), was also repressed after *Rfx6* deletion (Table S1), suggesting that later events in the exocytotic process may also be affected. Nevertheless, and despite multiple deficiencies in several steps controlling glucose-induced secretion, Rfx6 ^{Δ Beta} are only mildly glucose intolerant. As we did not reveal any increase in β cell mass, we suggest that Rfx6 ^{Δ Beta} mice compensate their impaired insulin secretion by other mechanisms, which may include an increased insulin sensitivity or higher secretion of glucocorticoids.

Importantly, ChIP experiments in the mouse β cell line Min6B1 suggest that Rfx6 binds to X-boxes in *Gck*, *Abcc8*, and the VDCC *Cacna1c* and *Cacnb2* genes. These results, together with the expression data of Rfx6 ^{Δ Beta} islets (RNA-seq), clearly suggest that Rfx6 controls β cell function in the adult by directly regulating the expression of key core components of the insulin-secretion pathway. Interestingly, *Gck* has been previously reported to be a direct target of Rfx3 (Ait-Lounis et al., 2010). The fact that we independently found, by an unbiased method (ChIP-seq), that Rfx6 binds to the very same region in the *Gck* promoter as Rfx3 strongly supports that Rfx transcription factors are important regulators of *Gck* expression. However, Rfx3 (*Rfx3* expression is unaffected in Rfx6 ^{Δ Beta} islets) was not sufficient to compensate for the absence of Rfx6, regarding *Gck* expression, suggesting that Rfx3-Rfx6 heterodimers bind to the *Gck* promoter for the optimal transcription of *Gck* gene. Binding of Rfx3 to *Abcc8* or voltage-dependent Ca^{2+} channel genes has not been reported, and although Rfx6 and Rfx3 might coregulate other targets genes, the extent of the functional redundancy of both transcription factors in β cell function and insulin secretion cannot be evaluated, as the phenotype of mice with a deletion of *Rfx3* in adult β cells has not been described. Likewise, whether a small but significant (~30%) decrease in *Rfx5* expression (Table S1) observed here also contributes to the phenotype of *Rfx6* null adult mice remains a question for the future.

Rfx6 Contributes to the Silencing of the Disallowed Genes and Maintenance of β Cell Maturity

A group of disallowed or “forbidden” housekeeping genes has recently been described to be selectively repressed in β cells. The potential role of their repression in β cell function and glucose homeostasis (discussed in detail in Pullen and Rutter, 2013) is still to be fully elucidated. However, it has been proposed for the two founder members, lactate dehydrogenase A (*Ldha*) and monocarboxylate transporter-1 (*MCT-1/Slc16a1*), that their downregulation in β cells prevents lactate and pyruvate, produced from muscle during exercise, from inappropriately stimulating insulin release (Thorrez et al., 2011). At present, it is unclear whether a general mechanism controls the silencing of all (or most) of the disallowed genes. However, epigenetic mechanisms such as the trimethylation of histone H3 on lysine 27 (H3K27me3) has been suggested to be the cause of disallowed gene repression, as H3K27me3 was found present on *Slc16a1* promoter (van Arensbergen et al., 2010) and a few other forbidden genes (*Cxcl12*, *Acot7*, *Nfib*, *Mgst1*, and *Maf*) (Pullen and Rutter, 2013). Likewise, DNA methylation is reported also to be involved in the repression of some (e.g., *Acot7*) (Dayeh et al., 2014), but not other (*Slc16a1/MCT1*) (Pullen et al., 2011), members of this family. On the other hand, microRNAs have also been reported to downregulate *Slc16a1* (Pullen et al., 2011). Our current study provides evidence of another mechanism whereby a transcription factor (Rfx6) may act as a master regulator, repressing disallowed genes and serving as a common mechanism for disallowed gene downregulation, as we found that ~79% of these genes were significantly upregulated in Rfx6 ^{Δ Beta} islets. We found that Rfx6 binds to an X-box located ~10 kb upstream of *Ldha* start site in Min6B1 cells, supporting direct repression. Along the same lines, and further supporting the notion that Rfx6 directly targets disallowed genes, we identified Rfx6 binding peaks in the vicinity of 32 out of 54 (59%) Rfx6-repressed disallowed genes (Table S3). Of note, increased expression of *Ldha* has also been described when *NeuroD1* is deleted in β cells (Gu et al., 2010). We thus cannot exclude that the repression of *NeuroD1*, observed in Rfx6 ^{Δ Beta} islets, also contributes to elevated *Ldha* transcript levels, a feature of neonatal β cells, in contrast to mature β cells, which have low amounts of *Ldha* (Gu et al., 2010; Sekine et al., 1994). Importantly, other members of the disallowed family have not been reported to be dysregulated in *NeuroD1*-deficient β cells. Taken together, our own and others' data suggest that repression of disallowed genes is achieved by multiple mechanisms including transcription factor-mediated repression of gene expression. Tissue-specific gene repression is thought to proceed during the maturation of islet cells during postnatal stages (Thorrez et al., 2011), a period when β cells acquire a mature GSIS (Blum et al., 2012). Interestingly, we did not observe any change in the expression of disallowed genes in the embryonic pancreas of *Rfx6*-deficient mice (data not shown). Thus, we propose that

(K and L) ChIP-seq (anti-HA) data showing Rfx6 binding peaks in *Abcc8* and *Gck* genes in 3HA-Rfx6 transfected Min6B1 cells.

(M and N) Quantitative ChIP (anti-Rfx6 antibody) in Min6B1 cells illustrating the binding of Rfx6 to X-boxes indicated in (K) and (L). PI and Rfx6 stand for preimmune and anti-Rfx6 serum.

Data are presented as mean \pm SD on n = 4–5 samples; **p < 0.01, *p < 0.05.

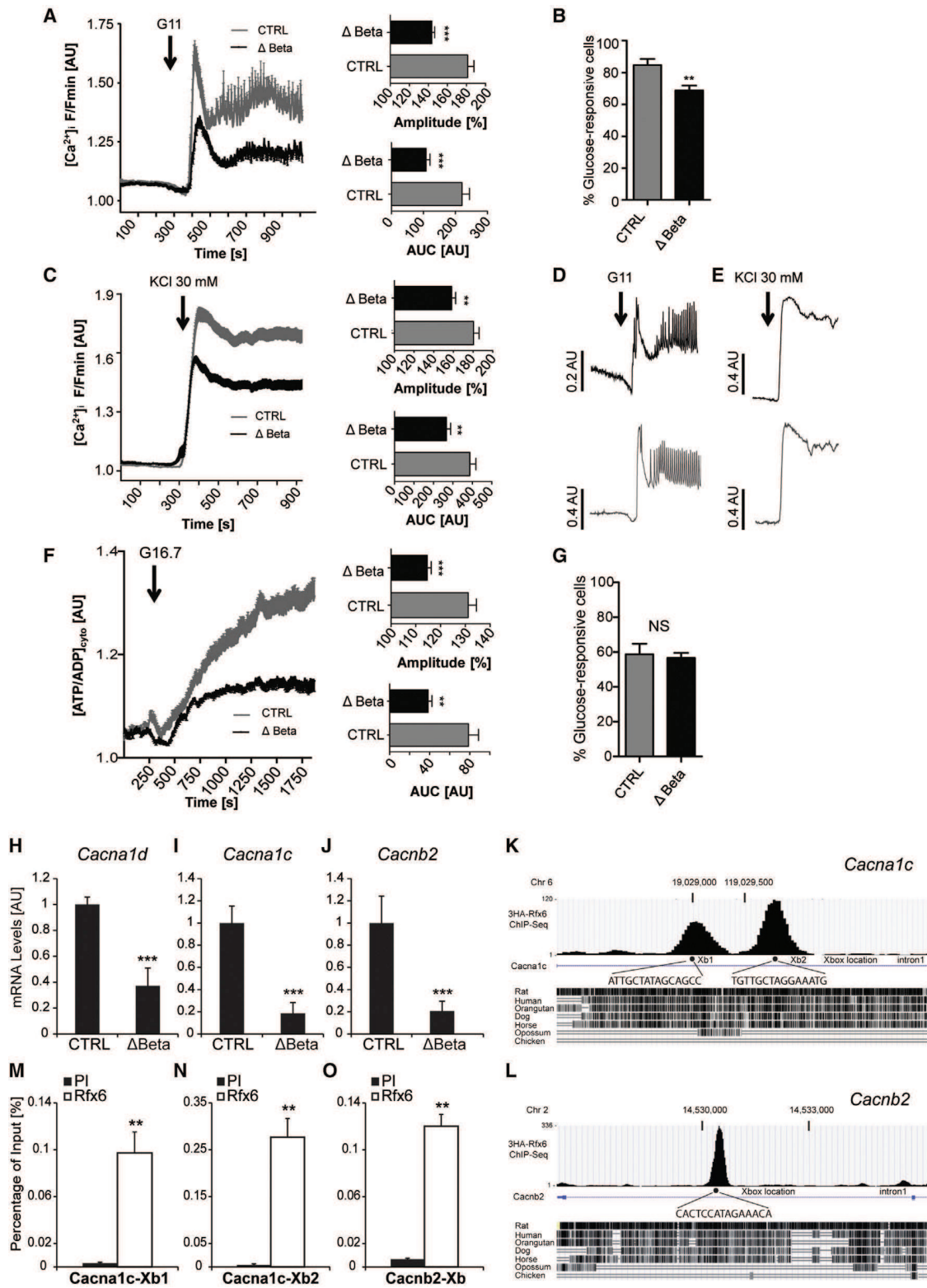


Figure 5. Impaired ATP/ADP Ratio and Calcium Trafficking in Glucose Stimulated β Cells in *Rfx6* ΔBeta Islets

(A) Mean (\pm SEM) Ca^{2+} traces following elevation of glucose from 3 mM to 11 mM ($n = 16$ islets from five mutants and six controls). Insets are the amplitude and area under the curves.

(legend continued on next page)

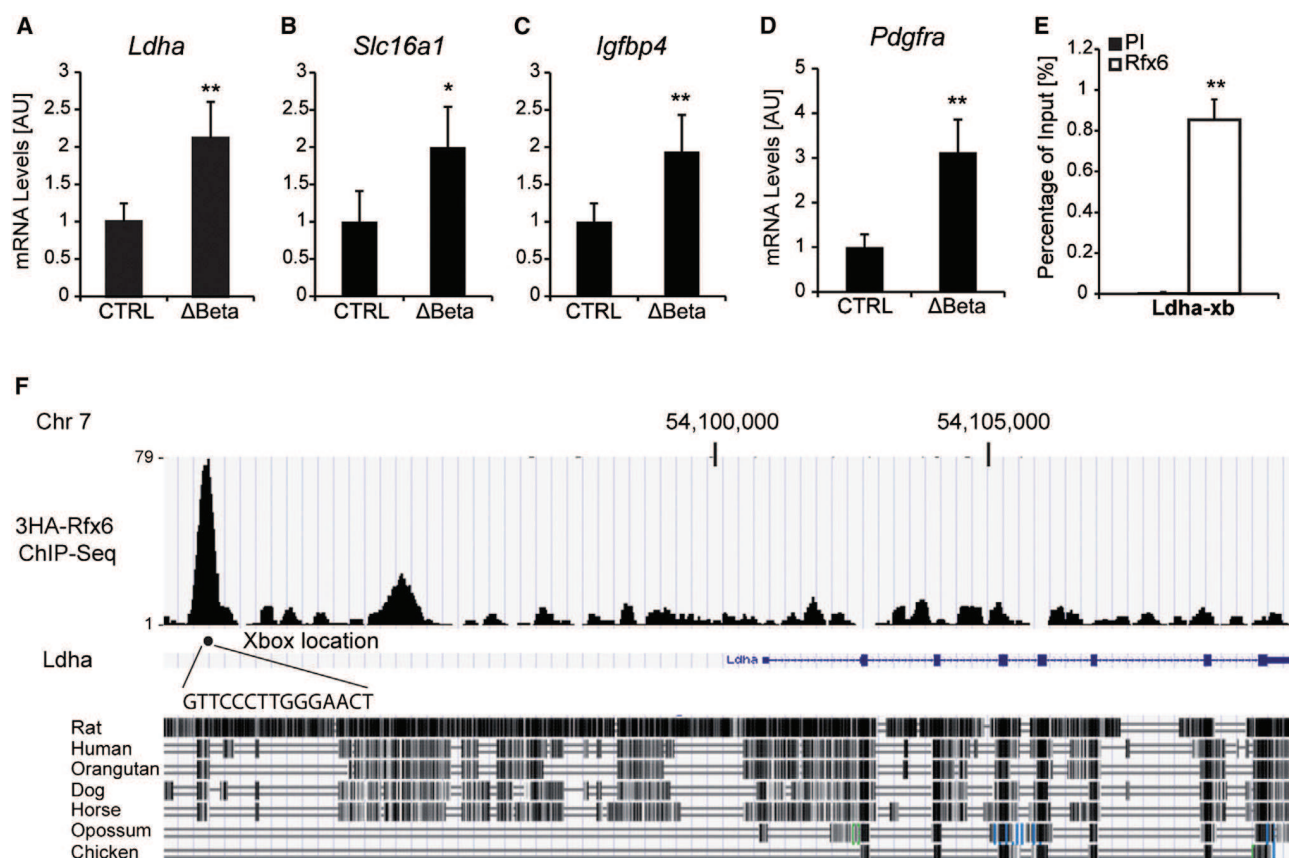


Figure 6. Rfx6 Targets and Represses Disallowed Genes in β Cells

(A–D) qRT-PCR showing the downregulation of *Ldha*, *Slc16a1*, *Igfbp4*, and *Pdgfra* in islet cells from Rfx6 Δ^{Beta} mice (9–11 weeks old), 5 days after tamoxifen injections compared to controls (n = 4).

(E and F) ChIP-PCR and ChIP-seq revealing the binding of Rfx6 on one X-box in *Ldha* gene in Min6B1 cells (n = 3).

Data are presented as mean \pm SD; **p < 0.01, *p < 0.05.

Rfx6 is essential to establish and maintain the repression of the disallowed genes and thereby the maturity of β cells. The extent to which perturbation of disallowed gene expression contributes to the phenotype of Rfx6 Δ^{Beta} mice remains to be studied. However, it was shown previously that overexpression of *Ldha* diminishes glucose-stimulated insulin secretion in islets (Ainscow et al., 2000; Ishihara et al., 1999). Moreover, forced overexpression of *Slc16a1* in β cells leads to insulin secretion in response to muscle-derived pyruvate (Pullen et al., 2012) and is thought to be responsible for exercise-induced hyperinsulin-

ism in humans (Otonkoski et al., 2007). In addition to *Ldha* and *Slc16a1*, several other disallowed genes potentially impact insulin secretion and glucose homeostasis, acting as second messengers controlling insulin secretion or regulating trafficking of insulin granules (see Pullen and Rutter, 2013 for a review). Thus, Rfx6 controls β cell maturity and glucose homeostasis by repression of the disallowed genes and activation of core component of the insulin secretion pathway, including glucokinase, the Abcc8/SUR1 subunit of K_{ATP} channels and voltage-gated Ca²⁺ channels.

(B) Proportion of fluo2-loaded cells that respond to 11 mM glucose.

(C) Mean (\pm SEM) Ca²⁺ traces following application of the depolarizing stimulus 30 mM KCl (n = 18 islets from same animals as A).

(D) Representative Ca²⁺ responses to 11mM glucose in a single islet (average of about 50 responsive cells per islet).

(E) Representative Ca²⁺ responses to 30 mM KCl in a single islet.

(F) Mean (\pm SEM) Perceval traces recording ATP dynamics following elevation of glucose from 3 mM to 16.7 mM (~250 cells from n = 12 islets from four mutants and four controls). Insets are the amplitude and area under the curves of cytosolic ATP/ADP rises ([ATP/ADP]_{cyto}).

(G) Proportion of Perceval-expressing cells, shown to represent beta cells (Hodson et al., 2014a) that respond to 16.7 mM glucose.

(H–J) qRT-PCR revealing the decrease in the transcription of VDCCs in Rfx6 Δ^{Beta} islets (n = 4).

(K–O) ChIP-seq (K and L) and quantitative ChIP (M–O) showing the binding of Rfx6 on X-boxes of *Cacna1c* and *Cacnb2* genes in Min6B1 cells (n = 3). PI and Rfx6 in (M) and (N) stand for preimmune and anti-Rfx6 serum.

Data are presented as mean \pm SD; ***p < 0.001, **p < 0.01.

We show here that Rfx6 is required for normal β cell identity, sustaining the expression of signature β cell genes (*Gck*, *Abcc8*) and repressing that of disallowed genes. These findings raise the possibility that changes in RFX6 expression may contribute to β cell failure in type 2 diabetes (T2D) in humans. To date, however, no such changes in *RFX6* mRNA have been reported in T2D versus healthy donor β cells (Marselli et al., 2010; Taneera et al., 2012), despite the increased expression of 15 disallowed genes in human T2D (Pullen and Rutter, 2013). However, we would point out that these studies involved a relatively small number of subjects, and therefore, we need to wait for the results of much larger ongoing trials to conclude whether subtle changes in *RFX6* levels are associated with T2D. Furthermore, changes in RFX6 activity resulting from altered subcellular localization or posttranslational modifications cannot be excluded in diseased islets. Interestingly, fasting hyperglycemia has been reported in a patient bearing a heterozygote mutation in the RFX6 gene and SNP susceptibility to T2D, supporting a role for RFX6 in β cell function in humans (Artuso et al., 2014).

Finally these results suggest that, in addition to having fewer β cells, sufferers of neonatal diabetes who carry *RFX6* mutations are also likely to have defects in those cells that remain (Mitchell et al., 2004). Whether these individuals may therefore be susceptible to pharmacological treatments, including sulphonylureas or GLP-1 analogs, which could prompt insulin release from the remaining β cells, may be worthy of investigation.

EXPERIMENTAL PROCEDURES

Immunostaining and Morphometric Analysis of Islets

Pancreata were fixed with 4% paraformaldehyde, cryo- or paraffin embedded, and stained with the primary and secondary antibodies listed in Supplemental Experimental Procedures. Antigen retrieval was used prior to staining for Rfx6. For bromodeoxyuridine (BrdU) detection assays, BrdU (50 mg/kg body weight) was injected 24 hr before sacrifice to assess proliferation in adult β cells. For α and β cell mass, quantification was performed every 2 mm and calculated as described in Figure S3. Four animals of each genotype were analyzed. Animal experiments were supervised by G.G. (agreement N°C67-59 approved by the Direction des Services Vétérinaires, Strasbourg, France) in compliance with the European legislation on the care and use of laboratory animals.

Insulin Secretion Assay

For each animal ($n = 4$ per genotype, 8–10 weeks old), quadruplicates of five starved islets were placed in Eppendorf tubes containing 2 ml of Krebs' buffer containing 2.8 mM glucose, 11 mM glucose, 16.8 mM glucose, or 2.8 mM glucose + 30 mM KCl and incubated for 1 hr, and the supernatant was collected to measure insulin secretion. Quadruplicates of five unstimulated islets were sonicated and extracted by acid-ethanol. Insulin in supernatants and islet lysates was measured by ELISA (ultrasensitive insulin ELISA, ALPCO). Secreted insulin was then normalized with total lysate insulin content and expressed as a percentage of total insulin.

To measure insulin secretion in vivo, 9- to 11-week-old males were fasted for 16 hr, and we collected blood from animals ($n = 6$ per genotype) before glucose injection and 5, 15, and 30 min after glucose injection. D-glucose solution (15%) was injected intraperitoneally at 3 g/kg body weight. Plasma was separated from blood by centrifugation, and circulating insulin was measured by ELISA (ultrasensitive insulin ELISA, ALPCO).

RNA Sequencing

Total RNA was extracted from adult islets from three controls and three mutants, and RNA integrity was assessed. After sequencing (HiSeq 2500, 50

base reads), reads were mapped onto the mm9 assembly of the mouse genome by using Tophat v1.4.1 (Trapnell et al., 2009) and the bowtie v0.12.7 aligner (Langmead et al., 2009). Only uniquely aligned reads have been retained for further analyses.

Calcium and ATP Imaging

Functional multicellular Ca^{2+} - and ATP/ADP-imaging were performed as previously described (Hodson et al., 2013; Hodson et al., 2014a, 2014b; Tarasov et al., 2012). Briefly, fluo2-loaded or Perceval-expressing islets were mounted in a custom-manufactured aluminum heated chamber and perfused with a HEPES-bicarbonate buffer (120 mM NaCl, 4.8 mM KCl, 24 mM NaHCO_3 , 0.5 mM Na_2HPO_4 , 5 mM HEPES, 3 mM D-glucose, 2.5 mM CaCl_2 , and 1.2 mM MgCl_2) saturated with 95% O_2 /5% CO_2 and adjusted to pH 7.4. During recording, islets were maintained at 36°C and drugs/glucose delivered through the perfusion system at the indicated concentrations. Excitation was performed using a 491 nm solid-state laser (Cobalt) coupled to a Yokogawa CSU10 Nipkow spinning disk head, and emitted signals were captured at 525 ± 50 nm using a Hamamatsu 16-bit electron-multiplying charge-coupled device. An adenoviral vector was used to deliver cDNA encoding Perceval into the first few islet layers (multiplicity of infection, 10–100; 48 hr incubation).

Statistics

Values are presented as mean \pm SD, and p values were determined using the two-tailed Student's t test with unequal variance. $p < 0.05$ was accepted as statistically significant.

ACCESSION NUMBERS

RNA-seq and ChIP-seq data have been deposited to the Gene Expression Omnibus under the accession numbers GSE59622 and GSE62844.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, and three tables and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2014.11.033>.

AUTHOR CONTRIBUTIONS

J.P. designed the experiments, acquired and analyzed the data, and wrote the manuscript; P.S. performed the Rfx6 ChIP on Min6B1 cells; A.M. participated in the acquisition and analysis of the data; D.J.H. performed and analyzed the ex vivo calcium and ATP imaging experiments; T.Y. analyzed ChIP-seq data; C.K. analyzed RNA-seq data; M.C.B. generated and characterized Ins1-CreERT2 mice; and G.G. and G.A.R. designed the experiments, interpreted the data, and wrote the manuscript.

ACKNOWLEDGMENTS

We are grateful to Dr. S. Yoshida for the Ngn3-Cre mice; the Mouse Clinical Institute for the Ins1-CreER^{T2} mice; Christelle Thibault, Bernard Jost, and Muriel Philipps of the Microarray and Deep Sequencing platform of the IGBMC; Dr. C. Wright for Pdx1 antibody; and the Beta Cell Biology Consortium for the c-peptide antibodies. This study was supported by the Institut National de la Santé et de la Recherche Médicale (INSERM) and by grants from the Agence Nationale pour la Recherche (ANR 11 BSV1 003-01 Rfx-Panclnt) and the Fondation pour la Recherche Médicale (FRM DEQ20110421295) to G.G. J.P. and P.S. are recipients of a fellowship from the Ministère de la Recherche and the Fondation ARC pour la recherche sur le Cancer. D.J.H. is a Diabetes UK R.D. Lawrence Fellow. G.A.R. thanks the MRC (UK) for Programme grant MR/J0003042/1, the BBSRC (UK) for a project grant (BB/J015873/1), the Royal Society for a Wolfson Research Merit Award, and the Wellcome Trust for a Senior Investigator Award (WT098424AIA).

Received: July 31, 2014
 Revised: October 27, 2014
 Accepted: November 20, 2014
 Published: December 11, 2014

REFERENCES

- Aftab, S., Semene, L., Chu, J.S., and Chen, N. (2008). Identification and characterization of novel human tissue-specific RFX transcription factors. *BMC Evol. Biol.* **8**, 226.
- Ainscow, E.K., Zhao, C., and Rutter, G.A. (2000). Acute overexpression of lactate dehydrogenase-A perturbs beta-cell mitochondrial metabolism and insulin secretion. *Diabetes* **49**, 1149–1155.
- Ait-Lounis, A., Bonal, C., Seguin-Estévez, Q., Schmid, C.D., Bucher, P., Herrera, P.L., Durand, B., Meda, P., and Reith, W. (2010). The transcription factor Rfx3 regulates beta-cell differentiation, function, and glucokinase expression. *Diabetes* **59**, 1674–1685.
- Artuso, R., Provenzano, A., Mazzinghi, B., Giunti, L., Palazzo, V., Andreucci, E., Blasetti, A., Chiuri, R.M., Gianiorio, F.E., Mandich, P., et al. (2014). Therapeutic implications of novel mutations of the RFX6 gene associated with early-onset diabetes. *Pharmacogenomics J.*, Published online July 22, 2014 <http://dx.doi.org/10.1038/tpj.2014.37>.
- Blum, B., Hrvatin, S.S., Schuetz, C., Bonal, C., Reznia, A., and Melton, D.A. (2012). Functional beta-cell maturation is marked by an increased glucose threshold and by expression of urocortin 3. *Nat. Biotechnol.* **30**, 261–264.
- Butler, A.E., Janson, J., Bonner-Weir, S., Ritzel, R., Rizza, R.A., and Butler, P.C. (2003). Beta-cell deficit and increased beta-cell apoptosis in humans with type 2 diabetes. *Diabetes* **52**, 102–110.
- 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.
- Concepcion, J.P., Reh, C.S., Daniels, M., Liu, X., Paz, V.P., Ye, H., Highland, H.M., Hanis, C.L., and Greeley, S.A. (2014). Neonatal diabetes, gallbladder agenesis, duodenal atresia, and intestinal malrotation caused by a novel homozygous mutation in RFX6. *Pediatr. Diabetes* **15**, 67–72.
- Dayeh, T., Volkov, P., Salö, S., Hall, E., Nilsson, E., Olsson, A.H., Kirkpatrick, C.L., Wollheim, C.B., Eliasson, L., Rönn, T., et al. (2014). Genome-wide DNA methylation analysis of human pancreatic islets from type 2 diabetic and non-diabetic donors identifies candidate genes that influence insulin secretion. *PLoS Genet.* **10**, e1004160.
- Gao, T., McKenna, B., Li, C., Reichert, M., Nguyen, J., Singh, T., Yang, C., Pannikar, A., Doliba, N., Zhang, T., et al. (2014). Pdx1 maintains β cell identity and function by repressing an α cell program. *Cell Metab.* **19**, 259–271.
- Gu, C., Stein, G.H., Pan, N., Goebbels, S., Hörnberg, H., Nave, K.A., Herrera, P., White, P., Kaestner, K.H., Sussel, L., and Lee, J.E. (2010). Pancreatic beta cells require NeuroD to achieve and maintain functional maturity. *Cell Metab.* **11**, 298–310.
- Hart, A.W., Mella, S., Mendrychowski, J., van Heyningen, V., and Kleinjan, D.A. (2013). The developmental regulator Pax6 is essential for maintenance of islet cell function in the adult mouse pancreas. *PLoS ONE* **8**, e54173.
- Hodson, D.J., Mitchell, R.K., Bellomo, E.A., Sun, G., Vinet, L., Meda, P., Li, D., Li, W.H., Bugliani, M., Marchetti, P., et al. (2013). Lipotoxicity disrupts incretin-regulated human β cell connectivity. *J. Clin. Invest.* **123**, 4182–4194.
- Hodson, D.J., Mitchell, R.K., Marselli, L., Pullen, T.J., Gimeno Brias, S., Semplici, F., Everett, K.L., Cooper, D.M., Bugliani, M., Marchetti, P., et al. (2014a). ADCY5 couples glucose to insulin secretion in human islets. *Diabetes* **63**, 3009–3021.
- Hodson, D.J., Tarasov, A.I., Gimeno Brias, S., Mitchell, R.K., Johnston, N.R., Haghollahi, S., Cane, M.C., Bugliani, M., Marchetti, P., Bosco, D., et al. (2014b). Incretin-modulated beta cell energetics in intact islets of Langerhans. *Mol. Endocrinol.* **28**, 860–871.
- Ishihara, H., Wang, H., Drewes, L.R., and Wollheim, C.B. (1999). Overexpression of monocarboxylate transporter and lactate dehydrogenase alters insulin secretory responses to pyruvate and lactate in beta cells. *J. Clin. Invest.* **104**, 1621–1629.
- lynedjian, P.B. (1993). Mammalian glucokinase and its gene. *Biochem. J.* **293**, 1–13.
- Jing, X., Li, D.Q., Olofsson, C.S., Salehi, A., Surve, V.V., Caballero, J., Ivarsson, R., Lundquist, I., Pereverzev, A., Schneider, T., et al. (2005). Cav2.3 calcium channels control second-phase insulin release. *J. Clin. Invest.* **115**, 146–154.
- Jonas, J.C., Sharma, A., Hasenkamp, W., Ilkova, H., Patané, G., Laybutt, R., Bonner-Weir, S., and Weir, G.C. (1999). Chronic hyperglycemia triggers loss of pancreatic beta cell differentiation in an animal model of diabetes. *J. Biol. Chem.* **274**, 14112–14121.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25.
- Li, C., Chen, P., Vaughan, J., Lee, K.-F., and Vale, W. (2007). Urocortin 3 regulates glucose-stimulated insulin secretion and energy homeostasis. *Proc. Natl. Acad. Sci. USA* **104**, 4206–4211.
- Lilla, V., Webb, G., Rickenbach, K., Maturana, A., Steiner, D.F., Halban, P.A., and Irminger, J.C. (2003). Differential gene expression in well-regulated and dysregulated pancreatic beta-cell (MIN6) sublines. *Endocrinology* **144**, 1368–1379.
- Marselli, L., Thorne, J., Dahiya, S., Sgroi, D.C., Sharma, A., Bonner-Weir, S., Marchetti, P., and Weir, G.C. (2010). Gene expression profiles of Beta-cell enriched tissue obtained by laser capture microdissection from subjects with type 2 diabetes. *PLoS ONE* **5**, e11499.
- Marselli, L., Bugliani, M., Suleiman, M., Olimpico, F., Masini, M., Petrin, M., Boggi, U., Filippini, F., Syed, F., and Marchetti, P. (2013). β -Cell inflammation in human type 2 diabetes and the role of autophagy. *Diabetes Obes. Metab.* **15** (Suppl 3), 130–136.
- Matschinsky, F.M. (2005). Glucokinase, glucose homeostasis, and diabetes mellitus. *Curr. Diab. Rep.* **5**, 171–176.
- McTaggart, J.S., Clark, R.H., and Ashcroft, F.M. (2010). The role of the KATP channel in glucose homeostasis in health and disease: more than meets the islet. *J. Physiol.* **588**, 3201–3209.
- Mitchell, J., Punthakee, Z., Lo, B., Bernard, C., Chong, K., Newman, C., Cartier, L., Desilets, V., Cutz, E., Hansen, I.L., et al. (2004). Neonatal diabetes, with hypoplastic pancreas, intestinal atresia and gall bladder hypoplasia: search for the aetiology of a new autosomal recessive syndrome. *Diabetologia* **47**, 2160–2167.
- Nenquin, M., Szollosi, A., Aguilar-Bryan, L., Bryan, J., and Henquin, J.C. (2004). Both triggering and amplifying pathways contribute to fuel-induced insulin secretion in the absence of sulfonylurea receptor-1 in pancreatic beta-cells. *J. Biol. Chem.* **279**, 32316–32324.
- Otonkoski, T., Jiao, H., Kaminen-Ahola, N., Tapia-Paez, I., Ullah, M.S., Parton, L.E., Schuit, F., Quintens, R., Sipilä, I., Mayatepek, E., et al. (2007). Physical exercise-induced hypoglycemia caused by failed silencing of monocarboxylate transporter 1 in pancreatic beta cells. *Am. J. Hum. Genet.* **81**, 467–474.
- Pearl, E.J., Jarikij, Z., and Horb, M.E. (2011). Functional analysis of Rfx6 and mutant variants associated with neonatal diabetes. *Dev. Biol.* **351**, 135–145.
- Pullen, T.J., and Rutter, G.A. (2013). When less is more: the forbidden fruits of gene repression in the adult β -cell. *Diabetes Obes. Metab.* **15**, 503–512.
- Pullen, T.J., Khan, A.M., Barton, G., Butcher, S.A., Sun, G., and Rutter, G.A. (2010). Identification of genes selectively disallowed in the pancreatic islet. *Islets* **2**, 89–95.
- Pullen, T.J., da Silva Xavier, G., Kelsey, G., and Rutter, G.A. (2011). miR-29a and miR-29b contribute to pancreatic beta-cell-specific silencing of monocarboxylate transporter 1 (Mct1). *Mol. Cell. Biol.* **31**, 3182–3194.
- Pullen, T.J., Sylow, L., Sun, G., Halestrap, A.P., Richter, E.A., and Rutter, G.A. (2012). Overexpression of monocarboxylate transporter-1 (SLC16A1) in mouse pancreatic β -cells leads to relative hyperinsulinism during exercise. *Diabetes* **61**, 1719–1725.
- Quintens, R., Hendrickx, N., Lemaire, K., and Schuit, F. (2008). Why expression of some genes is disallowed in beta-cells. *Biochem. Soc. Trans.* **36**, 300–305.

- Rahier, J., Guiot, Y., Goebbels, R.M., Sempoux, C., and Henquin, J.C. (2008). Pancreatic beta-cell mass in European subjects with type 2 diabetes. *Diabetes Obes. Metab.* *10* (Suppl 4), 32–42.
- Rosengren, A.H., Braun, M., Mahdi, T., Andersson, S.A., Travers, M.E., Shiget, M., Zhang, E., Almgren, P., Ladenvall, C., Axelsson, A.S., et al. (2012). Reduced insulin exocytosis in human pancreatic β -cells with gene variants linked to type 2 diabetes. *Diabetes* *61*, 1726–1733.
- Rutter, G.A. (2004). Visualising insulin secretion. The Minkowski Lecture 2004. *Diabetologia* *47*, 1861–1872.
- Rutter, G.A. (2010). Think zinc: new roles for zinc in the control of insulin secretion. *Islets* *2*, 49–50.
- Rutter, G.A. (2014). Dorothy Hodgkin Lecture 2014. Understanding genes identified by genome-wide association studies for type 2 diabetes. *Diabet. Med.* *31*, 1480–1487.
- Schulla, V., Renström, E., Feil, R., Feil, S., Franklin, I., Gjinovci, A., Jing, X.J., Laux, D., Lundquist, I., Magnuson, M.A., et al. (2003). Impaired insulin secretion and glucose tolerance in beta cell-selective Ca(v)1.2 Ca²⁺ channel null mice. *EMBO J.* *22*, 3844–3854.
- Sekine, N., Cirulli, V., Regazzi, R., Brown, L.J., Gine, E., Tamarit-Rodriguez, J., Girotti, M., Marie, S., MacDonald, M.J., Wollheim, C.B., et al. (1994). Low lactate dehydrogenase and high mitochondrial glycerol phosphate dehydrogenase in pancreatic beta-cells. Potential role in nutrient sensing. *J. Biol. Chem.* *269*, 4895–4902.
- Smith, S.B., Qu, H.Q., Taleb, N., Kishimoto, N.Y., Scheel, D.W., Lu, Y., Patch, A.M., Grabs, R., Wang, J., Lynn, F.C., et al. (2010). Rfx6 directs islet formation and insulin production in mice and humans. *Nature* *463*, 775–780.
- Soyer, J., Flasse, L., Raffelsberger, W., Beucher, A., Orvain, C., Peers, B., Ravassard, P., Vermot, J., Voz, M.L., Mellitzer, G., and Gradwohl, G. (2010). Rfx6 is an Ngn3-dependent winged helix transcription factor required for pancreatic islet cell development. *Development* *137*, 203–212.
- Spiegel, R., Dobbie, A., Hartman, C., de Vries, L., Ellard, S., and Shalev, S.A. (2011). Clinical characterization of a newly described neonatal diabetes syndrome caused by RFX6 mutations. *Am. J. Med. Genet.* *155A*, 2821–2825.
- Talchai, C., Xuan, S., Lin, H.V., Sussel, L., and Accili, D. (2012). Pancreatic β cell dedifferentiation as a mechanism of diabetic β cell failure. *Cell* *150*, 1223–1234.
- Taneera, J., Lang, S., Sharma, A., Fadista, J., Zhou, Y., Ahlqvist, E., Jonsson, A., Lyssenko, V., Vikman, P., Hansson, O., et al. (2012). A systems genetics approach identifies genes and pathways for type 2 diabetes in human islets. *Cell Metab.* *16*, 122–134.
- Tarasov, A.I., Semplici, F., Ravier, M.A., Bellomo, E.A., Pullen, T.J., Gilon, P., Sekler, I., Rizzuto, R., and Rutter, G.A. (2012). The mitochondrial Ca²⁺ uniporter MCU is essential for glucose-induced ATP increases in pancreatic β -cells. *PLoS ONE* *7*, e39722.
- Taylor, B.L., Liu, F.-F., and Sander, M. (2013). Nkx6.1 is essential for maintaining the functional state of pancreatic beta cells. *Cell Rep.* *4*, 1262–1275.
- Terauchi, Y., Sakura, H., Yasuda, K., Iwamoto, K., Takahashi, N., Ito, K., Kasai, H., Suzuki, H., Ueda, O., Kamada, N., et al. (1995). Pancreatic beta-cell-specific targeted disruption of glucokinase gene. Diabetes mellitus due to defective insulin secretion to glucose. *J. Biol. Chem.* *270*, 30253–30256.
- Thorrez, L., Laudadio, I., Van Deun, K., Quintens, R., Hendrickx, N., Granvik, M., Lemaire, K., Schraenen, A., Van Lommel, L., Lehnert, S., et al. (2011). Tissue-specific disallowance of housekeeping genes: the other face of cell differentiation. *Genome Res.* *21*, 95–105.
- Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* *25*, 1105–1111.
- van Arensbergen, J., García-Hurtado, J., Moran, I., Maestro, M.A., Xu, X., Van de Castele, M., Skoudy, A.L., Palassini, M., Heimberg, H., and Ferrer, J. (2010). Derepression of Polycomb targets during pancreatic organogenesis allows insulin-producing beta-cells to adopt a neural gene activity program. *Genome Res.* *20*, 722–732.
- Yang, S.N., and Berggren, P.O. (2006). The role of voltage-gated calcium channels in pancreatic beta-cell physiology and pathophysiology. *Endocr. Rev.* *27*, 621–676.
- Yoshida, S., Takakura, A., Ohbo, K., Abe, K., Wakabayashi, J., Yamamoto, M., Suda, T., and Nabeshima, Y. (2004). Neurogenin3 delineates the earliest stages of spermatogenesis in the mouse testis. *Dev. Biol.* *269*, 447–458.
- Zhang, C., Moriguchi, T., Kajihara, M., Esaki, R., Harada, A., Shimohata, H., Oishi, H., Hamada, M., Morito, N., Hasegawa, K., et al. (2005). MafA is a key regulator of glucose-stimulated insulin secretion. *Mol. Cell. Biol.* *25*, 4969–4976.

Cell Reports, Volume 9

Supplemental Information

Rfx6 Maintains the Functional

Identity of Adult Pancreatic β Cells

**Julie Piccand, Perrine Strasser, David J. Hodson, Aline Meunier, Tao Ye, Céline Keime,
Marie-Christine Birling, Guy A. Rutter, and Gérard Gradwohl**

Supplemental Experimental Procedures

Mouse strains, genotyping and manipulations

Animal experiments were supervised by G. Gradwohl (agreement N°C67-59 approved by the Direction des Services Vétérinaires, Strasbourg, France) in compliance with the European legislation on care and use of laboratory animals. Mice were maintained on C57/B6J genetic background on a 12/12 hour light/dark cycle with unlimited access to water and standard rodent chow diet. Ngn3-Cre mice are a gift from Dr. Shosei Yoshida (Yoshida et al., 2004), Ins1-CreER^{T2} mice were generated at the ICS (Illkirch, France) in the context of the CreER Zoo. Rfx6 Floxed mice were generated in collaboration with the ICS. Genomic tail DNA was analyzed by PCR using the primers described below. Rfx6^{fl/fl}; Ngn3-Cre (Rfx6^{ΔEndo}) were used for analysis of the endocrine phenotype and both Rfx6^{fl/fl} or Rfx6^{fl/+} littermates were used as controls. Rfx6^{fl/fl}; Ins1-CreER^{T2} (Rfx6^{ΔBeta}) and Rfx6^{fl/fl} (control) 8-10 weeks old mice were both injected subcutaneously with 100μl (1mg) of tamoxifen (Sigma, dissolved in filtered corn oil at 10mg/ml) daily during 3 days. The mice were analysed 5 days, 3 weeks or 1 month after the first injection. Note that only males were used for metabolic experiments whereas a mixture of male and females was used for transcripts analysis.

Mouse Islet Purification

Mouse islet purification was performed as previously described (Piccand et al., 2014). Briefly, mice were euthanized and injected with Type V Collagenase (Sigma C9263) solution directly into the common bile duct to perfuse the pancreas. Pancreas was dissected out and digested and islets were handpicked after several purification steps and used for RNA isolation or kept in culture over night before insulin secretion tests.

Glucose tolerance tests

For Oral Glucose Tolerance Test (OGTT), 16-hours fasted 12 weeks-old males received glucose by intragastric gavage (2g/kg body weight of 15% D-glucose). For Intraperitoneal Glucose Tolerance Test (IPGTT), 16-hours fasted 13 weeks-old males received glucose by intraperitoneal injection (2g/kg body weight of 15% D-glucose). For both GTT, circulating blood glucose was measured in tail blood at 0, 5, 15, 30, 45, 60 and 120 minutes using Accu-Check Performa (Roche).

Real time PCR analysis

Total RNA from newborn pancreas or purified islets was extracted using TRI Reagent (Invitrogen). Eventual genomic DNA contamination was removed by DNase I digestion (Roche). cDNA was synthesized using random primers (Roche) and Transcriptor Reverse Transcriptase (Roche). Quantitative PCR was carried out (LightCycler 480, Roche) with either Taqman probes (Applied Biosystems) or UPL probes and primers (Roche) as listed below. Gene expression levels were normalised to Rplp0.

ChIPSeq and ChIP-qPCR in Min6B1 cells

Min6B1 cells were grown in DMEM with 25mM glucose supplemented with 15% foetal calf serum, penicillin/streptomycin and 71 μ M β -mercaptoethanol, in a 5% CO₂ incubator at 37°C. Cells were transfected with pCMV-Tag2A-3HA-Rfx6 plasmid using Lipofectamine 2000 (Invitrogen) and harvested two days later. For ChIP, after removing the medium, cells (transfected and untransfected Min6B1) were cross-linked for 10min at room temperature with 1% formaldehyde. The reaction was quenched with 125mM glycine and cells were washed with PBS and harvested by scrapping and centrifugation. Cells were resuspended in buffer I (25mM Tris HCl pH8; 1.5mM MgCl₂; 10mM KCl; 0.1% NP40 (Sigma-Aldrich); Protease inhibitor cocktail (Roche)) using a Dounce homogenizer. After centrifugation, the pellet was resuspended in buffer II (1%SDS ; 10mM EDTA ; 50mM Tris HCl pH8 Protease inhibitor cocktail (Roche)). Chromatin was then sonicated using a Covaris E210 (USA) to generate chromatin fragments of 500 bp. Cell debris were removed by centrifugation at 11000g for 10min. Each ChIP was performed with 50 μ g of chromatin (4.10⁶ cells) and a 10% aliquot was removed for the input control. Samples were pre-cleared with Protein-G-sepharose beads and next anti-HA antibody (12CA5, IGBMC) was added and incubated overnight at 4°C in buffer III (16.7mM Tris HCl pH8 ; 0.01% SDS ; 1.1% Triton ; 1.2mM EDTA ; 167mM NaCl ; Protease inhibitor cocktail (Roche)). The following day, Protein-G-sepharose beads were added during 4h. After washes in buffer IV and V (20mM Tris HCl pH8; 0.1% SDS; 1% Triton; 2mM EDTA; 150mM (IV) or 500mM (V) NaCl; Protease inhibitor cocktail (Roche)) and finally in Tris-EDTA. Bound DNA was eluted in elution buffer (1%SDS, 100mM NaHCO₃) and reverse-crosslinked by proteinase K treatment overnight at 65°C. DNA was purified the following day by standard phenol-chloroform extraction. Briefly ChIPSeq was performed using an HiSeq 2500 (Illumina) sequencer and peak detection was performed using the MACS

software (Zhang et al. 2008). Peaks were then annotated using GPAT on the mouse genome (mm9). Ref Seq genes, genomic features and corresponding coordinates were downloaded from the USCC genome browser. Selected Rfx6 targets identified by ChIPSeq were confirmed by ChIP-qPCR (LightCycler 480, Roche) to determine the relative enrichment. The later experiments were performed on non-transfected Min6B1 cells using anti-Rfx6 (2767, IGBMC), and preimmune (2767, IGBMC) sera and appropriate UPL probes and primers (Roche) as listed below. ChIPSeq Data have been deposited at the Gene Expression Omnibus depository (GEO, Accession number GSE62844).

SUPPLEMENTAL REFERENCES

Piccand, J., Meunier, A., Merle, C., Jia, Z., Barnier, J.V., and Gradwohl, G. (2014). Pak3 promotes cell cycle exit and differentiation of beta-cells in the embryonic pancreas and is necessary to maintain glucose homeostasis in adult mice. *Diabetes* 63, 203-215.

Yoshida, S., Takakura, A., Ohbo, K., Abe, K., Wakabayashi, J., Yamamoto, M., Suda, T., and Nabeshima, Y. (2004). Neurogenin3 delineates the earliest stages of spermatogenesis in the mouse testis. *Developmental biology* 269, 447-458.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*;9(9):R137.

Primary and Secondary Antibodies used in this study

Antibody Name	Vendor/Provider	Dilution
Guinea pig@Insulin	Linco (4011-01F)	1/1000
Mouse@Insulin	Sigma (I-2018)	1/1000
Guinea pig@Glucagon	Linco (4031-01F)	1/2000
Guinea pig@PP	Linco (4041-01F)	1/1000
Rat@Somatostatin	Chemicon (MAB354)	1/500
Goat@chromograninA	Santa Cruz (sc-1488)	1/500
Rabbit@Rfx6	IGBMC (2766)	1/500
Rabbit@MafA	Bethyl (A300-611A)	1/2000
Rabbit@Pdx1	Dr. Chris Wright (Vanderbilt University)	1/2000
Rabbit@Glut2	Chemicon (AB1342)	1/500
Rabbit@C-peptide 1	BCBC (AB1044)	1/2000
Rabbit@C-peptide 2	BCBC (AB1042)	1/3000
Rat@BrdU	AbD Serotec (OBT0030S)	1/10
Donkey@rabbitDyLight549	Jackson Immuno Research (711-505-152)	1/500
Donkey@gpDyLight488	Jackson Immuno Research (706-486-148)	1/500
Donkey@goatDyLight549	Jackson Immuno Research (705-505-147)	1/500
Donkey@ratDyLight488	Jackson Immuno Research (712-485-153)	1/500
Donkey@gpDyLight549	Jackson Immuno Research (706-505-148)	1/500
Donkey@mouseDyLight488	Jackson Immuno Research (715-485-150)	1/500
Donkey@rat biotinylated	Jackson Immuno Research (712-065-153)	1/500
Streptavidin-Alexa488	Molecular Probes (S-11223)	1/500

Primers or TaqMan and UPL probes used in this study

Gene	Forward primer or TaqMan ID	Reverse primer	UPL Probe	Application
Ngn3-Cre	ctgcagtttagcagaacttcagagggga	atcaacgttttgttttcgga	NA	Genotyping
ERT	gcattaccgggtcgatgcaacgagtgatgag	aggatctctagccaggcaca	NA	Genotyping
Rfx6 flox	gaaggtgcaccataaaaagc	tataagccaccagggtcag	NA	Genotyping
Cacnb2-xb	cctggttgttttgcctctagc	gaaggcagctggggaacta	#33	ChIP-PCR
Cacna1c-xb1	gcttgctgtctcctgagtttc	cattactgcatttcctagcaaacac	#67	ChIP-PCR
Cacna1c-xb2	gctcctggcttgctgctaac	tttggtgggaaagcagagat	#105	ChIP-PCR
Ldha-xb	tgcctctgtcagccatcag	aagcagaaaaagcaacaacga	#41	ChIP-PCR
Abcc8-xb	tgaagagaccctgggttttat	gtatgtatacaaccagcctggaaa	#42	ChIP-PCR
Gck-xb	ggtcaccatagaaccacagg	caaccaggtgtggagtagatgtc		ChIP-PCR
Rfx6	tgaggaaagagaaaactggag	ggaaattttggcgaattgtc	#83	RT-qPCR
MafA	ctccagagccagggtggag	gtacagggtcccgcctcttg	#10	RT-qPCR
Ucn3	gacctgagcatttcactcc	cagaagtggcagcaggaagt	#105	RT-qPCR
Cacna1d	gaagctgcttgaccaagttgt	aactccccacggttacctc	#9	RT-qPCR
Cacna1c	ccaacctcatctcttctca	acatagtctgcattgcctaggat	#71	RT-qPCR
Cacnb2	gcaggagagccagatgga	tcctggctcctttccatag	#67	RT-qPCR
Ldha	ggcactgacgcagacaag	tgatcacctcgtaggcactg	#12	RT-qPCR
ChgA	Mm00514341_m1		NA	RT-qPCR
Ins1	Mm01259683_g1		NA	RT-qPCR
Gck	Mm00439129_m1		NA	RT-qPCR
Abcc8	Mm00803450_m1		NA	RT-qPCR
Kcnj11	Mm00440050_s1		NA	RT-qPCR
Slc2a2	Mm00446228_m1		NA	RT-qPCR
Pcsk1	Mm00479023_m1		NA	RT-qPCR
Gcg	Mm00801712_m1		NA	RT-qPCR
Ppy	Mm00435889_m1		NA	RT-qPCR
Sst	Mm00436671_m1		NA	RT-qPCR
Arx	Mm00545903_m1		NA	RT-qPCR
Pax4	Mm01159036_m1		NA	RT-qPCR
Pdx1	Mm00435565_m1		NA	RT-qPCR
Nkx6.1	Mm00454962_m1		NA	RT-qPCR
Insm1	Mm02581025_s1		NA	RT-qPCR
MafB	Mm0062748_s1		NA	RT-qPCR
NeuroD1	Mm00520715_m1		NA	RT-qPCR
Pax6	Mm00443081_m1		NA	RT-qPCR
Ngn3	Mm00437606_s1		NA	RT-qPCR
Rplp0	Mm01974474_gH		NA	RT-qPCR

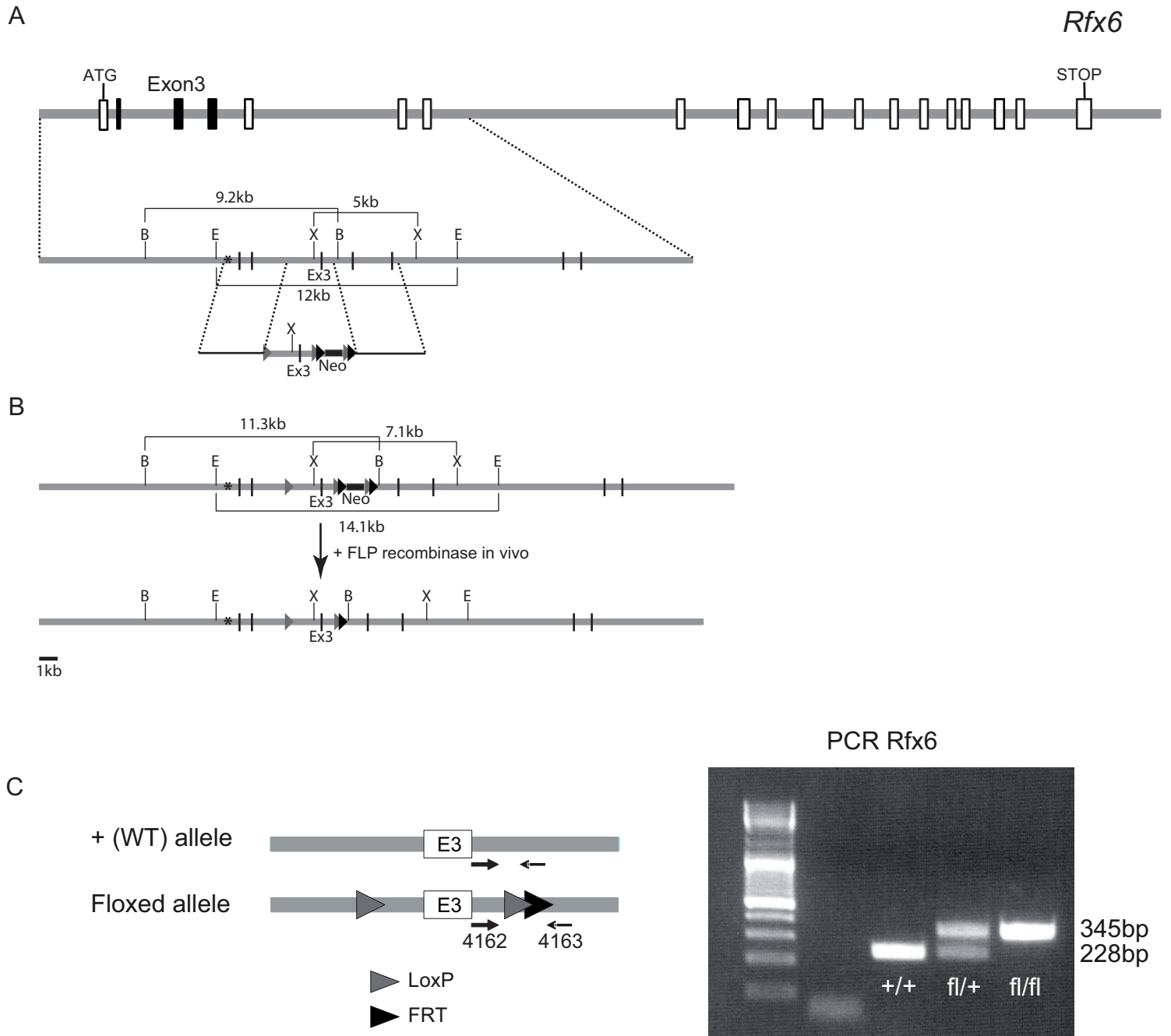


Figure S1: Generation of mice with a conditional *Rfx6* allele used in this study related to Experimental Procedures.

(A) Schema depicting the mouse *Rfx6* locus and targeting construct. The DNA binding domain is composed of the exons represented by black boxes. (B) Targeted *Rfx6* allele before and after the excision of the FRT flanked “Neo” selection cassette by the FLP recombinase. (A-B) Star indicates the position of the 5'- external probe used for Southern blot analysis. In B, the exons are represented as thick black bars. The Neomycin selection cassette is the black box, loxP and FRT sites are the grey and black triangles respectively. B: BsaI, E: EcoNI, X: XhoI. Scale is 1kb. (C) Genotyping strategy. Primers were designed to discriminate WT and floxed alleles (Supplemental Experimental Procedures). In the PCR gel on the right, the *Rfx6* floxed allele generates a band at 345bp and the WT allele at 228bp. Grey triangles represent loxP sites and the black triangle represents the remaining FRT site.

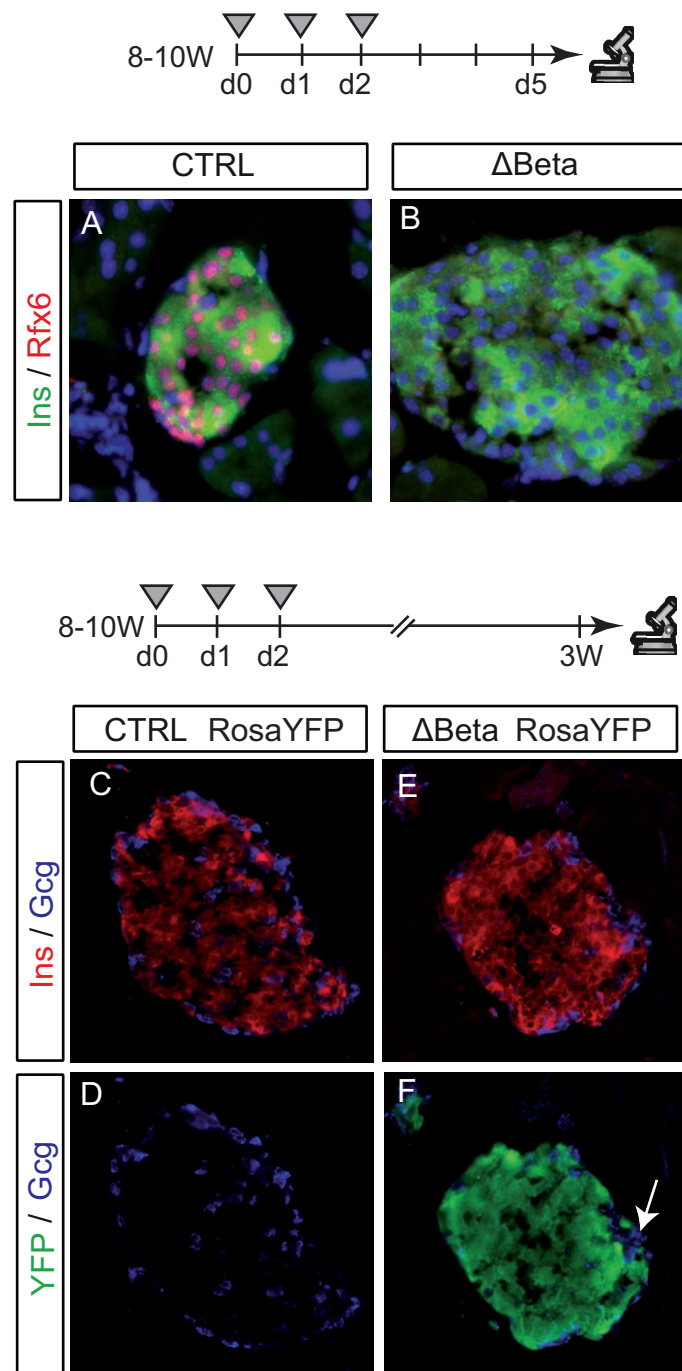


Figure S2: Data showing the efficient and specific deletion of Rfx6 in β -cells of Δ Beta mice upon tamoxifen injections related to Figure 2. (A-B) Rfx6 immunostaining is lost in insulin(Ins)-positive β -cells in Δ Beta mice ($Rfx6^{fl/fl}; Ins1-CreERT2$) 5 days after tamoxifen treatment in contrast to control mice ($Rfx6^{fl/fl}$). Grey triangles indicate the days of tamoxifen injections. (C-F) Tracing of the Cre activity of $Ins1-CreERT2$ mice in CTRL and Δ Beta mice crossed with the Cre reporter line $RosaYFP$. Strong YFP immunofluorescence in β -cells of Δ Beta; $RosaYFP$ mice (F) after tamoxifen treatment in contrast to control mice (D). Note that as expected α -cells (labeled by Glucagon, Gcg) are not marked by YFP (arrow in F).

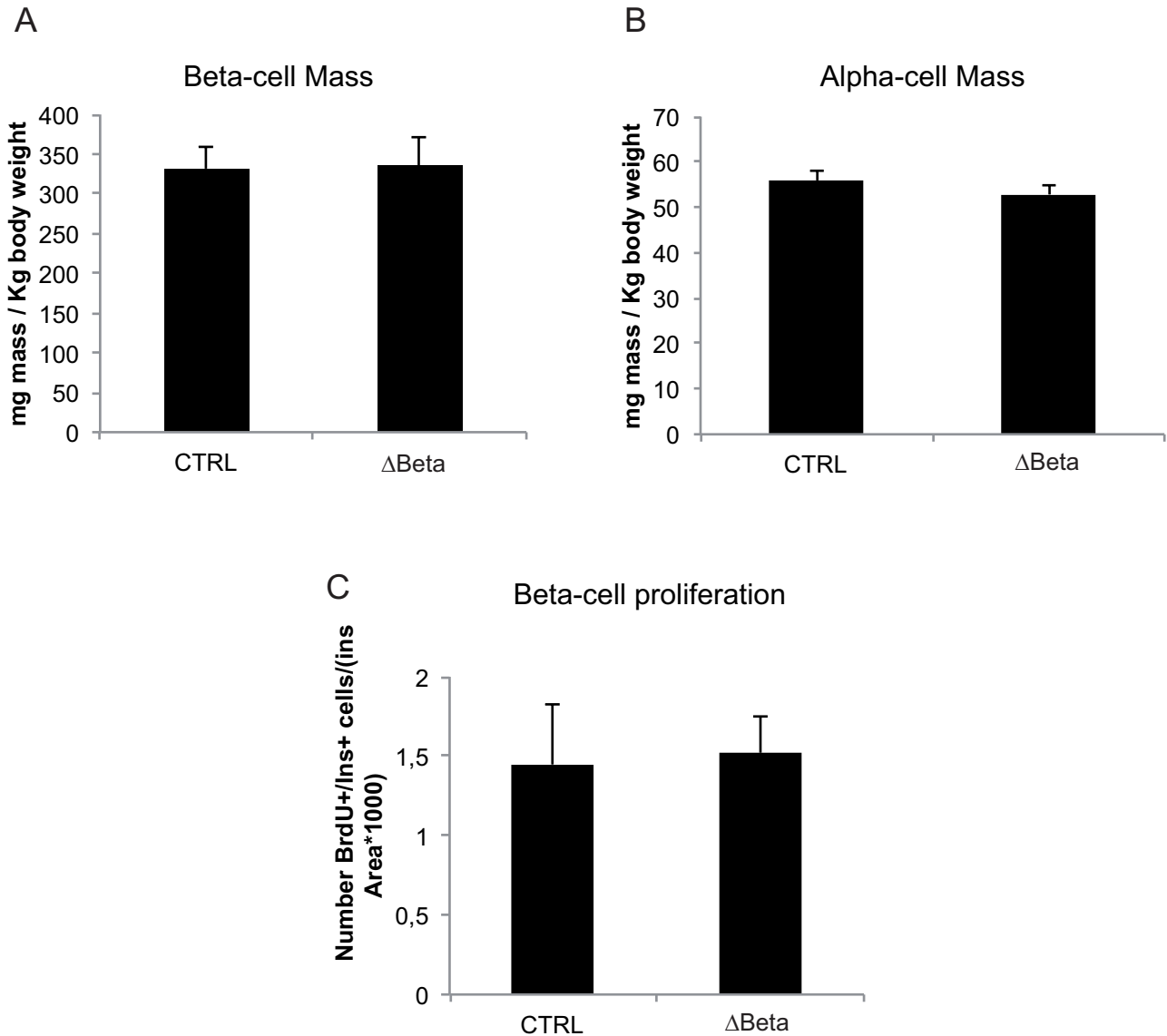
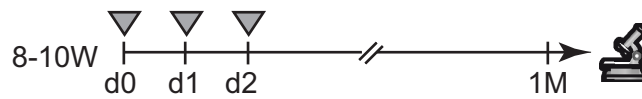


Figure S3: Data showing that alpha and beta cell mass as well as beta cell proliferation are not altered in adult Rfx6 Δ Beta mice one month after deletion of Rfx6 in beta cells, related to Figure 3.

Beta-cell mass (A) and alpha-cell mass (B) were quantified 1 month (M) after tamoxifen injections on n=4 control and Rfx6 Δ Beta adult mice from insulin and glucagon immunofluorescence area and pancreas weight. Beta cell mass = (Ins Area Fraction/100)*PancreasWeight(g)*1000)/(MouseWeight(g)/1000). Ins Area Fraction = (Ins Area) *100 / (Dapi Area). Alpha cell mass was calculated the same way except that Gcg area was taken into account instead of Ins area. C) Controls and Rfx6 Δ Beta mice were injected with BrdU 24h, before analyses and the number of Ins-positive/BrdU-positive cells was counted on n=4 animals 1 month after tamoxifen injections and the results normalized relative to the total Insulin area. Grey triangles indicate the days of tamoxifen injections.

	Gene name	Average expression CTRL	Average Expression Rfx6 Δ beta	FC	p-value
Rfx family members	Rfx1	60.45	44.22	-1.35	ns
	Rfx2	52.94	38.34	-1.35	ns
	Rfx3	1333.66	953.66	-1.37	ns
	Rfx4	1.20	1.77	1.28	ns
	Rfx5	659.73	454.97	-1.44	1.48E-03
	Rfx6*	1533.27	675.14	-2.23	2.40E-13
	Rfx7	130.23	118.77	-1.09	ns
	Rfx8	9.89	8.11	-1.18	ns
Beta cell function/transcription	Ins1	937814.39	498990.79	-1.85	2.74E-07
	Ins2	1914938.84	1896566.74	-1.01	ns
	Slc2a2	13015.56	9598.95	-1.34	4.43E-02
	Gck	2130.97	1344.81	-1.57	8.11E-05
	Abcc8	6540.14	4579.93	-1.41	1.31E-02
	Kcnj11	1197.27	1056.54	-1.13	ns
	Pcsk1	6639.56	7178.24	1.08	ns
	Pcsk2	26677.09	24193.81	-1.10	ns
	Slc30a8	19055.92	14614.63	-1.30	8.90E-03
	Chga	149119.97	129278.82	-1.15	ns
	Iapp	766631.71	678124.88	-1.13	ns
	Syp	3647.62	4329.68	1.18	ns
	Syt4	6151.99	5810.84	-1.06	ns
	Ucn3	3889.29	1769.32	-2.12	2.45E-06
	Neurod1	1764.47	971.99	-1.80	7.27E-09
	Pdx1	840.71	623.79	-1.33	ns
	Nkx6-1	1805.28	1630.66	-1.10	ns
	Isl1	3906.73	2125.09	-1.81	1.54E-06
	Pax6	6600.11	3210.75	-2.03	7.73E-14
	Insm1	1043.29	856.28	-1.21	ns
	Mafa	2701.13	1948.29	-1.33	ns
	Mafb	271.77	202.74	-1.33	ns
	Mnx1	225.32	134.05	-1.66	2.01E-05
	Myt1	1994.96	1037.75	-1.87	8.57E-05
	Neurog3	41.81	29.20	-1.33	ns
	Nkx2-2	1110.44	878.78	-1.23	ns
	Sox4	2551.92	2895.45	1.12	ns
	Foxa2	1550.68	1341.41	-1.15	ns

Table S1 : Expression of selected genes in Rfx6 Δ beta islets related to Figure 2 and 4
RNA Seq data on control and Rfx6 Δ beta mouse islets (n=3), 5 days after Tamoxifen treatment. FC: mutant /controls (CTRL); ns: non significant p value > 0.05; Average expression: normalized read counts divided by gene length. * Only exon-3 of Rfx6 is deleted in Rfx6 Δ beta mouse islets. Remaining exons (Rfx6 has 19 exons) can be transcribed and thus identified as “Rfx6” mRNA in the RNA sequencing experiments.

Gene name	Description	Fold Change Mutant/ctrl	p-value	Average expression (controls)	Average expression (mutants)
Cacna1e	calcium channel, voltage-dependent, R type, alpha 1E subunit	-4.90	4,89E-09	46.15	7.09
Cacnb2	calcium channel, voltage-dependent, beta 2 subunit	-2.10	3,79E-08	876.24	407.18
Cacna1c	calcium channel, voltage-dependent, L type, alpha 1C subunit	-2.17	1,54E-05	337.79	148.65
Cacna1b	calcium channel, voltage-dependent, N type, alpha 1B subunit	-1.82	6,28E-04	101.02	53.66
Cacna1a	calcium channel, voltage-dependent, P/Q type, alpha 1A subunit	-1.82	1,87E-03	2153.91	1129.47
Cacna2d3	calcium channel, voltage-dependent, alpha2/delta subunit 3	-2.10	2,78E-03	64.74	27.72
Cacna1d	calcium channel, voltage-dependent, L type, alpha 1D subunit	-1.66	3,52E-03	490.83	287.35
Cacnb1	calcium channel, voltage-dependent, beta 1 subunit	-1.54	7,14E-03	184.25	116.99
Cacnb3	calcium channel, voltage-dependent, beta 3 subunit	-1.15	ns	445.49	387.15
Cacna2d1	calcium channel, voltage-dependent, alpha2/delta subunit 1	-1.17	ns	1232.37	1051.85
Cacng8	calcium channel, voltage-dependent, gamma subunit 8	1.51	ns	0.88	4.36
Cacna1s	calcium channel, voltage-dependent, L type, alpha 1S subunit	1.48	ns	0.64	1.23
Cacna1h	calcium channel, voltage-dependent, T type, alpha 1H subunit	-1.14	ns	205.02	178.96
Cacng1	calcium channel, voltage-dependent, gamma subunit 1	1.24	ns	0.42	1.51
Cacna1i	calcium channel, voltage-dependent, alpha 1I subunit	-1.24	ns	102.85	77.25
Cacng2	calcium channel, voltage-dependent, gamma subunit 2	-1.25	ns	3.05	2.24
Cacnb4	calcium channel, voltage-dependent, beta 4 subunit	1.23	ns	1.35	1.73
Cacng4	calcium channel, voltage-dependent, gamma subunit 4	-1.20	ns	36.70	29.48
Cacna2d2	calcium channel, voltage-dependent, alpha 2/delta subunit 2	-1.18	ns	17.75	14.30
Cacna1g	calcium channel, voltage-dependent, T type, alpha 1G subunit	-1.09	ns	6.20	5.57
Cacng7	calcium channel, voltage-dependent, gamma subunit 7	1.13	ns	6.11	6.94
Cacna2d4	calcium channel, voltage-dependent, alpha 2/delta subunit 4	1.06	ns	5.23	5.67
Cacna1f	calcium channel, voltage-dependent, alpha 1F subunit	-1.05	ns	9.07	8.54
Cacng5	calcium channel, voltage-dependent, gamma subunit 5	-1.01	ns	1.88	1.84
Cacng3	calcium channel, voltage-dependent, gamma subunit 3	1.01	ns	1.92	1.90

Table S2: Down-regulation of voltage-dependent calcium channel (VDCC) genes in beta cells lacking Rfx6, related to Figure 5.

Table showing the expression (RNASeq) of voltage-dependent calcium channel genes in control (ctrl) and Rfx6 Δ beta mouse islets 5 days after Tamoxifen treatment. FC: mutant /control, ns: non specific p<0.05 . Average expression: normalized read counts divided by gene length.

Gene name	Description	Fold Change	Average expression controls	Average expression RfX6 Δ beta
MyI9	myosin, light polypeptide 9, regulatory	5.73	311.40	2431.82
Ogn #	Osteoglycin	5.54	16.63	148.53
Ly6c2	lymphocyte antigen 6 complex, locus C2	5.43	2.56	69.89
Islr #	immunoglobulin superfamily containing leucine-rich repeat	5.36	5.53	45.81
Tgm2	transglutaminase 2, C polypeptide	5.18	254.50	1972.31
Pdgfra (1)	platelet derived growth factor receptor, alpha polypeptide	5.13	20.74	234.54
Gas1 #	growth arrest specific 1	4.76	14.30	151.42
Gda	guanine deaminase	4.72	91.06	873.70
Pcolce #	procollagen C-endopeptidase enhancer protein	4.47	69.00	445.47
Igfbp4 (1) #	insulin-like growth factor binding protein 4	4.15	373.64	2227.03
Mylk #	myosin, light polypeptide kinase	3.99	65.44	308.06
Cd302 (1)	CD302 antigen	3.85	28.43	161.11
Mgst1 #	microsomal glutathione S-transferase 1	3.85	125.96	748.69
Cxcl12 (1)	chemokine (C-X-C motif) ligand 12	3.76	49.09	229.09
Arhgdib #	Rho, GDP dissociation inhibitor (GDI) beta	3.70	51.28	309.85
Itih5	inter-alpha (globulin) inhibitor H5	3.58	14.48	73.07
Tst #	thiosulfate sulfurtransferase, mitochondrial	3.56	42.63	184.24
Slc16a1 (1) #	solute carrier family 16 (monocarboxylic acid transporters), member 1	3.51	16.55	85.38
Fxyd1 #	FXDY domain-containing ion transport regulator 1	3.39	35.07	152.43
Ldha (1) #	lactate dehydrogenase A	3.11	657.33	2693.39
Fcgrt #	Fc receptor, IgG, alpha chain transporter	3.11	76.96	333.51
Maf #	avian musculoaponeurotic fibrosarcoma (v-maf) AS42 oncogene homolog	3.06	15.36	79.90
Sult1a1	sulfotransferase family 1A, phenol-preferring, member 1	3.02	83.74	356.79
Ddah1 #	dimethylarginine dimethylaminohydrolase 1	2.94	6.70	24.03
Gas6 #	growth arrest specific 6	2.91	228.93	793.31
Selenbp1	selenium binding protein 1	2.80	35.91	129.27
Car2 #	carbonic anhydrase 2	2.64	11.34	43.97
Gucy1a3 #	guanylate cyclase 1, soluble, alpha 3	2.51	55.33	149.02
Cat #	catalase	2.50	153.94	428.03
Ndrp2	N-myc downstream regulated gene 2	2.48	45.05	121.75
Nfib #	nuclear factor I/B	2.45	133.27	345.35
Yap1	yes-associated protein 1	2.43	124.39	348.35
Lmo4 #	LIM domain only 4	2.40	59.70	161.98
Tns1 #	tensin 1	2.31	71.00	168.93
Ly6a	lymphocyte antigen 6 complex, locus A	2.23	332.03	833.66
Smad3 #	MAD homolog 3 (Drosophila)	2.22	346.47	795.26
Hbb-b1	hemoglobin, beta adult major chain	2.21	426.06	1331.47
Hbb-b2	hemoglobin, beta adult minor chain	2.20	73.60	234.06
Parp3	poly (ADP-ribose) polymerase family, member 3	2.15	47.28	115.77
Fgf1 #	fibroblast growth factor 1	2.12	601.17	1325.74
Acot7 #	acyl-CoA thioesterase 7	2.03	73.13	161.04
Zfp3611 #	zinc finger protein 36, C3H type-like 1	2.03	41.87	91.95
Oat (1) #	ornithine aminotransferase	2.03	256.77	576.58
Galm #	galactose mutarotase	1.94	19.39	39.90
Ly6c1	lymphocyte antigen 6 complex, locus C1	1.81	326.64	626.69
Ak3	adenylate kinase 3	1.78	635.52	1161.10

Mgll #	monoglyceride lipase	1.70	462.81	831.93
Bloc1s1 #	biogenesis of lysosome-related organelles complex-1, subunit 1	1.70	121.03	215.11
Nola2 (Nhp2) #	NHP2 ribonucleoprotein homolog (yeast)	1.64	175.93	299.91
Plec #	plectin	1.50	190.71	286.16
Zyx	zyxin	1.40	57.71	82.45
C1qbp	complement component 1, q subcomponent binding protein	1.43	282.95	406.83
Uqcrb	ubiquinol-cytochrome c reductase binding protein	1.34	1598.62	2152.23
Ndrg4	N-myc downstream regulated gene 4	1.22	761.90	935.42
Cox5a	cytochrome c oxidase, subunit Va	ns	748.60	1008.20
Criz1 (Utp3) #	small subunit processome component, homolog (S. cerevisiae)	ns	365.34	407.60
Higd1a #	HIG1 domain family, member 1A	ns	110.53	126.48
Rpl24 #	ribosomal protein L24	ns	3565.55	3452.93
Rpl36 #	ribosomal protein L36	ns	151.04	144.09
Hba-a1	hemoglobin alpha, adult chain 1	ns	94.77	207.23
Hba-a2 #	hemoglobin alpha, adult chain 2	ns	18.56	87.10
Abca8a	ATP-binding cassette, sub-family A (ABC1), member 8a	ns	193.17	296.45
Mgst3 #	microsomal glutathione S-transferase 3	ns	215.97	306.02
Gsta4 #	glutathione S-transferase, alpha 4	ns	65.32	104.21
Fam59a #	family with sequence similarity 59, member A	ns	15.55	19.85
Aspa	Aspartoacylase	ns	33.72	30.19
Zdhhc9	zinc finger, DHHC domain containing 9	ns	45.10	51.59

Table S3: Rfx6 represses disallowed genes (data related to Figure 6)

Disallowed genes as defined by Pullen et al Islet 2010 and Thorrez et al Genome Research 2010 are up-regulated when Rfx6 is deleted in adult beta-cells (RNA Seq data on control and Δ beta mouse islets 5 days after Tamoxifen treatment. FC: mutant/controls, ns, non significant p value > 0.05. Average expression: normalized read counts divided by gene length. (1) disallowed genes identified in both studies. # potential direct target of Rfx6 based on the identification of a peak in the vicinity of these genes in Rfx6 ChIPSeq experiments performed in Min6B1 cells (see Methods).

A.4. Recherche des gènes cibles de RFX6 dans les ilots de Langerhans de souris

A.4.a. Stratégie expérimentale

L'identification des cibles régulées de RFX6 dans les cellules Min6b1 nous a permis de mettre en évidence un premier répertoire de gènes permettant de prédire son rôle dans les cellules beta pancréatiques. La réussite des expériences de CHIP Seq dans les Min6b1 à l'aide de l'anticorps anti-RFX6 a de ce fait permis de valider ce dernier pour cette technique. Cependant, bien que possédant la caractéristique principale des cellules beta de sécrétion d'insuline en réponse au glucose, les Min6b1 restent un modèle cellulaire différent de ce qui a lieu *in vivo* dans les ilots de Langerhans de souris. RFX6 régulerait ainsi des cibles spécifiques dans les ilots pancréatiques *in vivo*. De plus, ces ilots sont composés de plusieurs types cellulaires, dont les cellules beta sécrétrices d'insuline, qui expriment tous Rfx6 au stade adulte (Smith et al, 2010; Soyer et al, 2010). Par conséquent, les expériences de CHIP Seq dans les ilots pancréatiques permettront d'identifier les cibles de Rfx6 non seulement dans les cellules beta mais aussi dans les cellules alpha, delta, PP et epsilon.

La détermination des cibles de RFX6 *in vivo* à partir d'ilots de Langerhans de souris purifiés a donc été mise en place selon la procédure décrite dans le paragraphe Matériels et méthodes B.2. Pour cela, différentes quantités de matériel de départ ont été testées à partir de 30 à 70 souris (soit un équivalent de 5000 ou 10000 ilots par CHIP) ainsi que différentes concentrations en anticorps anti-RFX6 2767 (1/660^e ou 1/200^e) pour l'IP.

A.4.b. Identification et analyse des sites de liaison de RFX6 dans les ilots pancréatiques de souris

La mise au point du CHIP RFX6 dans les ilots pancréatiques de souris a pu être finalisée pour obtenir un nombre de séquences suffisamment élevé (25 millions de séquences minimum constitue le seuil de la Plateforme de Biopuces et Séquençage de l'IGBMC) pour nous permettre une analyse globale à partir de 10000 ilots et une concentration de 1/200^e de anti-RFX6 2767 pour un CHIP. Les résultats obtenus pour cette expérience sont résumés dans le **Tableau 9**. Une quantité d'ADN plus faible que la limite demandée par la Plateforme de Biopuces et Séquençage de l'IGBMC (10ng) a été produite après préparation de la bibliothèque d'ADN immunoprécipité, cependant après réalisation du séquençage à partir des 2,4ng d'ADN produits, les données obtenues ont permis l'identification de 10143 pics.

Les sites de liaisons de RFX6 dans les ilots pancréatiques sont regroupés autour des sites d'initiation de la transcription des gènes (TSS) (**Figure 40A**) et, pour 32 % d'entre eux,

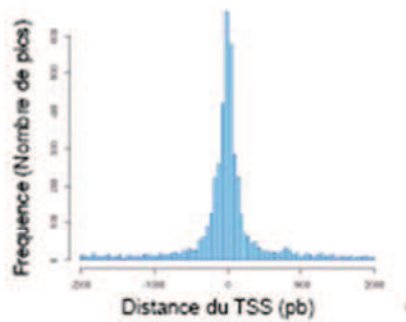
situés dans les régions promotrices et au niveau des TSS. 60% des pics sont situés à part égale dans les régions intergéniques ou introniques du génome (**Figure 40B**). Un motif de type « RFX » est retrouvé de manière enrichie sous 21% de la totalité des pics de ChIP Seq analysés (**Figure 40C**). Deux autres motifs de type NRF1 et CTCF/INSM1 sont également retrouvés dans respectivement 10 et 15% des pics (**Figure 40D**).

La protéine CTCF est un facteur de transcription impliqué dans de nombreux processus cellulaires, notamment la régulation de la transcription, et la régulation de l'architecture de la chromatine (Kim et al, 2015). En 2014, Fang et al montrent que lorsque Ctf est surexprimé chez la souris, l'expression de Pax6, qui joue un rôle clé dans le contrôle de la différenciation et de la fonction de la cellule beta, diminue, les souris présentent une homéostasie du glucose perturbée et souffrent d'hypoglycémie. CTCF semble donc impliquée dans la régulation de la fonction de la cellule beta (Fang et al, 2014). De plus, le motif consensus CTCF retrouvé dans 15 % des cibles est également associé au motif de fixation du facteur de transcription INSM1 dont la fonction et le rôle dans la différenciation endocrine et la fonction des cellules pancréatiques sont établis. En effet, dans les cellules endocrines du pancréas, l'expression d'Insm1 est notamment sous le contrôle de NGN3 pendant le développement et son expression est maintenue chez la souris adulte. L'absence d'Insm1 chez la souris provoque une sécrétion d'insuline en réponse au glucose perturbée (Jia et al, 2015; Lan & Breslin, 2009; Mellitzer et al, 2006; Osipovich et al, 2014). L'identification de ces motifs de fixation de CTCF et INSM1 dans les sites ciblés par RFX6 suggèrent une éventuelle action commune de ces facteurs de transcription et de RFX6 dans la régulation de la fonction de la cellule beta.

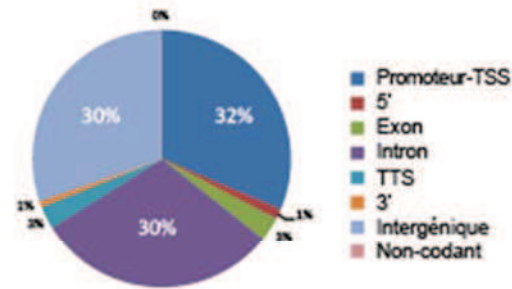
Jeu de données	Matériel biologique	Contrôle	Taille moyenne de l'ADN (pb)	Quantité totale d'ADN après préparation de la librairie (en ng)	Nombre de séquences	Nombre de séquences alignées	Séquences uniques	Nombre de pics identifiés
ChIP Rfx6	ilots pancréatiques de souris purifiés	input	355	2.4	68100395	49518093	1944586	10143

Tableau 9 : Résumé de l'expérience de ChIP Seq réalisée dans les ilots pancréatiques de souris.

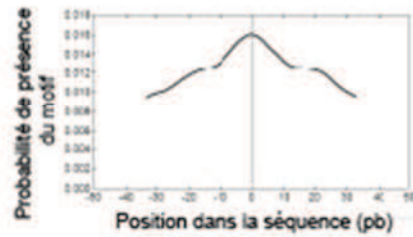
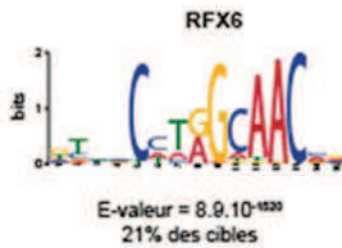
A Distribution des pics autour du TSS des gènes



B Régions liées par RFX6



C Analyse de motif *de novo*



Position	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
%A	11	9	16	17	0	29	16	46	3	10	93	99	0	36	27
%C	10	21	26	40	97	50	11	0	0	81	0	0	100	39	19
%G	43	9	20	17	0	1	1	54	97	3	3	1	0	13	42
%T	36	61	38	26	3	21	71	0	0	6	4	0	0	13	12

D Analyse de motif *de novo*

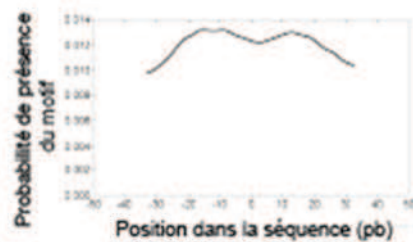
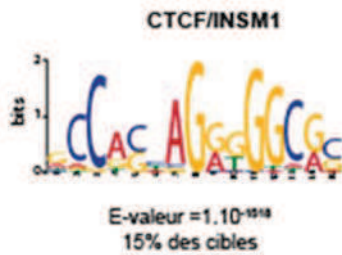
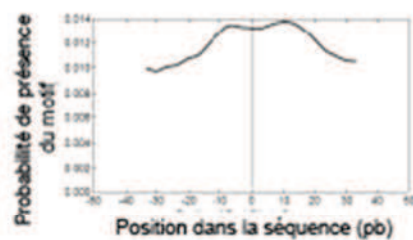
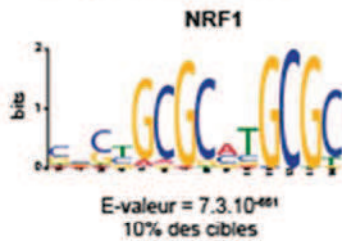
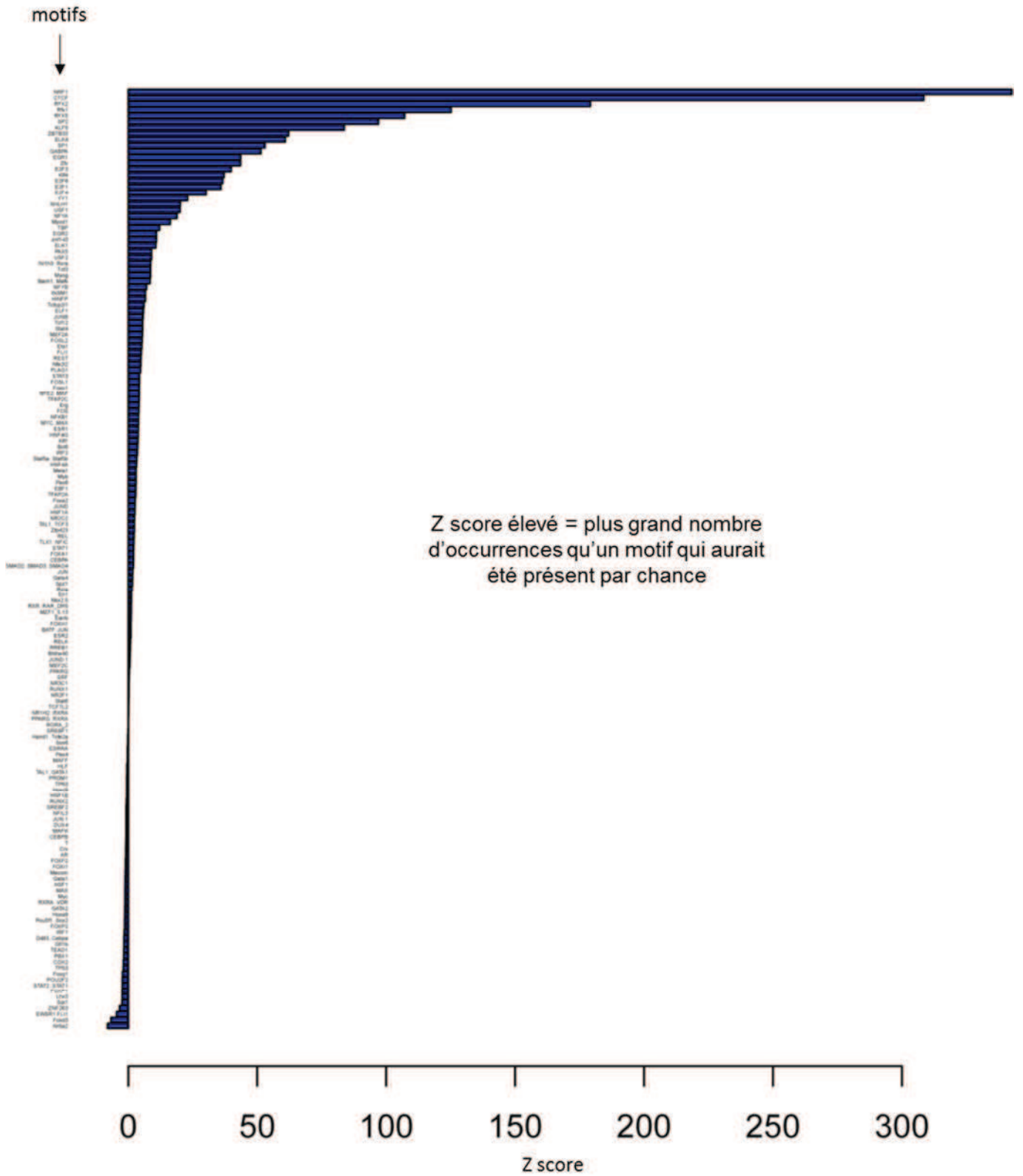


Figure 40 : 10143 sites de liaison de RFX6 sont identifiés dans les ilots pancréatiques de souris : (A) Distribution des sites de fixation de RFX6 autour des sites d'initiation de la transcription (TSS) des gènes. (B) Distribution des sites de fixation de RFX6 au sein du génome (TSS = site d'initiation de la transcription, TTS = site de terminaison de la transcription, 5' = région 5'UTR, 3' = région 3'UTR, non-codant = gènes non-codant). (C et D) Analyse de motif *de novo* dans les 100pb centrées sous le sommet des 10143 sites de fixation de RFX6 dans les ilots pancréatiques. Cette analyse révèle un motif consensus de fixation de RFX6 centralement enrichi sous 21% des pics de ChIP Seq, le pourcentage de présence de chaque base à chaque position est indiqué dans le tableau (C). Deux autres motifs consensus de type NRF1 et CTCF/INSM1 sont également enrichis sous les pics (D).

A.4.c. Analyse d'occurrence et de co-occurrence de motifs dans les cibles de RFX6 dans les ilots pancréatiques de souris

Dans le but d'identifier quels motifs connus (base de données Jaspar 2014) sont présents dans les séquences sous les 10143 pics de ChIP Seq anti-RFX6 (200pb centrées sous les sommets des pics) dans les ilots pancréatiques, une analyse d'occurrence de motifs a été mise en place (**paragraphe Matériels et méthodes C.4**) en association avec Stéphanie Legras, doctorante, de la plateforme de biopuces et séquençage de l'IGBMC. Des protéines ou cofacteurs agissant de concert avec RFX6 pourront ainsi être identifiés s'ils existent. Les résultats obtenus sont présentés en **Figure 41**. La présence ou occurrence des motifs est dénombrée et un Z score représentant la présence du motif plus ou moins élevée par rapport à ce qui aurait été obtenu par chance est calculé (**Figure 41A**). Les motifs occurrents de manière significative dans les cibles de RFX6 dans les ilots pancréatiques sont indiqués en **Figure 41B** et un histogramme est présenté pour les 10 motifs les plus occurrents (**Figure 41C**). Comme attendu et suggéré par l'analyse de motif *de novo* qui avait été réalisé sur le même jeu de données (**paragraphe Résultats A.4.b**), les motifs les plus occurrents sont les motifs de type « NRF1 », « CTCF » et « RFX ». L'occurrence significative des motifs qui a été décrite dans l'analyse réalisée dans les gènes ciblés dans les cellules Min6b1 et régulés dans les cellules beta pancréatiques par RFX6 (**paragraphe Résultats A.3.d**) est également retrouvée dans les gènes directement régulés par RFX6 dans les ilots pancréatiques pour les motifs Sp1 et 2, et E2F1 ou encore Insm1.

A



B

Motifs	Zscore	Nombre d'occurrence
NRF1	342.65	1089
CTCF	308.43	1500
RFX2	179.15	658
Rfx1	125.11	521
RFX5	107.13	294
SP2	96.99	960
KLF5	83.75	467
ZBTB33	62.07	65
ELK4	60.84	69
SP1	52.95	586
GABPA	51.36	61
EGR1	43.52	546
Zfx	43.47	130
E2F3	39.85	154
Klf4	37.12	139
E2F6	36.57	77
E2F1	35.90	95
E2F4	30.13	55
YY1	22.93	99
NHLH1	20.10	42
USF1	19.97	46
NFYA	18.83	40
Myod1	16.26	60
TBP	12.07	21
EGR2	10.91	150
znf143	10.87	53
ELK1	10.63	14
PAX5	8.98	20
USF2	8.93	28
Nr1h3..Rxra	8.62	26
Tcf3	8.57	50
Myog	8.55	31
Bach1..Mafk	8.39	36
NFYB	7.02	27
INSM1	6.60	18
HINFP	6.58	2
Tcfcp2l1	6.03	19
ELF1	5.82	37

Motifs	Zscore	Nombre d'occurrence
JUNB	5.70	18
Tcf12	5.50	29
Stat4	5.48	23
MEF2A	5.43	18
FOSL2	5.33	15
Ets1	5.17	39
FLI1	5.00	28
REST	4.86	24
Nfe2l2	4.79	25
PLAG1	4.68	26
STAT3	4.31	32
FOSL1	4.30	18
Foxo1	4.24	16
NFE2..MAF	4.21	23
TFAP2C	4.19	38
Erg	4.08	22
FOS	4.04	13
NFKB1	3.94	12
MYC..MAX	3.91	12
ESR1	3.83	27
HNF4G	3.78	25
Klf1	3.72	29
Bcl6	3.61	22
IRF2	3.49	12
Stat5a..Stat5b	3.26	19
HNF4A	3.15	23
Meis1	2.99	22
Myb	2.94	10
Pax6	2.94	4
EBF1	2.90	25
TFAP2A	2.86	22
Foxa2	2.63	12
JUND	2.56	10
HNF1A	2.52	6
NR2C2	2.40	30
TAL1..TCF3	2.34	9
Zfp423	2.31	13

C

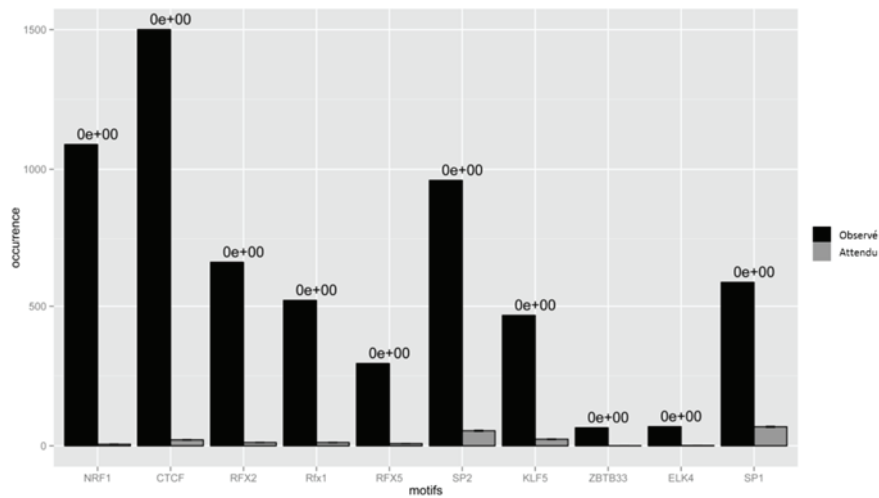


Figure 41 : Analyse d'occurrence de motifs dans les cibles de RFX6 dans les ilots pancréatiques. Les séquences sous les 10143 pics de CHIP Seq RFX6 dans les ilots pancréatiques sont analysées avec la base de données Jaspar 2014. **(A)** Histogramme où un Z score élevé signifie que l'occurrence du motif est plus fréquente que ce qui aurait été obtenu par chance, inversement un Z score peu élevé signifie que l'occurrence du motif est moins fréquente que ce qui aurait été obtenu par chance. **(B)** Les motifs sont classés par ordre décroissant de Z score et sont associés à leur nombre d'occurrences. Les motifs dont l'occurrence est significative sont présentés. **(C)** Histogramme d'occurrence des motifs. Le nombre d'occurrences attendu et observé est représenté pour les 10 motifs les plus occurants.

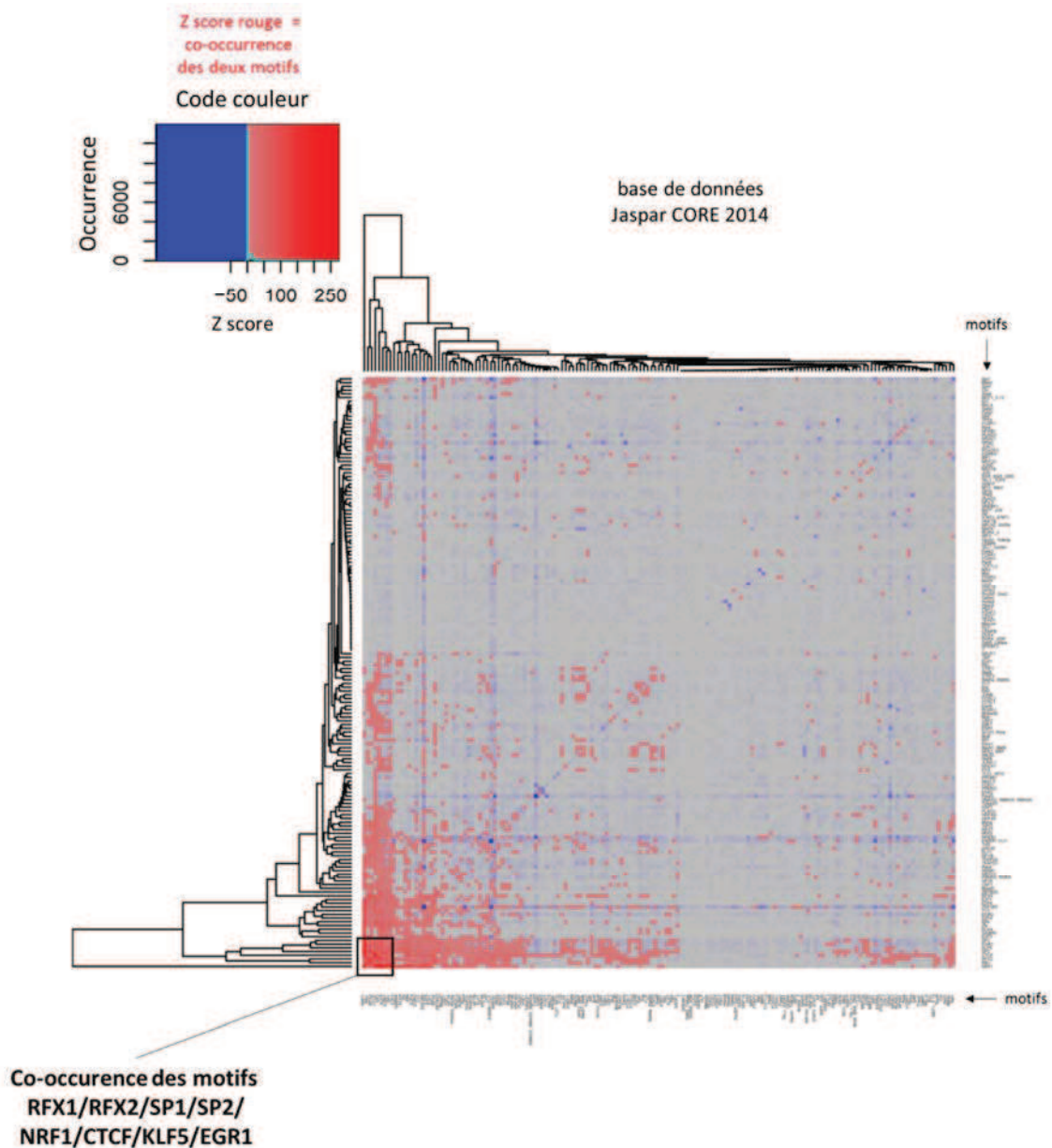
D'autre part, une analyse de co-occurrence de motifs a été réalisée et consiste à vérifier à quelle répétition deux motifs sont présents dans la même séquence. Les motifs dont la co-occurrence est la plus élevée avec les motifs de type « RFX » sont listés en **Figure 42A**. On retrouve dans cette liste les motifs les plus occurants identifiés dans l'analyse précédente. Le nombre de co-occurrence est particulièrement élevé pour les motifs « RFX » entre eux, ce qui avait également été observé dans les cellules Min6b1. Cela confirme que les ensembles de sites de liaison du même facteur de transcription ou de membres de la même famille de facteurs transcription, appelés « homotypic clusters » (Ezer et al, 2014), pourraient dans notre cas également permettre une régulation transcriptionnelle fine des gènes cibles de RFX6 dans les ilots pancréatiques. Enfin, l'analyse de co-occurrence est représentée par une heatmap où un Z score élevé représenté en rouge signifie que la co-occurrence des deux motifs apparaît plus fréquemment que ce qui aurait été obtenu par

chance (Figure 42B). Un zoom de cette représentation comprenant la co-occurrence avec les motifs RFX (encadrés en noir) est présentée en Figure 42C.

A

	SP1	SP2	NRF1	KLF5	EGR1	RFX2	RFX5	RFX1
RFX1	14	30	9	9	16	359	96	122
RFX2	19	37	11	13	20	168	108	359
RFX5	11	12	5	3	5	108	12	96

B



C

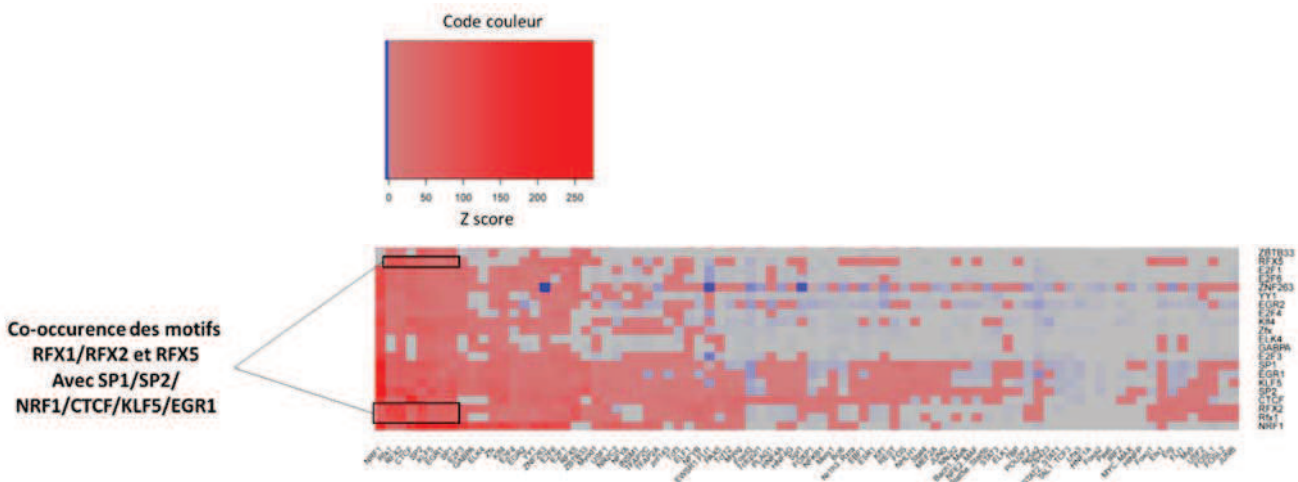


Figure 42 : Analyse de co-occurrence de motifs dans les cibles de RFX6 dans les ilots pancréatiques. Les séquences sous les 10143 pics de ChIP Seq RFX6 dans les ilots pancréatiques sont analysées avec la base de données de motifs Jaspar 2014. **(A)** Les motifs dont la co-occurrence est la plus élevée avec les motifs de type RFX sont présentés. **(B)** Heatmap où un Z score élevé représenté en rouge signifie que la co-occurrence des deux motifs apparaît plus fréquemment que ce qui aurait été obtenu par chance, inversement un Z score peu élevé représenté en bleu signifie que la co-occurrence des deux motifs apparaît moins fréquemment que ce qui aurait été obtenu par chance. **(C)** Zoom sur la région de la heatmap comprenant les motifs RFX.

Ces analyses d'occurrence et de co-occurrence ont ainsi permis de confirmer la présence enrichie de motifs « RFX » au niveau des sites de fixation de RFX6 dans les ilots pancréatiques et de mettre en avant d'éventuels cofacteurs tels que NRF1, CTCF ou les autres membres de la famille RFX par exemple. Cependant, d'autres analyses pourraient encore être développées telles que la recherche de co-régulation éventuelles de certaines cibles par test de transactivation ou une recherche des voies biologiques liées aux motifs hautement occursents permettant l'identification des fonctions encore non dévoilées de RFX6 dans les ilots pancréatiques.

A.4.d. Identification des gènes ciblés dans les ilots pancréatiques et régulés dans la cellule beta par RFX6

Dans le but de découvrir plus précisément le rôle de RFX6 *in vivo* dans les ilots pancréatiques, et selon la même procédure que celle décrite dans le paragraphe Résultats A.3.c réalisée sur les sites de fixation de RFX6 identifiés dans les cellules Min6b1, la

recherche des gènes directement régulés par RFX6 dans les ilots pancréatiques de souris a été effectuée. Pour cela une comparaison des données de CHIP Seq obtenues dans les ilots de Langerhans et de RNA Seq chez la souris delta beta a donc été réalisée et montre que 633 gènes sont directement régulés par RFX6 dont 390 sont réprimés et 243 sont activés (Figure 43).

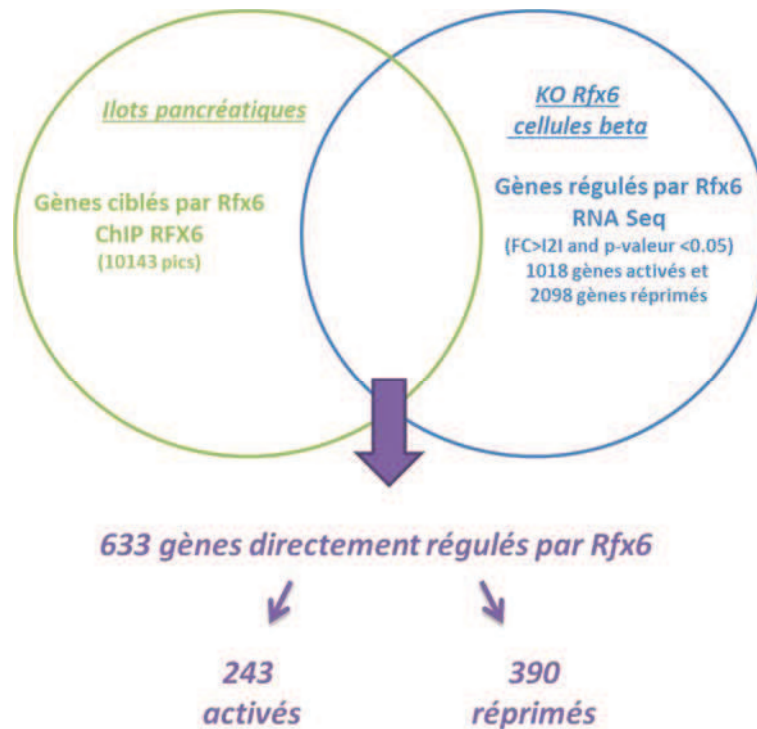


Figure 43 : Identification des gènes ciblés dans les ilots pancréatiques et régulés dans la cellule beta par RFX6. La compilation des données de transcriptomique (gènes dont l'expression est dérégulée (RNA Seq) dans les cellules beta pancréatiques chez la souris delta beta, $|FC|>2$ et $p\text{-valeur}<0,05$) et des sites de liaisons de RFX6 (7524 gènes associés aux 10143 pics identifiés pour le CHIP Seq RFX6 dans les ilots pancréatiques) permet l'identification de 633 gènes potentiellement directement régulés par RFX6 dont 243 sont activés et 390 sont réprimés.

Parmi ceux-ci, les 100 gènes dont les variations d'expression positives ou négatives les plus importantes observées par RNA Seq dans les cellules beta lorsque Rfx6 est inactivé sont présentés en Figure 44.

Gène	Identifiant Ensembl	Description	Moyenne Expression CTRL	Moyenne Expression MUT	FC	p-valeur	Position génomique du pic de ChIP Seq anti-RFX6
Vwf	ENSMUSG00000001930	Von Willebrand factor homolog	93.75	1156.24	8.1	2.03E-13	Promoteur-TSS et intron
Nnmt	ENSMUSG000000032271	nicotinamide N-methyltransferase	6.47	155.95	7.1	6.43E-08	Intergénique
C1qtnf9	ENSMUSG000000071347	C1q and tumor necrosis factor related protein 9	8.23	104.01	6.4	2.19E-08	Intron
Col18a1	ENSMUSG000000001435	collagen, type XVIII, alpha 1	31.05	338.65	6.4	3.13E-09	Intron
Dpysl3	ENSMUSG000000024501	dihydropyrimidinase-like 3	57.94	527.66	6.3	1.26E-10	Intron
Rtkn2	ENSMUSG000000037846	rhotekin 2	78.96	527.36	6.2	7.64E-32	Intron
Gzmk	ENSMUSG000000042385	granzyme K	3.06	31.85	6.1	7.84E-09	Promoteur-TSS
Esr1	ENSMUSG000000019768	estrogen receptor 1 (alpha)	56.31	554.35	6.1	5.24E-09	Intron
Ptgis	ENSMUSG000000017969	prostaglandin I2 (prostacyclin) synthase	20.05	404.22	6.0	1.36E-06	Intron
Dcn	ENSMUSG000000019929	decorin	265.05	3714.50	5.8	8.62E-07	Intergénique
Loxl1	ENSMUSG000000032334	lysyl oxidase-like 1	13.72	153.18	5.7	2.91E-07	Exon
Mmrn1	ENSMUSG000000054641	multimerin 1	4.29	48.35	5.7	2.91E-07	Intergénique
Lgals7	ENSMUSG000000053522	lectin, galactose binding, soluble 7	3.45	115.81	5.7	4.59E-06	Intergénique
Sfrp2	ENSMUSG000000027996	secreted frizzled-related protein 2	15.14	216.45	5.6	1.52E-06	Intergénique
Aldh1a2	ENSMUSG000000013584	aldehyde dehydrogenase family 1, subfamily A2	11.42	128.69	5.5	7.52E-07	Intron
Fmod	ENSMUSG000000041559	fibromodulin	0.62	10.82	5.5	3.36E-06	Intergénique
Slpi	ENSMUSG000000017002	secretory leukocyte peptidase inhibitor	72.65	1090.39	5.4	3.49E-06	Promoteur-TSS
Ly6c2	ENSMUSG000000022584	lymphocyte antigen 6 complex, locus C2	2.56	69.89	5.4	7.70E-06	Intergénique
Rprm	ENSMUSG000000075334	reprimin, TP53 dependent G2 arrest mediator candidate	1.27	19.02	5.4	3.19E-06	Intergénique
B3gat2	ENSMUSG000000026156	beta-1,3-glucuronyltransferase 2 (glucuronosyltransferase S)	1.48	15.38	5.4	5.03E-07	Intergénique
Itgb2	ENSMUSG000000000290	integrin beta 2	28.25	291.49	5.3	1.05E-06	Intron
Tmem100	ENSMUSG000000069763	transmembrane protein 100	2.65	23.84	5.3	2.72E-07	5'UTR
Pdgfra	ENSMUSG000000029231	platelet derived growth factor receptor, alpha polypeptide	20.74	234.54	5.1	4.05E-06	Intron
Wfdc1	ENSMUSG000000023336	WAP four-disulfide core domain 1	9.51	64.86	5.1	9.30E-10	Intron
Cdh13	ENSMUSG000000031841	cadherin 13	11.17	83.26	5.1	4.36E-08	Intron
Mxd3	ENSMUSG000000021485	Max dimerization protein 3	4.54	30.72	5.0	3.39E-09	Intron
Fndc1	ENSMUSG000000071984	fibronectin type III domain containing 1	51.10	499.23	5.0	2.93E-06	Intron
Mrc1	ENSMUSG000000026712	mannose receptor, C type 1	23.53	384.63	5.0	1.76E-05	Intron
S100a9	ENSMUSG000000056071	S100 calcium binding protein A9 (calgranulin B)	170.26	1694.68	4.9	5.40E-06	Intergénique
Fam198b	ENSMUSG000000027955	family with sequence similarity 198, member B	15.27	126.14	4.8	2.64E-06	Intergénique
Col1a1	ENSMUSG000000001506	collagen, type I, alpha 1	802.44	5784.43	4.8	4.57E-07	Promoteur-TSS
Gas1	ENSMUSG000000052957	growth arrest specific 1	14.30	151.42	4.8	1.34E-05	Intergénique
Pcdh7	ENSMUSG000000029108	protocadherin 7	8.24	68.26	4.8	3.02E-06	Intron
Rab32	ENSMUSG000000019832	RAB32, member RAS oncogene family	13.35	105.25	4.6	5.88E-06	Promoteur-TSS
Capp	ENSMUSG000000056737	capping protein (actin filament), gelsolin-like	142.20	1018.37	4.5	2.29E-06	Promoteur-TSS
Cldn10	ENSMUSG000000022132	claudin 10	34.34	257.21	4.5	5.22E-06	Intron
Rerg1	ENSMUSG000000092164	RERG/RAS-like	2.05	26.89	4.5	4.64E-05	Intergénique
Dkk3	ENSMUSG000000030772	dickkopf homolog 3 (Xenopus laevis)	15.01	84.20	4.5	7.30E-09	Intron et intergénique
Tfec	ENSMUSG000000029553	transcription factor EC	0.95	14.69	4.4	7.15E-05	Intergénique
Slfn2	ENSMUSG000000072620	schlafen 2	27.37	185.39	4.4	2.12E-06	Intergénique
Islr2	ENSMUSG000000051243	immunoglobulin superfamily containing leucine-rich repeat 2	1.13	7.07	4.4	5.45E-07	Exon et intergénique
Ntn1	ENSMUSG000000020902	netrin 1	24.30	145.88	4.4	2.06E-07	Intergénique
Bnc1	ENSMUSG000000025105	basonuclin 1	1.31	33.55	4.4	0.00011802	Intergénique
Mmp9	ENSMUSG000000017737	matrix metalloproteinase 9	18.50	145.40	4.4	1.64E-05	Exon
Mki67	ENSMUSG000000031004	antigen identified by monoclonal antibody Ki 67	25.75	130.70	4.3	4.95E-11	Intergénique
Nov	ENSMUSG000000037362	nephroblastoma overexpressed gene	34.40	185.81	4.3	1.07E-08	Intergénique
Ch1	ENSMUSG000000030077	cell adhesion molecule with homology to L1CAM	16.59	96.29	4.3	5.57E-07	Intergénique
Ifitm3	ENSMUSG000000025492	interferon induced transmembrane protein 3	166.87	1058.44	4.2	5.39E-06	TTS
Ear11	ENSMUSG000000047222	eosinophil-associated, ribonuclease A family, member 11	0.48	24.99	4.2	0.000183732	Intergénique
Gpr84	ENSMUSG000000063234	G protein-coupled receptor 84	0.00	8.05	4.1	0.000193216	Exon

Gène	Identifiant Ensembl	Description	Moyenne Expression CTRL	Moyenne Expression MUT	FC	p-valeur	Position génomique du pic de ChIP Seq anti-RFX6
Cps1	ENSMUSG00000025991	carbamoyl-phosphate synthetase 1	74.50	0.07	-84.1	7.26E-48	Intron
Fabp2	ENSMUSG00000023057	fatty acid binding protein 2, intestinal	148.59	0.48	-35.9	7.42E-29	Intron
Mlxipl	ENSMUSG00000005373	MLX interacting protein-like	1595.32	84.51	-17.7	5.35E-148	Intron
Muc4	ENSMUSG000000079620	mucin 4	569.11	25.30	-13.8	9.73E-22	Intron
Armc4	ENSMUSG000000061802	armadillo repeat containing 4	30.64	2.36	-9.7	9.06E-25	Promoteur-TSS
2210407C18Rik	ENSMUSG000000037145	RIKEN cDNA 2210407C18 gene	86.00	3.13	-9.2	1.11E-10	Intergénique
RP23-329D14.2	ENSMUSG000000093739		22.87	0.00	-8.0	1.58E-08	Intron
Cdx2	ENSMUSG000000029646	caudal type homeobox 2	7.31	0.00	-7.6	4.92E-08	Intergénique
Dcdc5	ENSMUSG000000074981	doublecortin domain containing 5	81.94	9.17	-7.5	1.12E-23	Promoteur-TSS
Gm16308	ENSMUSG000000090237	predicted gene 16308	25.32	0.89	-7.4	2.07E-08	Intron
Pck1	ENSMUSG000000027513	phosphoenolpyruvate carboxykinase 1, cytosolic	23.55	1.19	-6.9	4.07E-08	Intergénique
Reln	ENSMUSG000000042453	reelin	264.23	30.27	-5.9	1.51E-09	Intron
Gpr179	ENSMUSG000000070337	G protein-coupled receptor 179	167.30	22.78	-5.9	1.59E-13	Intron
Adh4	ENSMUSG000000037797	alcohol dehydrogenase 4 (class II), pi polypeptide	24.73	0.00	-5.5	7.33E-06	Intron
Abcc2	ENSMUSG000000025194	ATP-binding cassette, sub-family C (CFTR/MRP), member 2	238.59	29.47	-5.5	2.95E-09	Intron
Th	ENSMUSG000000000214	tyrosine hydroxylase	675.91	119.69	-5.4	2.31E-49	Intron
Gm281	ENSMUSG000000084902	predicted gene 281	197.81	35.87	-5.3	1.96E-36	Intron
Baiap3	ENSMUSG000000047507	BAI1-associated protein 3	1366.19	245.47	-5.1	6.15E-23	Intron
5330413P13Rik	ENSMUSG000000086822	RIKEN cDNA 5330413P13 gene	152.73	26.78	-5.1	2.01E-18	Promoteur-TSS
Relt	ENSMUSG000000008318	RELT tumor necrosis factor receptor	1747.65	326.04	-5.0	5.28E-29	Intron
Prss32	ENSMUSG000000048992	protease, serine, 32	18.20	0.50	-5.0	2.17E-05	Promoteur-TSS
2810442I21Rik	ENSMUSG000000087060	RIKEN cDNA 2810442I21 gene	80.56	14.11	-4.9	8.10E-14	Intron
Mfi2	ENSMUSG000000022780	antigen p97 (melanoma associated) identified by monoclonal antibodies 133.2 and 96.5	12.78	1.91	-4.9	1.27E-08	3'UTR
Cacna1e	ENSMUSG000000004110	calcium channel, voltage-dependent, R type, alpha 1E subunit	46.15	7.09	-4.9	4.89E-09	Intergénique et intron
Mogat2	ENSMUSG000000052396	monoacylglycerol O-acyltransferase 2	24.89	2.82	-4.6	4.87E-06	Intergénique
4930578M01Rik	ENSMUSG000000087446	RIKEN cDNA 4930578M01 gene	241.00	49.96	-4.6	1.99E-32	Intergénique et promoteur-TSS
Fam135b	ENSMUSG000000036800	family with sequence similarity 135, member B	199.20	42.48	-4.6	7.20E-57	Intergénique
4930555F03Rik	ENSMUSG000000031559	RIKEN cDNA 4930555F03 gene	76.12	13.68	-4.4	1.09E-08	Intron
Ttc16	ENSMUSG000000039021	tetratricopeptide repeat domain 16	44.38	9.22	-4.3	7.14E-14	Promoteur-TSS
Hsph1	ENSMUSG000000029657	heat shock 105kDa/110kDa protein 1	6149.83	1383.59	-4.2	1.26E-19	Promoteur-TSS
Gm5577	ENSMUSG000000084950	predicted gene 5577	137.69	31.20	-4.1	2.50E-19	Promoteur-TSS et intron
Ffar1	ENSMUSG000000044453	free fatty acid receptor 1	975.95	232.17	-4.0	9.24E-34	Intergénique
Krt20	ENSMUSG000000035775	keratin 20	26.81	4.24	-3.9	3.64E-05	Intergénique
SNORA73	ENSMUSG000000065158	Small nucleolar RNA SNORA73 family	942.26	221.79	-3.9	1.02E-14	TTS
Fbxw24	ENSMUSG000000062275	F-box and WD-40 domain protein 24	77.00	9.74	-3.9	0.000146755	Intergénique
Tcp11	ENSMUSG000000062859	t-complex protein 11	138.98	34.24	-3.9	1.66E-28	Intergénique
9330101J02Rik	ENSMUSG000000049571	RIKEN cDNA 9330101J02 gene	24.43	5.39	-3.9	2.73E-09	Promoteur-TSS et intron
Acta1	ENSMUSG000000031972	actin, alpha 1, skeletal muscle	28.03	5.98	-3.8	6.65E-08	Intron
Dok7	ENSMUSG000000044716	docking protein 7	24.38	5.48	-3.7	5.73E-07	Intron
Zfp92	ENSMUSG000000031374	zinc finger protein 92	372.86	94.35	-3.7	2.79E-16	Promoteur-TSS
Mesp1	ENSMUSG000000030544	mesoderm posterior 1	77.75	19.58	-3.7	1.11E-14	Exon
Tnr	ENSMUSG000000015829	tenascin R	531.84	134.97	-3.6	1.45E-10	Intergénique et intron
Glb1l2	ENSMUSG000000036395	galactosidase, beta 1-like 2	1574.52	439.42	-3.5	3.93E-22	Promoteur-TSS et intron
Plcxd3	ENSMUSG000000049148	phosphatidylinositol-specific phospholipase C, X domain containing 3	421.85	121.88	-3.4	2.12E-36	Intron
Npy	ENSMUSG000000029819	neuropeptide Y	713.54	181.22	-3.4	7.69E-07	Intergénique et intron
Meg3	ENSMUSG000000021268	maternally expressed 3	19363.66	5495.86	-3.3	3.36E-15	Intergénique
Gm239	ENSMUSG000000034057	predicted gene 239	4.23	0.42	-3.3	0.001482696	Intron
Gm19357	ENSMUSG000000093629	predicted gene, 19357	26.77	6.83	-3.3	3.87E-06	Intergénique
Speer2	ENSMUSG000000063163	spermatogenesis associated glutamate (E)-rich protein 2	39.28	8.33	-3.1	0.000385978	Intergénique
Fxyd6	ENSMUSG000000066705	FXYD domain-containing ion transport regulator 6	1107.74	341.76	-3.1	2.23E-18	Promoteur-TSS et intron

Figure 44 : Gènes ciblés dans les ilots pancréatiques et régulés dans la cellule beta par RFX6. Les 50 gènes les plus surexprimés (premier tableau) ou sous-exprimés (deuxième tableau) (RNA Seq) dans les cellules beta pancréatiques chez la souris n'exprimant pas RFX6 dans les cellules beta (MUT) par rapport à la souris sauvage (WT) sont listés avec leur variation d'expression (FC) et la p-valeur associée. La position du site de fixation de RFX6 identifié lors du ChIP RFX6 dans les ilots pancréatiques est indiquée pour chaque gène.

Enfin, un classement des gènes ciblés dans les ilots pancréatiques et régulés dans la cellule beta par RFX6 est effectué par voies biologiques (Gene Ontology) (**Figure 45**). Les 10 gènes les plus dérégulés dans chaque liste sont présentés en **Figure 45A** et la totalité des gènes impliqués dans la sécrétion d'insuline et l'homéostasie du glucose est présentée en **Figure 45B**.

A

Voies biologiques	% des gènes ciblés et régulés par RFX6	Gènes ciblés et régulés par RFX6
Homéostasie du glucose	2.1	Bad ; Pparg
		Stk11 ; Stat3 ; Rph3al ; Ptch1 ; Gck ; Pax6 ; Cacna1c ; Ffar1 ; Cacna1e ; Pck1 ; Mlxipl
Sécrétion d'insuline	5.1	Rest ; Arrb1 ; Blk ; Inhbb ; Pim3 ; Cpt1a ; Serp1 ; Ppp3ca ; Bad ; Eif2ak3
		Gck ; Glp1r ; Syt9 ; Ghsr ; Hmgn3 ; Oit1 ; Cacna1c ; Gpr ; Ffar1 ; Cacna1e
Régulation de la transcription	54.2	Esr1 ; Sfrp2 ; Mxd3 ; Col1a1 ; Dkk3 ; Tfec ; Bnc1 ; Ccna2 ; Ptrf ; Fgfr2
		Zfp541 ; Zfp174 ; Grhl1 ; Agap2 ; Per3 ; Barhl1 ; Mesp1 ; Zfp92 ; Cdx2 ; Mlxipl
Cycle cellulaire	22.1	Rprm ; Cdh13 ; Gas1 ; Gsg2 ; Ncaph ; Ccna2 ; Birc5 ; Casc5 ; Cks2 ; Dlgap5
		Cntrob ; Bcl2l11 ; Haus5 ; Mdm4 ; BC024139 ; Nanog ; 2410004A20Rik ; Cspp1 ; Crocc ; Mlxipl

Résultats-Recherche des gènes cibles du facteur de transcription RFX6

B

SECRETION D'INSULINE							
Gène	Identifiant Ensembl	Description	Moyenne Expression CTRL	Moyenne Expression MUT	FC	p-valeur	Position génomique du pic de ChIP Seq anti-RFX6
Rest	ENSMUSG00000029249	RE1-silencing transcription factor	18.88	58.54	2.6	0.000489274	Intergénique
Arrb1	ENSMUSG00000018909	arrestin, beta 1	753.76	1948.76	2.5	1.35E-08	Promoteur-TSS
Blk	ENSMUSG00000014453	B lymphoid kinase	1.54	10.15	2.5	0.018155648	TTS
Inhbb	ENSMUSG000000037035	inhibin beta-B	29.03	70.58	2.1	0.008277884	Intergénique
Pim3	ENSMUSG000000035828	proviral integration site 3	610.11	1082.10	1.7	0.012859894	Promoteur-TSS et exon
Cpt1a	ENSMUSG000000024900	carnitine palmitoyltransferase 1a, liver	266.92	439.29	1.6	0.014039923	Intron
Serp1	ENSMUSG000000027808	stress-associated endoplasmic reticulum protein 1	6684.80	10255.67	1.5	0.00000592	Promoteur-TSS
Ppp3ca	ENSMUSG000000028161	protein phosphatase 3, catalytic subunit, alpha isoform	1499.58	2172.78	1.4	0.000211182	Intron
Bad	ENSMUSG000000024959	BCL2-associated agonist of cell death	487.89	635.72	1.3	0.000308632	TTS
Eif2ak3	ENSMUSG000000031668	eukaryotic translation initiation factor 2 alpha kinase 3	1553.64	2002.86	1.3	0.015071449	Intron
Oxct1	ENSMUSG000000022186	3-oxoacid CoA transferase 1	5552.09	6691.36	1.2	0.00382301	Promoteur-TSS
Irs1	ENSMUSG000000055980	insulin receptor substrate 1	114.00	87.31	-1.3	0.02939742	Promoteur-TSS
Slc30a8	ENSMUSG000000022315	solute carrier family 30 (zinc transporter), member 8	19055.92	14614.63	-1.3	0.008895781	Intergénique
Snap25	ENSMUSG000000027273	synaptosomal-associated protein 25	1976.84	1485.16	-1.3	0.045847988	Intergénique et intron
Rims2	ENSMUSG000000037386	regulating synaptic membrane exocytosis 2	565.95	410.57	-1.4	0.032127359	Promoteur-TSS
Dgat1	ENSMUSG000000022555	diacylglycerol O-acyltransferase 1	760.11	545.88	-1.4	0.000198633	Promoteur-TSS
Mgea5	ENSMUSG000000025220	meningioma expressed antigen 5 (hyaluronidase)	1086.36	774.29	-1.4	0.00000118	Promoteur-TSS
Abcc8	ENSMUSG000000040136	ATP-binding cassette, sub-family C (CFTR/MRP), member 8	6540.14	4579.93	-1.4	0.013125843	Intron
Sirt4	ENSMUSG000000029524	sirtuin 4 (silent mating type information regulation 2 homolog) 4 (S. cerevisiae)	374.51	261.68	-1.4	0.000374994	Intergénique
Rapgef4	ENSMUSG000000049044	Rap guanine nucleotide exchange factor (GEF) 4	6952.18	4735.31	-1.5	0.0000986	Intron
Htt	ENSMUSG000000029104	huntingtin	1263.38	848.02	-1.5	0.005451179	Promoteur-TSS
Irs2	ENSMUSG000000038894	insulin receptor substrate 2	471.33	315.47	-1.5	0.000936225	Intron
Gck	ENSMUSG000000041798	glucokinase	2130.97	1344.81	-1.6	0.0000811	Intron et 5'UTR
Glp1r	ENSMUSG000000024027	glucagon-like peptide 1 receptor	5989.60	3404.97	-1.7	0.010632697	Intron et intergénique
Syt9	ENSMUSG000000062542	synaptotagmin IX	156.39	87.70	-1.7	0.014305292	Intron
Ghr	ENSMUSG000000051136	growth hormone secretagogue receptor	44.13	25.18	-1.7	0.000499342	Intron
Hmgn3	ENSMUSG000000066456	high mobility group nucleosomal binding domain 3	7401.27	3914.30	-1.9	5.82E-17	Intergénique
Oit1	ENSMUSG000000021749	oncprotein induced transcript 1	149.37	70.47	-2.0	0.000169382	Intron
Cacna1c	ENSMUSG000000051331	calcium channel, voltage-dependent, L type, alpha 1C subunit	337.79	148.65	-2.2	0.0000154	Promoteur-TSS
Gipr	ENSMUSG000000030406	gastric inhibitory polypeptide receptor	1531.66	498.03	-3.0	8.22E-20	Intron
Ffar1	ENSMUSG000000044453	free fatty acid receptor 1	975.95	232.17	-4.0	9.24E-34	Intergénique
Cacna1e	ENSMUSG000000004110	calcium channel, voltage-dependent, R type, alpha 1E subunit	46.15	7.09	-4.9	4.89E-09	Intron et intergénique

HOMEOSTASIE DU GLUCOSE							
Gène	Identifiant Ensembl	Description	Moyenne Expression CTRL	Moyenne Expression MUT	FC	p-valeur	Position génomique du pic de ChIP Seq anti-HA-RFX6
Bad	ENSMUSG000000024959	BCL2-associated agonist of cell death	487.89	635.72	1.3	0.000308632	TTS
Pparg	ENSMUSG000000000440	peroxisome proliferator activated receptor gamma	162.72	193.25	1.2	0.042367955	Promoteur-TSS
Stk11	ENSMUSG000000003068	serine/threonine kinase 11	1780.29	1450.70	-1.2	0.000000026	Promoteur-TSS et intron
Stat3	ENSMUSG000000004040	signal transducer and activator of transcription 3	7568.46	5948.86	-1.3	0.0000052	Intron et TTS
Rph3al	ENSMUSG000000020847	rabphilin 3A-like (without C2 domains)	2893.11	2180.00	-1.3	0.008895606	Promoteur-TSS et intergénique
Ptch1	ENSMUSG000000021466	patched homolog 1	460.22	333.50	-1.4	0.038348603	Intron et intergénique
Gck	ENSMUSG000000041798	glucokinase	2130.97	1344.81	-1.6	0.0000811	Intron et 5'UTR
Pax6	ENSMUSG000000027168	paired box gene 6	6600.11	3210.75	-2.0	7.73E-14	Intron et promoteur-TSS
Cacna1c	ENSMUSG000000051331	calcium channel, voltage-dependent, L type, alpha 1C subunit	337.79	148.65	-2.2	0.0000154	Promoteur-TSS
Ffar1	ENSMUSG000000044453	free fatty acid receptor 1	975.95	232.17	-4.0	9.24E-34	Intergénique
Cacna1e	ENSMUSG000000004110	calcium channel, voltage-dependent, R type, alpha 1E subunit	46.15	7.09	-4.9	4.89E-09	Intergénique et intron
Pck1	ENSMUSG000000027513	phosphoenolpyruvate carboxykinase 1, cytosolic	23.55	1.19	-6.9	4.07E-08	Intergénique
Mxipl	ENSMUSG000000005373	MLX interacting protein-like	1595.32	84.51	-17.7	5.35E-148	Promoteur-TSS

Figure 45 : Classement par voies biologiques des gènes ciblés dans les îlots pancréatiques et régulés dans la cellule beta par RFX6. Les gènes ciblés par RFX6 (ChIP RFX6 dans les îlots pancréatiques) et surexprimés (en bleu) ou sous-exprimés (en orange) en absence de RFX6 (RNA Seq delta beta) dans les cellules beta pancréatiques sont listés et classés par voies biologiques (Gene Ontology). **(A)** Le pourcentage du nombre total de gènes ciblés et régulés par RFX6 (633 gènes au total) et les 10 gènes les plus dérégulés dans chaque liste sont présentés. Les gènes en gras présentent une variation d'expression significative supérieure à deux (RNA Seq IFCI>2 et p-valeur<0,05). **(B)** Listes complètes des gènes impliqués dans la sécrétion d'insuline et l'homéostasie du glucose. Les niveaux d'expression dans le RNA Seq dans les îlots pancréatiques

chez la souris n'exprimant pas RFX6 dans les cellules beta (MUT) par rapport à la souris sauvage (WT) sont listés avec leur variation d'expression (FC) et la p-valeur associée. La position génomique du site de fixation de RFX6 identifié lors du ChIP RFX6 dans les ilots pancréatiques est indiquée pour chaque gène.

Parmi les gènes mis en avant lors de cette analyse, on en compte 32 (5,1%) ayant un rôle connu dans la sécrétion d'insuline et 13 (2,1%) ayant un rôle connu dans l'homéostasie du glucose. De nombreux gènes ont également une fonction connue dans la régulation de la transcription (54,2%) ou le cycle cellulaire (22,1%).

Les gènes décrits ci-dessous impliqués spécifiquement dans la sécrétion d'insuline et l'homéostasie du glucose, sont directement régulés dans les cellules beta pancréatiques par RFX6 et pourraient expliquer le rôle précis de RFX6 dans la cellule beta.

Ce sont tout d'abord des gènes impliqués dans le métabolisme glucidique tels que *Ffar1*, *Glpr*, *Irs 1* et *2*, *Hmgn3* (décrits dans le paragraphe Résultats A.3.c), *Slc30a8* ou *Pparg* qui sont mis en évidence.

Le gène ***Slc30a8*** (Solute carrier family 30A8) a été identifié et lié au diabète de type 2 lors d'une étude de type GWAS (genome wide association study) en 2007 (Sladek et al, 2007). Le rôle de *Slc30a8* dans la régulation de la sécrétion d'insuline et de l'homéostasie du glucose est controversé puisque son absence dans les ilots pancréatiques isolés de souris provoque une augmentation de la sécrétion d'insuline alors que les souris n'exprimant plus ce gène présentent des taux d'insuline plasmatiques diminués et une intolérance au glucose. D'autre part, *Slc30a8* a récemment été identifié comme protecteur contre le diabète de type 2 chez l'homme (Flannick et al, 2014; Rutter & Chimienti, 2015).

Pparg (peroxisome proliferator-activated receptor gamma) est fortement exprimé dans les adipocytes et participe à l'assimilation des acides gras libres et des glucides, il protège également contre la résistance à l'insuline. Il est de ce fait utilisé en thérapie contre le diabète de type 2 et le syndrome métabolique (Jay & Ren, 2007). Par surexpression de *Pparg* spécifiquement dans les cellules beta de souris obèses, une intolérance au glucose, une réduction de la masse des cellules beta, une stimulation de l'apoptose des cellules beta ainsi qu'un taux d'insuline plasmatique diminué sont observés (Hogh et al, 2014) ce qui confirme son rôle clé dans le métabolisme du glucose. D'autre part, il a également été identifié, tout comme *NRF1*, comme régulateur clé, de nombreux gènes dérégulés spécifiquement chez des patients diabétiques et pré-diabétiques (Patti, 2004).

Les gènes circadiens **Per2** et **Per3** qui régulent la sécrétion d'insuline sont également retrouvés parmi les gènes directement régulés par RFX6 dans les ilots pancréatiques. En effet, le rythme circadien a été relié au métabolisme cellulaire (Kovac et al, 2009) et des souris déficientes pour **Per2** présentent une sécrétion d'insuline fortement stimulée en présence de glucose (Zhao et al, 2012).

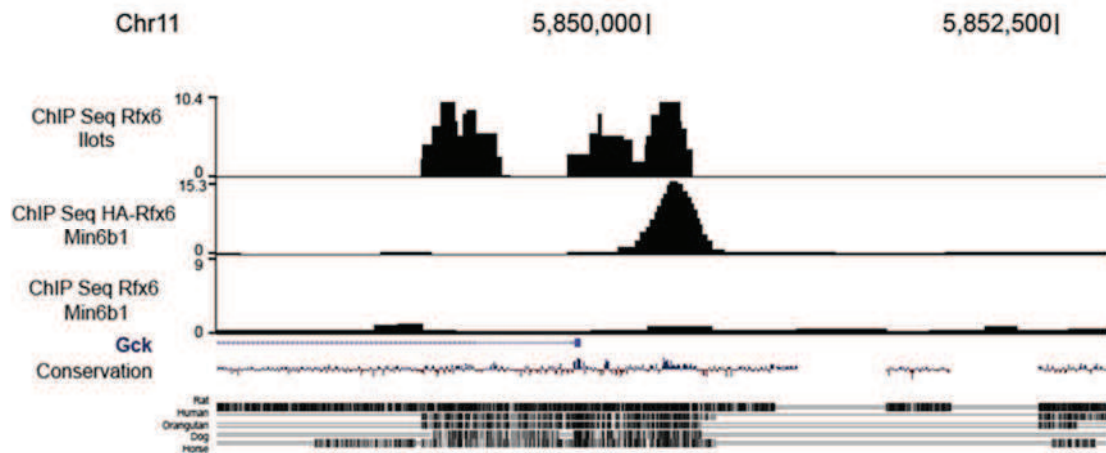
D'autres facteurs de transcription impliqués dans le développement du pancréas sont également directement régulés par RFX6 dans les ilots pancréatiques tels que **Pax6** ou **Myt1**.

Pax6 (paired box 6) est un facteur de transcription impliqué dans la régulation du développement des ilots pancréatiques et de la transcription des gènes des hormones pancréatiques (Sander et al, 1997). **Pax6** réprime notamment l'activation du gène de l'insuline par **NeuroD1** ou **Pdx1** dans des cellules d'insulinome de rat (Wolf et al, 2010). **Pax6** permet également la maintenance de la fonction de la cellule beta par régulation de l'expression de différents gènes clés tels que **Glut2**, **PC1/3**, **Pdx1**, **MafA** ou **NeuroD1** (Sander et al, 1997; Wen et al, 2009). Enfin, la régulation de son expression par le glucose et son action sur le gène suppresseur de tumeur **PTEN** et la viabilité cellulaire a été décrite plus récemment dans le modèle cellulaire **INS-1E** (Balakrishnan et al, 2014).

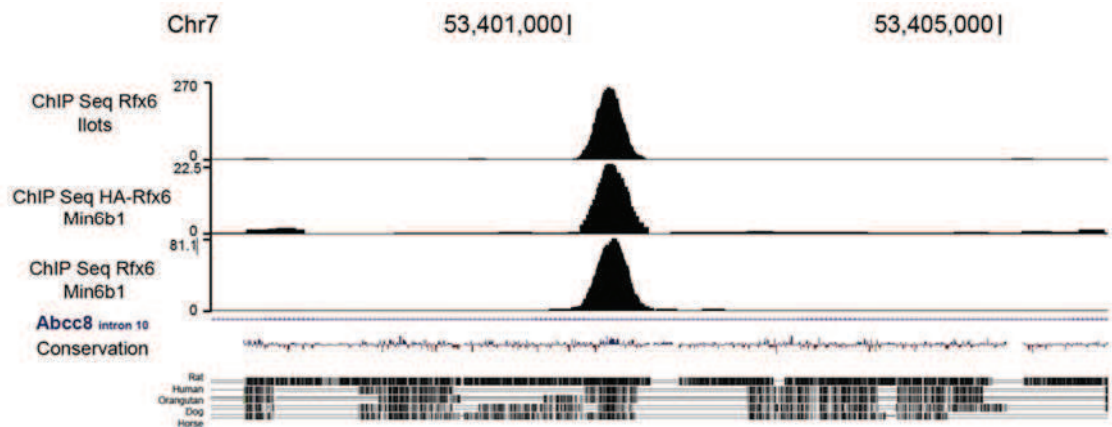
Myt1 (myelin transcription factor 1) est exprimé dans le pancréas lors du développement dans les progéniteurs endocrines mais également dans les cellules différenciées. Lorsque ce gène est inactivé chez la souris, les cellules des ilots pancréatiques expriment plusieurs hormones. Par inactivation spécifique de **Myt1** dans le pancréas, les cellules des ilots expriment également plusieurs hormones et présentent une réduction d'expression du transporteur de glucose **Glut2** (Wang et al, 2007). Il a également été montré que **Myt1** et **Ngn3** régulent positivement leurs expressions pour promouvoir la différenciation endocrine (Wang et al, 2008).

Cette recherche de gènes cibles dans les ilots pancréatiques aura également permis de confirmer que les gènes cibles de RFX6 que nous avons mis en avant dans la lignée cellulaire **Min6b1** dans la publication de 2014 dans *Cell Reports* (Piccand et al, 2014) décrite dans le paragraphe Résultats A.3.e sont retrouvés *in vivo* pour la plupart. Les sites de fixation de RFX6 au niveau des différents gènes d'intérêt sont représentés dans la **Figure 46**.

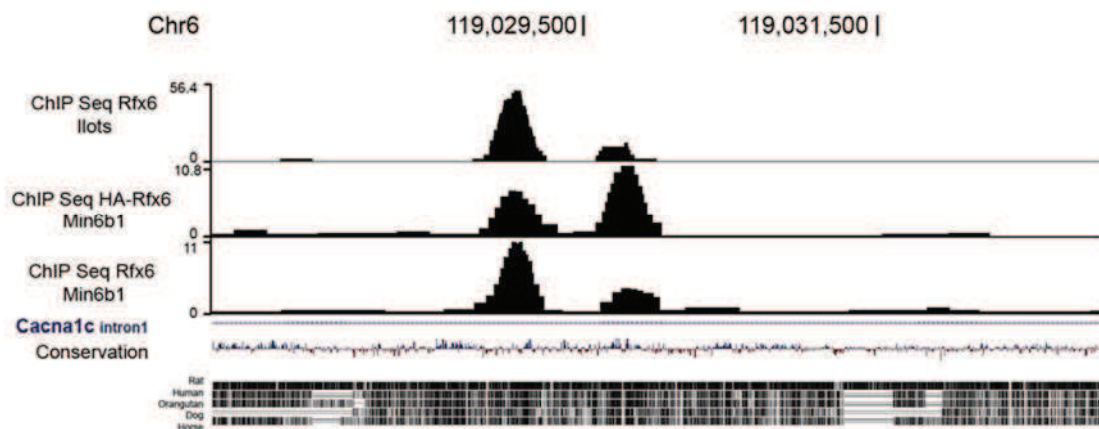
A / RFX6 se fixe en amont du TSS du gène de la glucokinase



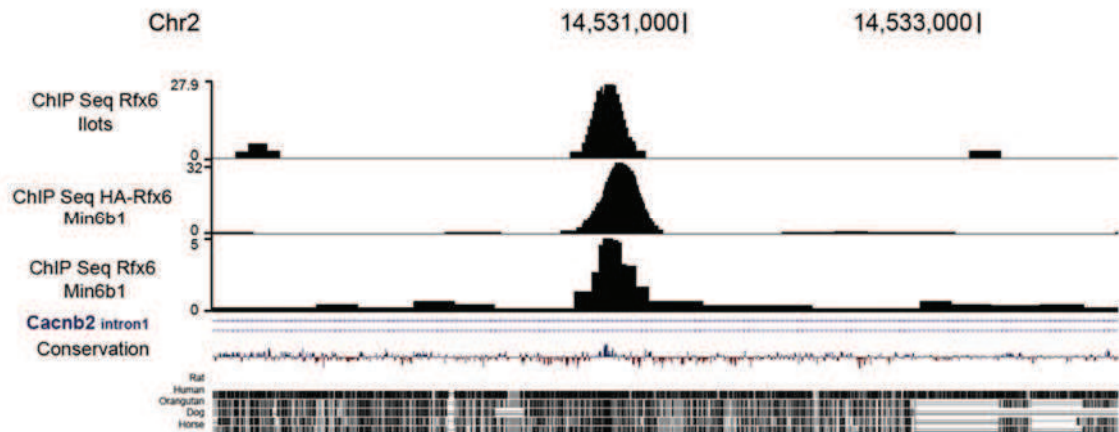
B / RFX6 se fixe sur l'intron 10 de Abcc8



C / RFX6 se fixe sur l'intron 1 de Cacna1c



D / RFX6 se fixe sur l'intron 1 de Cacnb2



E / RFX6 se fixe en amont de Ldha

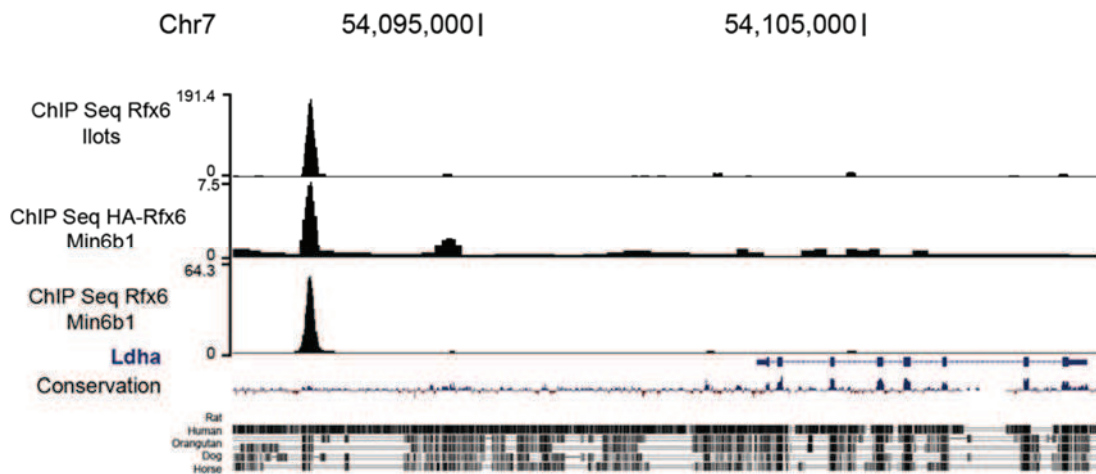


Figure 46 : Comparaison des sites de fixation de RFX6 dans les cellules Min6b1 et les ilots pancréatiques au niveau des gènes clés de la fonction et de l'identité des cellules beta adultes. Les sites cibles de RFX6 en amont du TSS de la glucokinase (A), dans l'intron 10 de Abcc8 (B), au niveau du premier intron de Cacna1c (C) et Cacnb2 (D) et 10 kb en amont de Ldha (E) sont identiques dans la lignée cellulaire Min6b1 *in vitro* et dans les ilots de Langerhans *in vivo*.

Le rôle de RFX6 dans le maintien de l'identité et de la fonction de la cellule sécrétrice d'insuline a ainsi été confirmé.

A.4.e. Détermination de réseaux fonctionnels parmi les gènes ciblés dans les ilots pancréatiques et régulés dans la cellule beta par RFX6

Enfin, pour aller plus en avant dans la découverte de la fonction précise de RFX6 dans la cellule beta adulte, une analyse fonctionnelle des 633 gènes potentiellement directement régulés par RFX6 dans les cellules beta pancréatiques (**Figure 43**) a été réalisée avec l'aide d'Ingenuity Pathway Analysis (IPA <http://www.ingenuity.com/products/ipa>). IPA est une application logicielle basée sur le Web pour l'analyse, l'intégration et l'interprétation des données provenant d'expérience de génomique. IPA permet notamment de créer des réseaux fonctionnels parmi une liste de gènes d'intérêt.

Le résultat de l'analyse IPA des gènes ciblés dans les ilots pancréatiques et régulés dans la cellule beta par RFX6 est résumée en **Figure 47**. Les gènes sont ainsi regroupés par principales voies cellulaires, fonctions moléculaires et cellulaires, développement et fonction des systèmes physiologiques et enfin par réseaux fonctionnels.

Les voies cellulaires principales représentées dans notre jeu de données sont étonnamment peu liées à des fonctions pancréatiques ou à des organes associés tels que ceux du système digestif. En effet, la voie cellulaire principale associée à notre ensemble de gènes est la signalisation du guidage axonal. Les fonctions moléculaires et cellulaires associées aux gènes directement régulés par RFX6 regroupent le mouvement, la croissance ainsi que la prolifération et la morphologie cellulaire, mais aussi le développement, la survie et la mort cellulaire. En ce qui concerne le groupe développement et fonction des systèmes physiologiques, ce sont ici les systèmes cardiovasculaires et respiratoires de même que le développement de l'organisme, embryonnaire et tissulaire qui sont soulignés.

Parmi les molécules régulées, **VWF** et **NNMT** sont les plus réprimées et ciblées par RFX6. La glycoprotéine VWF permet notamment l'activation des plaquettes dans le phénomène de coagulation (Lenting et al, 2012; Lenting et al, 2015). NNMT a été pour sa part étudiée dans des cas de maladie de Parkinson (Matsubara et al, 2002) mais également dans le pancréas et le tissu adipeux, où elle joue un rôle dans le développement du syndrome métabolique (Giuliante et al, 2015). L'étude du métabolome de cellules pancréatiques cancéreuses a de plus été associée à une augmentation importante de l'expression de NNMT (Bi et al, 2014). Inversement, ce sont **CPS1** et **FABP2**, qui sont les molécules les plus activées et ciblées par RFX6 dans notre jeu de données. CPS1 est principalement connu pour son rôle dans le foie où elle participe au cycle de l'urée (Martinez et al, 2010) et un SNP de cette protéine a également été lié à une étude concentrée sur la sensibilité à l'insuline et le diabète de type2 (Xie et al, 2013). FABP2 est pour sa part associé

aux maladies métaboliques et plus précisément à l'obésité et à la stéatose hépatique non alcoolique (Thumser et al, 2014). Cependant, les niveaux d'expression de ces quatre protéines sont très faibles dans les cellules beta de souris *in vivo*, suggérant que ce changement d'expression n'est peut-être pas significativement enrichissant pour la définition du rôle de RFX6 dans la cellule beta.

Top Canonical Pathways		
Name	p-value	Overlap
Axonal Guidance Signaling	4,27E-05	6,6 % 28/423
Agranulocyte Adhesion and Diapedesis	7,79E-05	9,2 % 15/163
Wnt/-catenin Signaling	8,95E-05	9,1 % 15/165
Human Embryonic Stem Cell Pluripotency	1,17E-04	9,8 % 13/132
Hepatic Fibrosis / Hepatic Stellate Cell Activation	1,84E-04	8,5 % 15/176

Molecular and Cellular Functions		
Name	p-value	#Molecules
Cellular Movement	5,05E-03 - 7,62E-12	102
Cellular Growth and Proliferation	5,03E-03 - 4,46E-09	148
Cellular Development	5,02E-03 - 6,08E-07	136
Cell Morphology	4,80E-03 - 2,65E-06	111
Cell Death and Survival	5,04E-03 - 3,61E-06	117

Physiological System Development and Function		
Name	p-value	#Molecules
Cardiovascular System Development and Function	5,14E-03 - 1,17E-13	99
Organismal Development	5,36E-03 - 1,17E-13	189
Embryonic Development	5,09E-03 - 4,00E-08	140
Tissue Development	5,36E-03 - 6,56E-08	158
Respiratory System Development and Function	1,52E-03 - 3,61E-07	36

Top Analysis-Ready Molecules	
Fold Change up-regulated	
Molecules	Exp. Value
VWF	↑ 8,078
NNMT	↑ 7,143
C1QTNF9B	↑ 6,445
COL18A1	↑ 6,443
DPYSL3	↑ 6,262
RTKN2	↑ 6,156
GZMK	↑ 6,121
ESR1	↑ 6,088
PTGIS	↑ 5,980
DCN	↑ 5,787

Fold Change down-regulated	
Molecules	Exp. Value
CPS1	↓ -84,113
FABP2	↓ -35,912
MLXIPL	↓ -17,664
Muc4	↓ -13,756
ARMC4	↓ -9,748
Z210407C18Rik	↓ -9,236
CDX2	↓ -7,608
DCDC5	↓ -7,535
Gm16308	↓ -7,367
PCK1	↓ -6,902

Top Networks	
ID Associated Network Functions	Score
1 Digestive System Development and Function, Organismal Development, Developmental Disorder	37
2 Cardiovascular System Development and Function, Organismal Development, Lipid Metabolism	34
3 Endocrine System Disorders, Gastrointestinal Disease, Metabolic Disease	30
4 Cell Signaling, Molecular Transport, Nucleic Acid Metabolism	29
5 Cell Signaling, Carbohydrate Metabolism, Molecular Transport	27

Figure 47 : Analyse Ingenuity des 633 gènes ciblés dans les îlots pancréatiques et régulés dans la cellule beta par RFX6. (<http://www.ingenuity.com/products/ipa>)

Enfin, la construction de réseaux fonctionnels par IPA permet l'identification de fonctions cellulaires principalement enrichies dans les gènes ciblés dans les îlots pancréatiques et régulés dans la cellule beta par RFX6. Les réseaux fonctionnels construits regroupent ainsi les fonctions fondamentales de développement et de fonction du système digestif, développement de l'organisme, trouble du développement (réseau 1),

développement et fonction du système cardiovasculaire, développement de l'organisme, métabolisme des lipides (réseau 2), troubles du système endocrine, maladies gastro-intestinales, maladies métaboliques (réseau 3), signalisation cellulaire, transport moléculaire, métabolisme des acides nucléiques (réseau 4) et signalisation cellulaire, transport moléculaire, et métabolisme des acides nucléiques (réseau 5).

Le réseau numéro 3, qui est centralisé autour des troubles du système endocrine, des maladies gastro-intestinales et des maladies métaboliques est illustré en **Figure 48**. Il comprend notamment *Pdgfra* étudié pour son rôle dans la prolifération des cellules beta (Chen et al, 2011) et identifié comme gène dit « disallowed » des cellules beta tout comme *Gas1* également présent dans ce réseau (Pullen et al, 2010; Thorrez et al, 2011). Le réseau numéro 2 qui est centralisé sur les fonctions de développement et fonction du système cardiovasculaire, développement de l'organisme et métabolisme des lipides est illustré en **Figure 49**. Celui-ci possède notre gène d'intérêt *Mxipl* (**étudié dans le paragraphe Résultats B**) qui, de manière étonnante, est le seul gène activé par RFX6 dans ce réseau.

Réseau 3 : Troubles du système endocrine, Maladie gastrointestinales, Maladies métaboliques

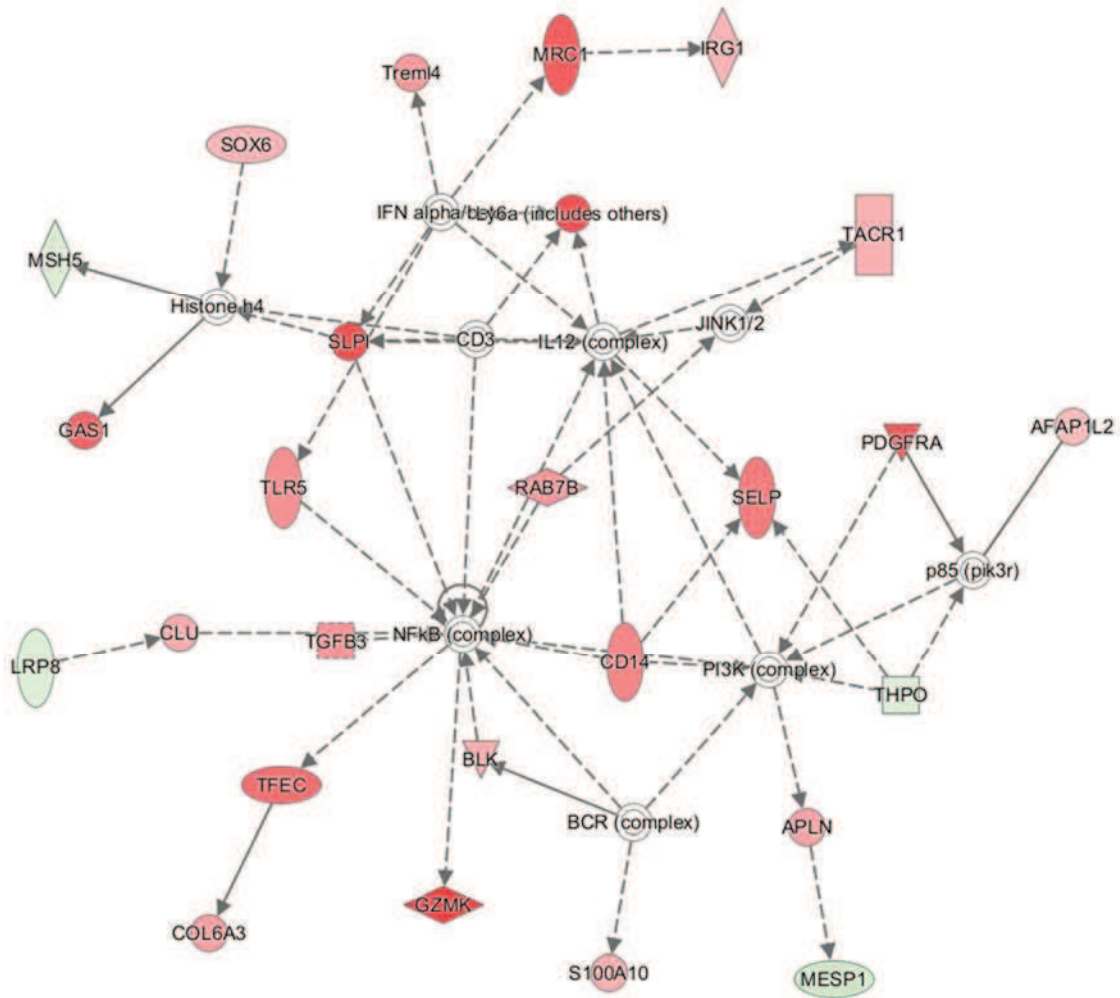


Figure 48 : Réseau fonctionnel regroupant les gènes ciblés dans les îlots pancréatiques et régulés dans la cellule beta par RFX6 et impliqués dans les troubles du système endocrine, les maladies gastro-intestinales, et les maladies métaboliques. Les gènes de ce réseau fonctionnel indiqués en rouge sont réprimés par RFX6 alors que ceux en vert sont activés par RFX6. Les gènes indiqués sur fond blanc ne sont pas directement régulés par RFX6. Ce réseau fonctionnel a été construit avec l'aide d'Ingenuity Pathway Analysis <http://www.ingenuity.com/products/ipa>. Les interactions directes (traits pleins) et indirectes (traits pointillés) entre les différentes protéines sont représentées sur ce schéma.

Réseau 2 : Développement et fonction du système cardiovasculaire, Développement de l'organisme, Métabolisme des lipides

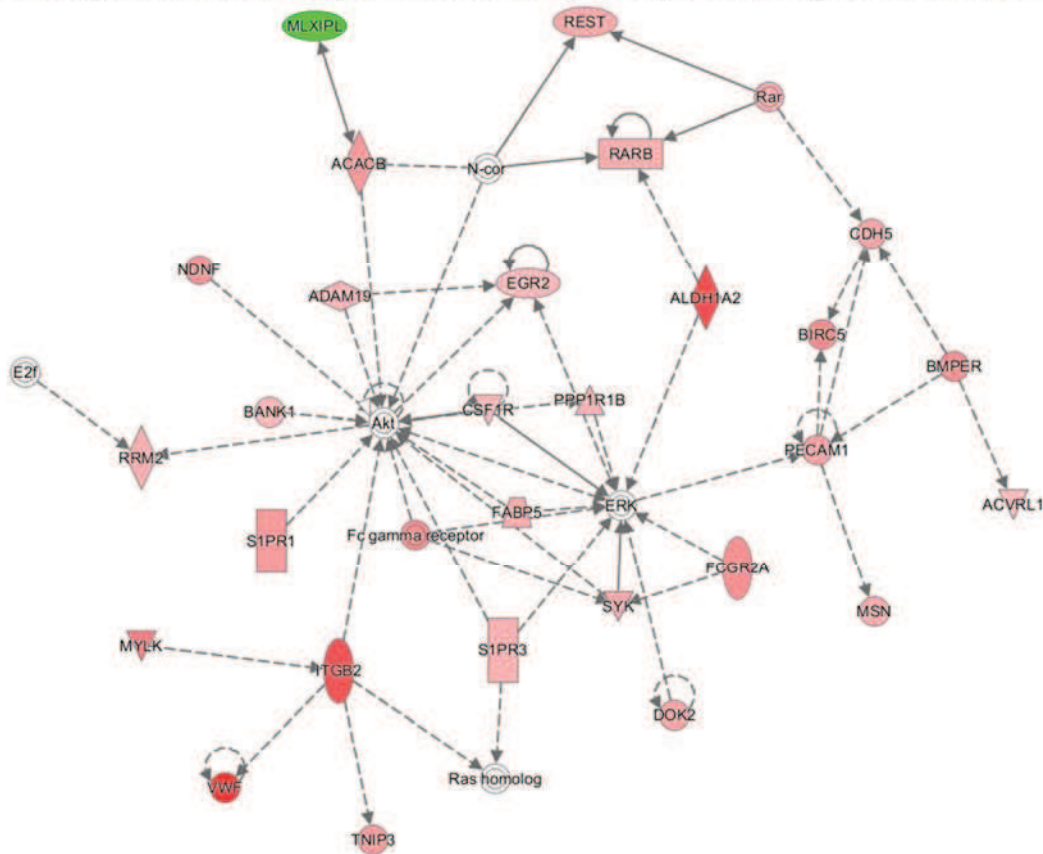


Figure 49 : Réseau fonctionnel regroupant les gènes ciblés dans les îlots pancréatiques et régulés dans la cellule beta par RFX6 et impliqués dans le développement et la fonction du système cardiovasculaire, le développement de l'organisme, et le métabolisme des lipides. Tous les gènes de ce réseau fonctionnel indiqués en rouge sont réprimés par RFX6 alors que MLXIPL (en vert) est quant à lui, activé par RFX6. Les gènes indiqués sur fond blanc ne sont pas directement régulés par RFX6. Ce réseau fonctionnel a été construit avec l'aide d'Ingenuity Pathway Analysis <http://www.ingenuity.com/products/ipa>. Les interactions directes (traits pleins) et indirectes (traits pointillés) entre les différentes protéines sont représentées sur ce schéma.

Ces analyses fonctionnelles n'ont ainsi hélas pas permis l'identification de fonctions moléculaires et cellulaires précises et hautement enrichies, liées aux gènes ciblés dans les îlots pancréatiques et régulés dans la cellule beta par RFX6, et pertinentes dans les fonctions pancréatiques.

Les données de CHIP-Seq dans les îlots pancréatiques de souris et dans le modèle cellulaire beta Min6b1 ont ainsi permis l'analyse globale des sites de fixation de RFX6, et

montrent que RFX6 se fixe de manière préférentielle dans les régions intergéniques ou introniques du génome mais également au niveau des promoteurs des gènes. Les sites de fixation de Rfx6 sont regroupés autour des sites d'initiation de la transcription des gènes. Différentes analyses *de novo* du motif de fixation de RFX6 sur l'ADN ont aussi permis de révéler un motif consensus de type « RFX » centralement enrichi sous le sommet des pics de CHIP-Seq (**Figure 40**). Par compilation des données de transcriptomique dans la souris n'exprimant pas Rfx6 spécifiquement dans la cellule beta et des gènes ciblés par RFX6 dans les ilots pancréatiques, les gènes directement régulés par RFX6 ont pu être mis en avant ainsi que son rôle fondamental dans la fonction et le maintien de l'identité de la cellule beta sécrétrice d'insuline.

B. Mlxipl est une cible principale de RFX6 dans la cellule beta adulte

Pour identifier les gènes directement régulés par RFX6, une comparaison globale des gènes ciblés par RFX6 avec les gènes sous- ou surexprimés dans les cellules beta de souris n'exprimant pas Rfx6 spécifiquement dans ces cellules, a été réalisée et l'un des candidats majeur mis en avant par ces études est le facteur de transcription Mlxipl qui est fortement sous-exprimé (17,7x) dans les cellules beta pancréatiques de la souris delta beta. Mlxipl avait déjà été mis en avant dans la même analyse réalisée dans le modèle cellulaire beta Min6b1. Bien que le rôle de Mlxipl soit bien défini dans le foie, moins de choses sont connues concernant son rôle dans le pancréas et plus particulièrement dans les cellules productrices d'insuline. Nous avons choisi d'étudier plus précisément la régulation de Mlxipl par RFX6 et d'essayer de comprendre dans quelle mesure Mlxipl participe au phénotype observé chez la souris lorsque Rfx6 est délété.

B.1. Mlxipl est exprimé dans différents organes

Mlxipl est un facteur de transcription à hélice-boucle-hélice basique et glissière à leucine appartenant à la superfamille Myc/Max/Mad (Grandori et al, 2000). Il porte divers noms dont MLXIPL (Mlx Interacting Protein Like), ChREBP (Carbohydrate Response Element Binding Protein), ou encore WBSR14 (Williams Beuren Syndrome Critical Region 14) et a été associé au Syndrome de Williams Beuren (OMIM#194050) qui consiste en un trouble du développement neuronal caractérisé par différentes affections dont une intolérance au glucose et un diabète chez 75% des patients (Poerber, 2010; Poerber et al, 2010). Ce syndrome est causé par diverses délétions dans la région chromosomique 7q11.23 chez l'homme (de Luis et al, 2000).

Le locus de Mlxipl représente 33kb d'ADN génomique et le gène possède 17 exons chez l'homme et la souris. Un transcrit codant de 3,6kb est produit chez la souris (ENSMUST0000005507.6). Mlxipl est exprimé de façon ubiquitaire pendant le développement ainsi qu'à l'âge adulte, principalement dans le rein, le muscle, le tissu adipeux, le foie, l'intestin grêle, et dans le pancréas (Cairo et al, 2001; de Luis et al, 2000; Iizuka et al, 2004; Wang & Wollheim, 2002).

L'interaction de MLXIPL avec la protéine MLX (Max Like protein X) en hétérodimère (Stoekman et al, 2004) est nécessaire à leur fonction de transactivation des gènes dont l'expression dépend du glucose. Cet hétérodimère lie ainsi les motifs de type E-box dans l'ADN pour réprimer la transcription des gènes ciblés (Cairo et al, 2001) au niveau de régions ChoRE (Carbohydrate Response Element). Comme le suggère Shih et al. (Shih et al, 1995),

ChREBP joue son rôle par liaison à deux motifs ChoRE (CACGTG) séparés de 5 nucléotides. En 2011, une étude par ChIP Seq a permis l'identification du motif (CABGTG-nnCnG-nGnSTG) comme site de liaison à l'ADN de MLXIPL dans la totalité des gènes cibles dans le foie humain, et corrèle ce motif (Ma et al, 2006) avec la séquence ChoRE consensus précédemment définie (CAYGNG-n5-CNCRTG) (Jeong et al, 2011) pour Mlxipl.

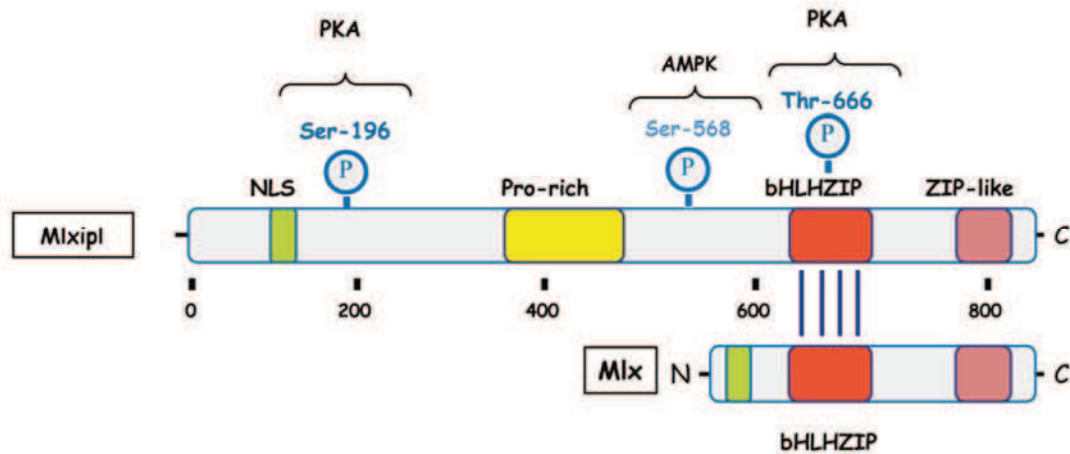


Figure 50 : La protéine MLXIPL interagit avec la protéine MLX pour réguler ses gènes cibles. La phosphorylation des Sérines 196 et 568 et de la Thréonine 666 inhibe l'activité de Mlxipl à faible concentration en glucose. (bHLHZIP (basic loop-helix-leucine-zipper) ; Pro-rich (proline rich domain) ; NLS (nuclear localization signal) ; ZIP-like (leucine zipper like motif)). Adapté de (Postic et al, 2007)

B.2. Régulation de l'activité de MLXIPL en fonction du glucose

Le contrôle de l'activité du facteur de transcription MLXIPL est régulé par sa localisation intracellulaire. En effet, en présence de fortes concentrations en glucose, celui-ci est dirigé vers le noyau où il peut activer ou réprimer la transcription de ses gènes cibles.

L'un des gènes ciblé et transactivé par Mlxipl dans le foie est la L-pyruvate kinase (LPK). Son action est régulée par son état de phosphorylation et rehaussée par une forte concentration en glucose (Yamashita et al, 2001). Mlxipl contient au moins 3 sites de phosphorylation par PKA (cAMP-dependent protein kinase) et deux d'entre eux jouent un rôle important dans l'inhibition de l'activation de LPK dépendante au glucose médiée par l'AMPc. La phosphorylation de la Ser196 proche du NLS (signal de localisation nucléaire) inhibe l'import nucléaire de la protéine. La phosphorylation de la Thr666 abolit sa capacité de liaison à l'ADN ainsi que son activité de transactivation de LPK (Kawaguchi et al, 2001;

Yamashita et al, 2001). D'autre part, un autre site de phosphorylation par l'AMPK (AMPc-activated protein kinase) au niveau de la Ser568 permet l'inhibition de la capacité de liaison à l'ADN de Mlxipl et de la régulation positive de la transcription de LPK (Kawaguchi et al, 2002a). Ces études montrent que l'AMPc active PKA et l'AMPK, qui phosphorylent et donc inhibent Mlxipl qui ne peut plus activer LPK quand la concentration en glucose est faible. A haute concentration en glucose, le Xylulose-5-phosphate (X5P) active PP2A (protein Phosphatase 2A) (Kabashima et al, 2003) qui déphosphoryle Mlxipl et l'active. Par l'analyse de divers mutants de Mlxipl, Tsatsos et al. (Tsatsos & Towle, 2006) ont vérifié ce modèle et suggèrent qu'un second mécanisme indépendant de la phosphorylation par ces protéines kinases mais qui pourrait être en lien avec un autre domaine structurel de la protéine, est impliqué dans sa régulation.

Plus récemment, et contrairement aux modèles proposés jusque-là, il a été proposé que PP2A n'est pas absolument nécessaire à l'activation de Mlxipl en réponse au glucose, et que la déphosphorylation de la Ser196 n'est pas suffisante à la translocation nucléaire de Mlxipl sans stimuli de glucose. Enfin, cette étude démontre le rôle essentiel du Glucose-6-phosphate (G6P) et non du X5P dans l'activation de Mlxipl en réponse au glucose dans les cellules du foie (Dentin et al, 2012).

La régulation de l'activité de Mlxipl par le contrôle de sa localisation intracellulaire a lieu également via son interaction avec différents partenaires tels que les protéines 14-3-3 et NIF3L1 (Merla et al, 2004). Cette publication suggère également que la régulation du métabolisme du glucose par Mlxipl supporte l'hypothèse que le phénotype diabétique observé chez les patients atteints du Syndrome de Williams Beuren pourrait être lié à l'hémizygotie de ce gène entraînant une réduction de sa réponse fonctionnelle (régulation d'enzymes impliquées dans la synthèse des triglycérides à partir d'un excès de glucides) aux fortes concentrations en glucose.

La translocation nucléaire de Mlxipl n'est cependant pas suffisante à la stimulation de son activité dépendante au glucose. Mlxipl possède un module de détection du glucose qui lui permet d'activer ses gènes cibles en présence d'une forte concentration en glucose. Il est composé d'un domaine LID (Low-glucose Inhibitory Domain) inhibant spécifiquement à faible concentration en glucose le second domaine GRACE (Glucose Response Activation Conserved Element) possédant l'activité de transactivation, et permettant ainsi une activation spécifique de Mlxipl uniquement en présence de fortes concentrations de glucose (Davies et al, 2008; Li et al, 2006a). Cette activité, qui a été étudiée dans la lignée cellulaire beta INS-1E, et qui est indépendante du changement de localisation cellulaire de Mlxipl, semble ainsi impliquée dans la régulation de l'expression des gènes sensibles au glucose.

La régulation de l'expression de Mlxipl en fonction du glucose est donc complexe et permet par ce biais un contrôle fin de son activité dépendant de la disponibilité du glucose pour l'organisme.

B.3. Rôle de MLXIPL dans le métabolisme des lipides et du glucose

Comme indiqué auparavant, Mlxipl est fortement exprimé dans le foie et le tissu adipeux, organes dans lesquels ce gène a été le plus étudié. Le foie permet de maintenir l'homéostasie de l'énergie d'un organisme en réponse à son régime. Lorsque la nutrition est riche, l'excès de glucides est converti en triglycérides dans le foie via la lipogenèse et par l'induction des enzymes lipogéniques et glycolytiques telles que la pyruvate kinase (LPK), l'acétyl-coenzyme A ou l'acide gras synthase (FAS). L'alimentation riche en sucres promeut l'activation d'enzymes nécessaire pour la conversion du glucose en acides gras dans l'organisme. Parmi les enzymes de la lipogenèse, dont l'expression est augmentée à haute concentration en glucose, on trouve la pyruvate kinase (LPK), l'acide gras synthase (FAS), l'acétyl coA carboxylase 1 (ACC1) et l'acide malique. Lorsque Mlxipl est inactivé, l'expression de ces enzymes de la lipogenèse est réduite (Iizuka et al, 2004). Dans le foie de rat et dans des hépatocytes primaires, Yamashita et al. ont en effet montré trois ans auparavant que Mlxipl cible et transactive l'expression du gène L-pyruvate kinase (LPK) dans le foie et que son action est régulée par son état de phosphorylation et rehaussée par une forte concentration en glucose (Yamashita et al, 2001).

L'étude de souris n'exprimant pas Mlxipl a permis de mettre en évidence son rôle clé dans le métabolisme du glucose et des lipides. En effet, les souris KO pour Mlxipl sont viables mais présentent des taux de glucose et d'insuline plasmatiques plus élevés que les souris sauvages suggérant une glycolyse hépatique réduite. Le stockage du glucose en triglycéride est donc diminué ainsi que les voies de lipogenèse et de glycolyse hépatique. Sous un régime alimentaire riche en glucose et fructose, les souris KO Mlxipl présentent un défaut de synthèse des acides gras qui provoque un manque d'énergie et de production de chaleur induisant le développement d'une hypothermie et un taux de plus de 50% de mort en moins d'une semaine. Sous un régime riche en sucres complexes, une intolérance au glucose et plus faiblement à l'insuline apparaît chez les souris KO pour Mlxipl. Cette étude conclut que Mlxipl est le facteur de transcription clé régulant l'induction de la lipogenèse et de la glycolyse hépatique en présence de fortes concentrations en sucres et permet ainsi leur conversion en graisse dans le foie (Iizuka et al, 2004).

Par comparaison transcriptomique d'hépatocytes cultivés sous faible ou haute concentration en glucose et en présence ou non d'un dominant négatif de MLX (empêchant

le dimère MLX-MLXIPL de lier l'ADN), les gènes régulés par celui-ci ont été mis en avant (Ma et al, 2006). Cette étude confirme que Mlxipl cible directement tout le répertoire des gènes de la lipogenèse hépatique. De même, en 2011, l'identification des gènes ciblés et régulés par Mlxipl dans une lignée cellulaire hépatique confirme à échelle génomique son rôle clé dans le métabolisme des lipides et démontre également son implication dans les voies de transport, du développement et de la motilité cellulaire. De plus, parmi les gènes ciblés par Mlxipl, 47% sont surexprimés à haute concentration en glucose et 53% sont sous-exprimés, attestant l'importance de Mlxipl dans la régulation directe des gènes sensibles au glucose (Jeong et al, 2011).

Par ailleurs, un lien entre Mlxipl et obésité est révélé par l'étude de Iizuka et al. en 2006 qui montre par croisement de souris obèses (leptin- ob/ob) et déficientes pour Mlxipl, que celui-ci joue un rôle prédominant dans l'allègement du phénotype diabétique de ces souris. Une réduction de la lipogenèse et une augmentation de la tolérance au glucose et du contrôle de l'appétit est observé et suggère ainsi qu'une réduction d'activité de Mlxipl pourrait avoir un effet bénéfique dans le traitement de l'obésité (Iizuka et al, 2006). Enfin, chez des patients obèses, l'expression de Mlxipl est augmentée dans le foie et diminuée dans le tissu adipeux par rapport à des patients sains, ce qui est compatible avec l'augmentation de la lipogenèse hépatique et la diminution de la lipogenèse adipocytaire retrouvée chez ces patients. D'autre part, cette étude révèle une augmentation d'expression de Mlxipl lors de la différenciation de pré-adipocytes, suggérant un rôle de ce facteur de transcription dans l'adipogenèse également (Hurtado del Pozo et al, 2011). La mauvaise assimilation du glucose par les organes périphériques est l'un des symptômes précoces du diabète de type 2. Lorsque le gène Glut4, qui permet l'assimilation du glucose par le tissu adipeux et les muscles sous stimuli d'insuline, est sous-exprimé, le développement d'un diabète précoce caractérisé par une perturbation de la sensibilité à l'insuline est observé (Abel et al, 2001). La régulation de Mlxipl par GLUT4 est décrite en 2012 (Herman et al, 2012). Cette étude révèle également que la régulation de l'expression de Mlxipl par le glucose provoque l'activation de deux isoformes alpha et beta du gène, qui diffèrent par un promoteur alternatif. La diminution de l'expression spécifique de Mlxipl beta et sa corrélation avec une résistance à l'insuline a permis ainsi aux auteurs de proposer cette nouvelle isoforme comme une cible pour traiter le diabète (Herman et al, 2012).

Un autre organe clé impliqué dans le développement du diabète et contrôlant la régulation de l'homéostasie du glucose est le pancréas dans lequel Mlxipl est également exprimé.

B.4. MLXIPL et le pancréas

Le niveau d'expression de Mlxipl est identique dans le foie et le pancréas chez l'homme et la souris (Metukuri et al, 2012). Sa régulation et sa fonction ont donc également été étudiées dans le pancréas.

Contrôle de l'activité de Mlxipl dans le pancréas

Comme indiqué précédemment dans le paragraphe Résultats B.2, une controverse concernant le métabolite signal de l'activation de Mlxipl dans le foie existe entre le G6P et le X5P. En 2010, c'est le rôle clé du G6P dans l'activation de Mlxipl dans la lignée cellulaire INS-1E beta qui est démontré (Li et al, 2010). Plus récemment, ce sont les rôles du G6P et du X5P dans la régulation de Mlxipl dans une lignée cellulaire beta INS-1E qui ont été précisés (Iizuka et al, 2013). Les auteurs utilisent dans cette étude un analogue du glucose, le 2-deoxy-glucose, qui lorsqu'il est catabolisé ne produit que du G6P contrairement au glucose qui lui est catabolisé en G6P mais également en X5P permettant ainsi d'étudier le rôle précis du G6P dans les cellules beta. La relation directe entre le G6P et la fonction de transactivation de Mlxipl n'a cependant pas pu être démontrée. De même le rôle du X5P est étudié par surexpression de la xylulokinase (XYLB) qui permet la conversion du xylitol en X5P dans les cellules beta. Dans ces conditions, l'expression des gènes cibles de Mlxipl est stimulée dans les cellules INS1-E mais toujours dans une moindre mesure par rapport à la stimulation dépendante du glucose.

Ces résultats, associés à ceux de Li et al. (Li et al, 2010), montrent que dans la cellule beta ce sont à la fois le G6P et le X5P qui régulent l'induction de Mlxipl en réponse au glucose.

Comme cité dans le paragraphe précédent (**Résultats B.2**), la régulation de l'activité de Mlxipl est également contrôlée par sa localisation intracellulaire. Par une étude dans la lignée cellulaire beta Min6b1, Noordeen et al. ont identifié la sorcine comme une ancre cytoplasmique pour Mlxipl dans les cellules beta. En effet, à haute concentration en glucose, suite à l'influx de calcium dans la cellule beta, un changement conformationnel de la sorcine a lieu et permet le relâchement de Mlxipl qui entre dans le noyau et active l'expression de ses gènes cibles (Noordeen et al, 2012).

Tout comme il a été montré dans le foie, Mlxipl cible les gènes de la lipogenèse dans les cellules beta. Sa translocation nucléaire, ainsi que la régulation des enzymes LPK et FAS notamment, sont mis en évidence dans la lignée cellulaire beta Min6 sous stimuli de glucose induisant une augmentation du contenu en acide gras. Il est également prouvé que la

sécrétion d'insuline en réponse au glucose est augmentée en absence de Mlxipl (da Silva Xavier et al, 2006). Un effet inverse est observé lorsque Mlxipl est surexprimé dans la lignée cellulaire beta INS-1 et dans les ilots de souris où un diabète apparaît (Poungvarin et al, 2012).

Mlxipl et la cellule beta sécrétrice d'insuline

Mlxipl est aussi un acteur clé de la fonction de la cellule beta adulte puisque son inactivation stimule l'expression de gènes clés tels que Pdx1, MafA, Gck et l'insuline (da Silva Xavier et al, 2010), suggérant qu'une surexpression de Mlxipl pourrait être à l'origine d'un phénotype aggravé de diabète de type 2 et qu'une inactivation de Mlxipl pourrait probablement permettre le développement d'un traitement antidiabétique. L'identification du réseau des gènes impliqués dans la préservation de la sécrétion d'insuline en réponse au glucose dans le modèle cellulaire beta Min6 a également mis en évidence le rôle clé de Mlxipl dans cette fonction (Yamato et al, 2013).

D'autre part, Mlxipl réprime l'expression de ARNT/HIF1 beta, acteur clé de la fonction de la cellule beta, dont l'expression diminue chez les patients diabétiques de type 2 et est nécessaire à la sécrétion d'insuline en réponse au glucose (Gunton et al, 2005; Noordeen et al, 2010). En effet, par recherche des gènes cibles de Mlxipl dans les cellules Min6, les auteurs identifient ARNT/HIF1 beta. Ils confirment qu'à forte concentration en glucose ou lorsque Mlxipl est surexprimé, l'expression de ARNT/HIF1 beta diminue dans les cellules beta et qu'en absence de Mlxipl, son expression augmente (Noordeen et al, 2010). Mlxipl régule aussi COUP-TFII qui est essentiel au métabolisme des cellules beta et à leur sécrétion d'insuline en réponse au glucose (Perilhou et al, 2008b). La régulation du facteur de transcription COUP-TFII est dépendante du glucose et de l'insuline et implique Mlxipl et Foxo1 (Perilhou et al, 2008a) dans le pancréas et le foie. Mlxipl joue également un rôle clé dans le métabolisme des acides gras par la répression du facteur de transcription PPAR alpha en réponse au glucose dans les cellules beta pancréatiques (Boergesen et al, 2011). De plus, en présence de fort taux de lipides, PPAR alpha a été proposé comme un protecteur de la fonction des cellules beta et l'identification de sa régulation par Mlxipl renforce l'idée d'un traitement antidiabétique (da Silva Xavier et al, 2010; Iizuka et al, 2006) basée sur une inactivation de ce dernier.

Par ailleurs, les mécanismes sous-jacents au phénomène de glucotoxicité semblent impliquer Mlxipl. La glucotoxicité consiste en une hyperglycémie chronique entraînant une altération des cellules beta sécrétrices d'insuline. En effet, une étude de gain de fonction de Mlxipl dans la lignée cellulaire beta INS-1, mais également dans les ilots pancréatiques de

souris et humains, montre que la surexpression de Mlxipl récapitule le phénotype de glucotoxicité notamment par activation des voies de la lipogenèse, du stress oxydatif, de l'apoptose et par perturbation de la synthèse et de la sécrétion d'insuline en réponse au glucose (Poungvarin et al, 2012). L'action de médiateur de Mlxipl sur la glucotoxicité, mais également sur la lipotoxicité, renforce encore son rôle dans le développement de l'insulinorésistance et du DNID.

Enfin, la régulation du gène préapoptotique Txnip par Mlxipl a été révélée en 2009 par Cha-Molstad et al. (Cha-Molstad et al, 2009). Ce gène avait préalablement été identifié comme stimulant la mort cellulaire beta induite, la glucotoxicité et la perte de masse des cellules beta observée chez les patients diabétiques (Chen et al, 2008a; Chen et al, 2008b; Minn et al, 2005a; Minn et al, 2005b), il est également lié à une hypertriglycéridémie et une intolérance au glucose chez l'homme (Zhao et al, 2015). Dans cette étude de 2009, Mlxipl est identifié comme le facteur de transcription responsable de l'induction d'expression de Txnip en réponse au glucose dans les cellules beta murines et humaines. Les modifications de chromatine associées à cet évènement ont été caractérisées et révèlent une interaction entre Mlxipl et p300 permettant l'acétylation de l'histone H4 de Txnip et sa transcription. Plus récemment, la régulation négative d'expression de Txnip par Foxo1 a été mise en évidence et suggère que Foxo1 et Mlxipl rivalisent pour réguler l'expression de ce gène dans les cellules beta pancréatiques (Kibbe et al, 2013) et que cette association n'est pas limitée à ce seul gène mais touche également Lpk, un autre gène cible de Mlxipl.

En 2012, l'étude de Soggia et al. (Soggia et al, 2012) a montré que le glucose via Mlxipl régule aussi la différenciation des cellules beta et pointe ainsi un éventuel rôle de ce gène lors du développement. Pendant l'embryogenèse, l'expression de Mlxipl est enrichie dans le lignage endocrine et provoque l'induction de NeuroD1 (cible de Ngn3) lors du développement. Son expression augmente tout au long de la différenciation des cellules, et par augmentation de l'expression de ce gène en présence de glucose, la différenciation des cellules est stimulée. Cependant, le rôle précis de Mlxipl lors du développement du pancréas n'est pas encore connu.

Des connaissances approfondies au sujet de la prolifération des cellules beta et de son contrôle permettraient une avancée en faveur du développement de thérapies pour le diabète (Kulkarni et al, 2012). En 2012, Metukuri et al. (Metukuri et al, 2012) proposent pour la première fois le rôle essentiel de Mlxipl dans le contrôle de la prolifération des cellules beta en réponse au glucose. Un rôle clé de Mlxipl dans la prolifération de cellules cancéreuses avait déjà été montré (Tong et al, 2009a). Dans cette publication de 2012, les auteurs montrent que l'expression de Mlxipl dans le foie et les îlots pancréatiques est

équivalente. Par des études de perte et de gain de fonction, ils montrent que la prolifération des cellules beta INS1-E en réponse au glucose est affectée en l'absence de Mlxipl et que l'expression de gènes régulateurs du cycle cellulaire, tels que les cyclines Ccnd2, Ccna2, CCne1 ou Cdk4 et Cdk6 est diminuée. L'effet inverse est observé lorsque Mlxipl est surexprimé pour la cycline Ccnd2 et l'expression d'inhibiteurs du cycle cellulaires tels que p57, p15, p18 ou p19 diminue. Mlxipl contrôle ainsi la prolifération des cellules beta en réponse au glucose.

Dans la cellule beta sécrétrice d'insuline, les différentes études citées montrent que Mlxipl est impliqué dans de nombreux processus cellulaires tels que les voies de prolifération, la différenciation, la glucotoxicité ou encore la sécrétion d'insuline. L'identification de Mlxipl comme une cible majeure de RFX6 dans les cellules beta, nous a poussés à étudier plus précisément la régulation de Mlxipl par RFX6 et d'essayer de comprendre dans quelle mesure Mlxipl participe au phénotype observé chez la souris lorsque Rfx6 est délété.

B.5. RFX6 lie le premier intron de Mlxipl

Les résultats de CHIP Seq obtenus dans les paragraphes Résultats A.3 et A.4, sont présentés en **Figure 51A**. RFX6 se lie au niveau de l'intron 1 de Mlxipl dans les cellules Min6b1 mais également *in vivo* dans les ilots pancréatiques.

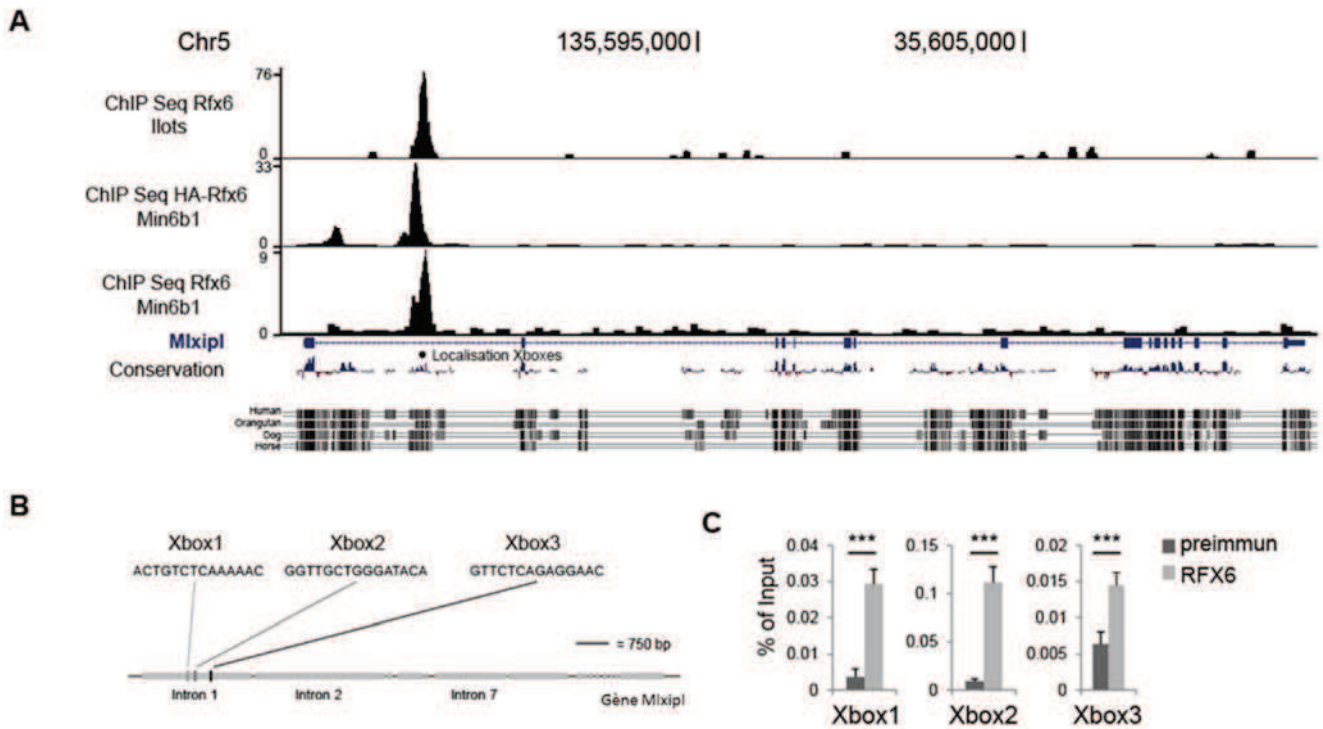


Figure 51 : RFX6 lie trois Xbox localisées dans le premier intron de Mlxipl. (A) Pics de séquences enrichies après ChIP anti-HA-RFX6 dans les cellules Min6b1 transfectées avec pCMVTag2A-Flag-3HA-Rfx6-2A-mCherry (milieu) ou ChIP anti-RFX6 dans les cellules Min6b1 (bas) ou les ilots de Langerhans (haut). **(B)** Représentation des trois motifs Xbox localisés dans le premier intron de Mlxipl. **(C)** Confirmation de liaison de RFX6 aux trois Xbox identifiées en (B) par ChIP qPCR anti-RFX6 réalisé dans les cellules Min6b1, (FC Xbox1 = 7,9 ; Xbox2 = 12,2 ; Xbox3 = 2,3) (n=5 pour Xbox 1 et 2 et n=5 pour Xbox3). Les résultats sont exprimés en moyenne \pm SD *p<0,05 ; **p<0,01 ; ***p<0,001 ; ns : non significatif, FC : variation d'expression.

Lors de la construction du motif *de novo* présent sous les pics de ChIP Seq anti-Rfx6 identifiés dans les cellules Min6b1 (**Figure 29**), c'est le motif Xbox2 qui avait été identifié comme site de liaison de RFX6 dans l'intron 1 de Mlxipl et ce motif est également retrouvé sous le pic de ChIP Seq RFX6 dans les ilots de Langerhans. Par une recherche complémentaire de motifs de type « RFX » avec comme matrice la séquence consensus RYYNYYN(1,3)RRNRAC publiée par Laurençon et al. (Laurençon et al, 2007) en 2007 et à l'aide de l'interface Web Fuzznuc (<http://emboss.bioinformatics.nl/cgi-bin/emboss/fuzznuc>) dans la totalité de la séquence présente sous le pic d'intérêt (ChIP RFX6 dans les Min6b1), deux sites supplémentaires ont également été identifiés : Xbox1 et Xbox3. Ces trois sites de liaison de RFX6 identifiés au niveau de l'intron 1 de Mlxipl sont représentés en **Figure 51B** et

Figure 52. Les Xbox 1 et 3, contrairement à la Xbox 2, ne sont pas présentes sous le pic de ChIP Seq identifié dans les îlots pancréatiques.

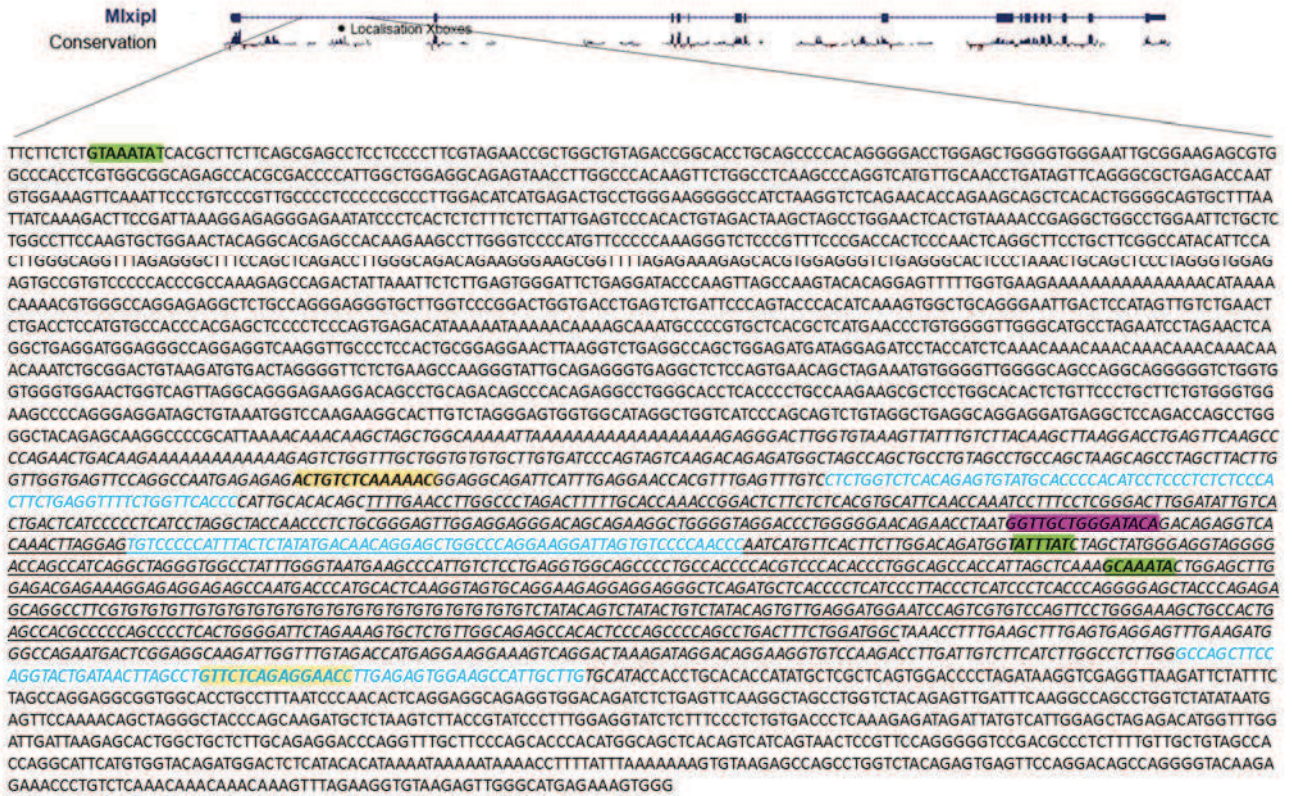


Figure 52 : Séquence nucléotidique de la région où se fixe RFX6 dans l'intron 1 de Mlxipl. L'ADN sous le pic du ChIP RFX6 dans les cellules Min6b1 est représenté en italique et l'ADN sous le pic du ChIP RFX6 dans les îlots pancréatiques est souligné. Les amplicons utilisés pour le ChIP qPCR sont indiqués en bleu. RFX6 se fixe sur les Xbox1 (surlignée en orange), 2 (surlignée en rose) et 3 (surlignée en orange) indiquées dans cette séquence. (sites Foxa1/2 surlignés en vert (Gao et al, 2010))

La fixation significative de RFX6 à ces trois Xbox a également été confirmée par ChIP qPCR (ChIP RFX6 dans les cellules Min6b1) en **Figure 51C**, où une liaison préférentielle de ce FT à la Xbox2 est observée (FC = 12,2). Cette confirmation n'a été réalisée pour l'instant que dans la lignée Min6b1 et devra être testée également dans les îlots pancréatiques. Ces trois sites de liaison de RFX6 sont présents dans une région hautement conservée de l'intron

1 de Mlxipl, où la Xbox 2 est conservée à 100% entre la souris et l'homme, la Xbox3 est conservée à 57% et la Xbox1 n'est pas conservée (**Figure 53**).



Figure 53 : Alignement inter-espèces des Xbox identifiées dans la région de fixation de RFX6 dans l'intron 1 de Mlxipl. Les Xbox d'intérêt sont surlignées en bleu. La Xbox 1 n'est pas conservée alors que les Xbox 2 et 3 présentent respectivement 100% et 57% de conservation entre la souris et l'homme.

Ces résultats suggèrent ainsi que la Xbox 2, localisée dans le premier intron de Mlxipl et identifiée par CHIP Seq RFX6, à la fois dans la lignée cellulaire beta Min6b1 mais également dans les ilots pancréatiques, revêt un aspect essentiel pour la régulation de ce gène par RFX6.

B.6. Rfx6 régule positivement l'expression de Mlxipl dans la cellule beta adulte

Mlxipl est une cible de RFX6 dans les cellules beta et son expression est diminuée de 17,7 fois quand RFX6 n'est plus exprimé dans la souris delta beta (RNA Seq). Les résultats présentés en **Figure 54** confirment que l'inactivation de Rfx6 entraîne une diminution d'expression de Mlxipl. En effet, dans les cellules Min6b1 traitées par siRfx6, la production de protéine RFX6 est affectée et son expression diminue de 45% (**Figure 54A et B**). Et dans ces mêmes conditions, l'expression de Mlxipl diminue de 50% (**Figure 54B à gauche**). La diminution d'expression de Mlxipl suite à l'inactivation de Rfx6 a également été confirmée en qPCR chez la souris delta beta (**Figure 54B à droite**) où une perte quasi complète de Mlxipl est observée (FC = -544).

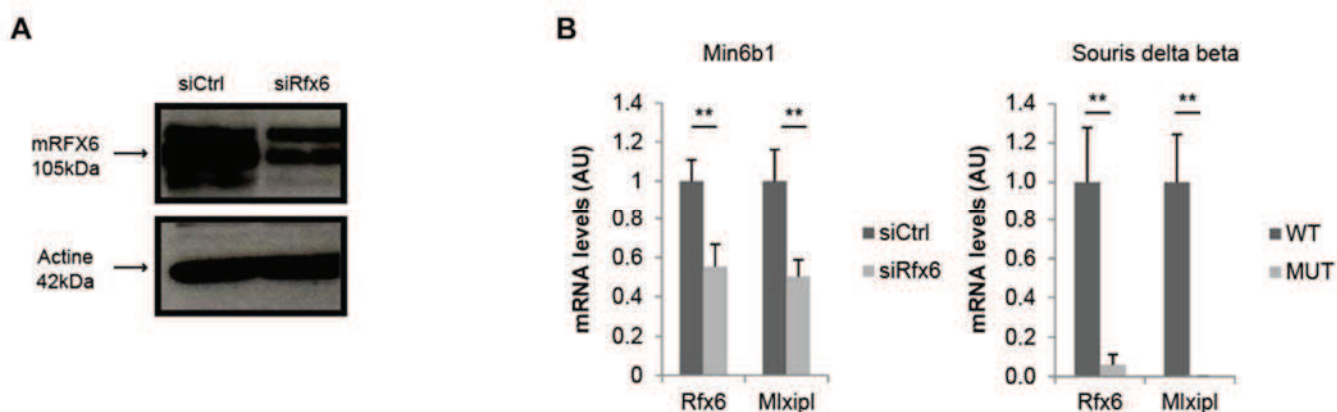


Figure 54 : RFX6 régule positivement l'expression de Mlxipl chez la souris. (A) Western blot anti-Rfx6 et anti-Actine sur 15 µg d'extraits protéique totaux de cellules Min6b1 transfectées par 75nM de siRfx6 ou siCtrl. La production de protéine RFX6 est diminuée dans les cellules Min6b1 traitées par 75nM de siRfx6. **(B)** L'expression de Mlxipl diminue quand l'expression de Rfx6 est réduite dans les cellules Min6b1 et chez la souris. Une diminution d'expression de Rfx6 de 45% dans les cellules Min6b1 48h après traitement par siRfx6 (75nM) entraîne une diminution significative d'expression de 50% de Mlxipl (n=3). Ces résultats sont confirmés *in vivo* chez la souris delta beta (MUT = Rfx6 fl/fl ; Ins1-CrERT) qui n'exprime plus Rfx6 5 jours après injection de tamoxifène (n=4). Les résultats sont exprimés en moyenne ± SD, **p<0,01.

Ces résultats montrent ainsi que RFX6 est un activateur de l'expression de Mlxipl dans les ilots pancréatiques de souris et les cellules Min6b1, probablement par liaison aux sites Xbox identifiés par CHIP Seq (**paragraphe Résultats B.5**). RFX6 lie ainsi le premier intron

de Mlxipl et active son expression. Suite à ces résultats, la fonction de transactivation doit être testée entre les éléments régulateurs de Mlxipl contenant les Xbox identifiées par ChIP-Seq et RFX6 pour la protéine murine mais également humaine.

B.7. La Xbox2 de Mlxipl promeut sa transactivation par la protéine RFX6 de souris et humaine

Pour déterminer si les éléments Xbox identifiés par ChIP Seq dans l'intron 1 de Mlxipl permettent effectivement de transactiver l'expression de ce gène, des tests luciférase (entre les éléments régulateurs de Mlxipl contenant les Xbox identifiées par ChIP-Seq et la protéine RFX6) ont été réalisés.

Premièrement, l'activité de transactivation de RFX6 sur la Xbox2, clonée en quatre copies en amont d'un promoteur minimal et du gène de la luciférase, a été testée. Comme indiquée dans la **Figure 55A et B**, la protéine RFX6 de souris mais également humaine (**dont la stratégie de clonage est présentée dans le paragraphe Résultats B.8.a**) transactive l'expression de Mlxipl via l'élément régulateur que constitue la Xbox2 présente dans son premier intron.

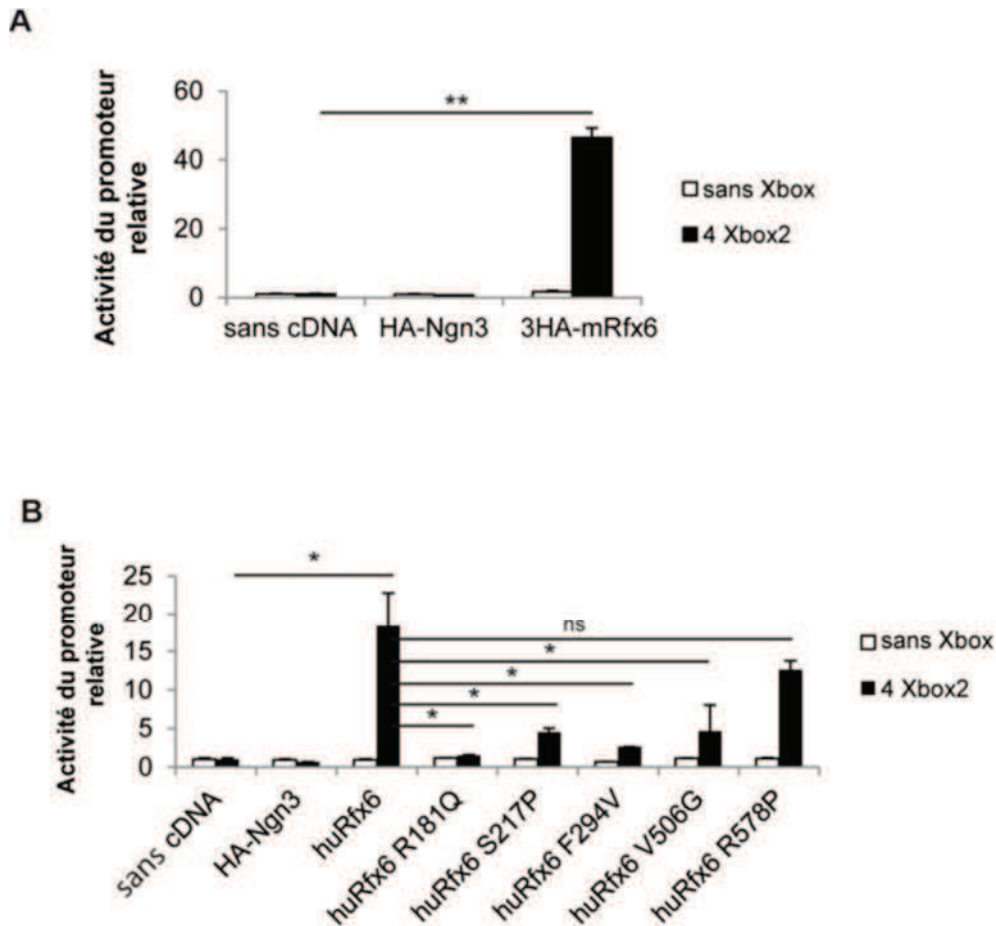


Figure 55 : Rfx6 transactive l'expression de Mlxipl via sa liaison à la Xbox2 localisée dans son premier intron. (A) La protéine RFX6 de souris promeut l'expression de la luciférase par sa liaison à la Xbox2 (n=3, les données sont normalisées par rapport aux valeurs sans ADNc, HA-Ngn3 = contrôle négatif). **(B)** La protéine RFX6 humaine promeut l'expression de la luciférase par sa liaison à la Xbox2 (n=3, les données sont normalisées par rapport aux valeurs sans ADNc, HA-Ngn3 = contrôle négatif). Toutes les protéines RFX6 humaines mutantes montrent une activité significativement réduite (R181Q FC = -12,1 ; S217P FC = -4 ; F294V FC = -6,9 ; V506G FC = -3,9 ; R578P FC = -1,4) sauf la protéine RFX6 humaine mutante R578P. Les résultats sont exprimés en moyenne \pm SD *p<0,05 ; **p<0,01 ; ***p<0,001, FC : variation d'expression.

La Xbox2 de l'intron 1 de Mlxipl peut ainsi être considérée comme site activateur de l'expression de ce gène. Cette fonction de transactivation devra par la suite être testée sur la région complète de l'intron 1 de Mlxipl contenant les trois Xbox identifiées auparavant et décrites dans la **Figure 56**. Des analyses prévues sur cette région complète ou délétaée spécifiquement pour une, plusieurs ou la totalité des Xbox, seront réalisées dans un avenir proche et permettront de spécifier les résultats obtenus concernant la régulation de l'expression de Mlxipl par RFX6, ainsi que l'aspect essentiel de la Xbox2 identifiée par ChIP

Seq RFX6 dans les cellules Min6b1. Par ailleurs, dans les cellules beta pancréatiques, l'expression de Mlxipl est régulée par Foxa1 et Foxa2, qui se lient également au niveau de deux régions conservées dans le premier intron du gène. Une inhibition des deux gènes Foxa1 et 2 dans les ilots pancréatiques induit une perturbation de la sécrétion d'insuline en réponse au glucose ainsi qu'une diminution d'expression de Mlxipl de 13 fois (Gao et al, 2010). La co-transactivation de Mlxipl par RFX6 et FOXA1/2 pourra ainsi également être testée.

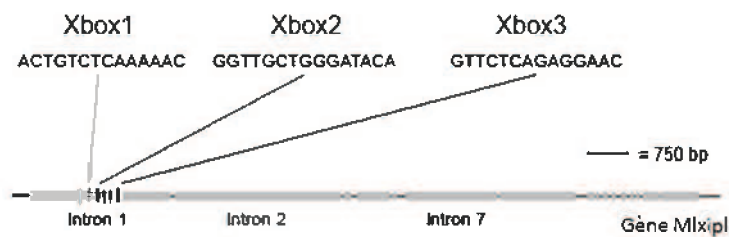


Figure 56 : Représentation des motifs de fixation de facteurs de transcriptions localisés dans le premier intron de Mlxipl. Des sites de fixation de type « RFX » (noir) et « FOXA1/2 » (vert) sont localisés dans le premier intron de Mlxipl.

Ce système permettra par la même méthode de mesurer la fonctionnalité transcriptionnelle de la protéine humaine RFX6 possédant des mutations ponctuelles. L'étude de ces mutations ponctuelles et de leurs conséquences sur la fonction de la protéine RFX6 est présentée dans le paragraphe suivant.

B.8. Les mutations ponctuelles humaines de Rfx6 affectent la fonction de transactivation du gène Mlxipl

Diverses mutations de Rfx6 chez l'homme ont en effet, été identifiées et reliées au syndrome de Mitchell-Riley, caractérisé notamment par un diabète néonatal, une atrésie de l'intestin grêle et une malabsorption intestinale (Paragraphe Introduction D.3). Ces mutations sont représentées dans la **Figure 57**. Pour les mutations S217P et R181Q, des tests de liaison à l'ADN ont déjà été réalisés et montrent que la mutation R181Q, présente au sein du domaine de liaison à l'ADN de RFX6, inhibe totalement la capacité de liaison à l'ADN de la protéine mutée. Alors que la mutation S217P, qui est localisée entre le domaine de liaison à

l'ADN et les domaines de dimérisation de RFX6, n'affecte que très peu la liaison à l'ADN de RFX6 (Smith et al, 2010).

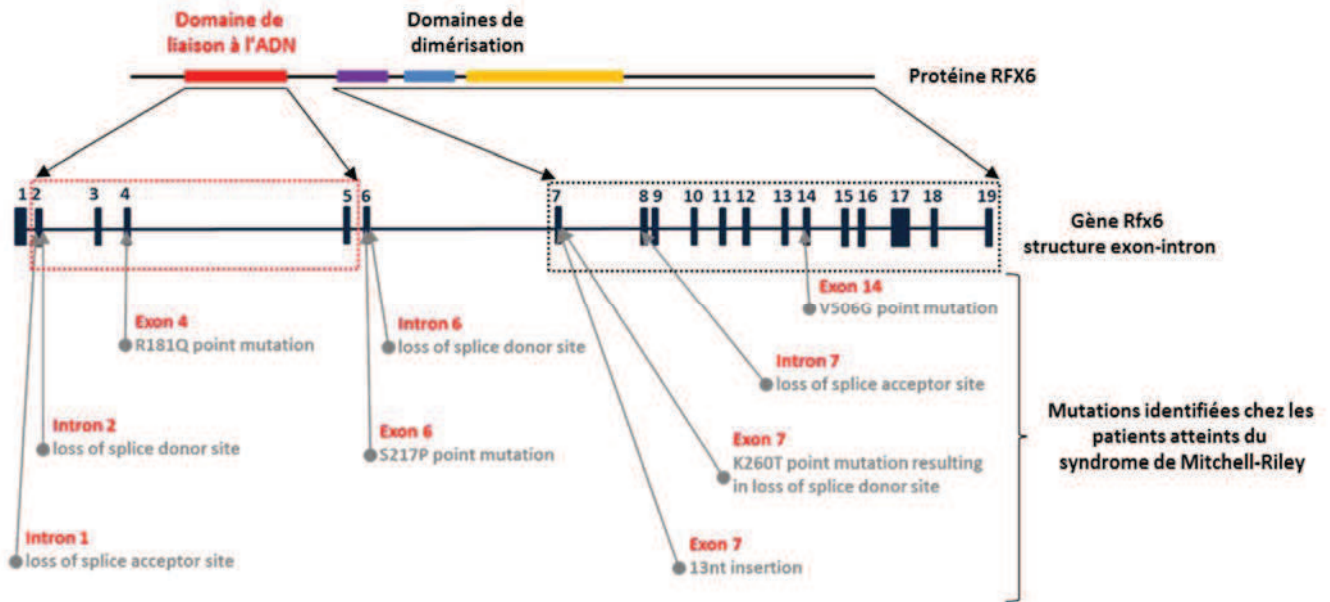


Figure 57 : Mutations du gène Rfx6 humain reliées au syndrome de Mitchell-Riley.

Chaque mutation est décrite et localisée sur la structure du gène Rfx6 humain. Cette figure a été construite à l'aide des différentes publications faisant référence à ces mutations (Chandra et al, 2014; Chappell et al, 2008; Concepcion et al, 2014; Martinovici et al, 2010; Mitchell et al, 2004; Smith et al, 2010; Spiegel et al, 2011). Deux mutations supplémentaires au niveau des exons 17 et 18 publiées en août 2015, ne figurent pas sur ce schéma (Sansbury et al, 2015).

D'autre part, trois nouvelles mutations humaines de Rfx6 ont été identifiées chez des patients diabétiques dans le laboratoire de P. Froguel par le groupe de Martine Vaxillaire à Lille.

Par séquençage direct des 19 exons de Rfx6 chez 22 patients de famille référencée comme atteinte de diabète MODY (diabète apparu avant 30 ans) et 86 patients diabétiques adultes, ce groupe a identifié deux nouvelles mutations hétérozygotes dans ce gène : F294V et R578P. De même, dans une cohorte de 32 patients possédant pour certains des anomalies de développement du pancréas et/ou un diabète apparu pendant l'enfance, une troisième mutation homozygote V506G a également été identifiée chez un patient présentant un diabète néonatal et atteint d'une atresie de l'intestin (jéjunum). Celle-ci a récemment été publiée (Chandra et al, 2014) et étudiée en collaboration avec l'équipe de R. Scharfmann qui

montre que cette mutation est à l'origine d'un défaut de synthèse de l'insuline due à une incapacité de RFX6 à transactiver son expression dans une lignée cellulaire beta humaine.

Dans le but de déterminer le lien causal précis entre sévérité de la pathologie et mutations de Rfx6, la relation entre la structure et la fonction de ces protéines mutées est actuellement en cours d'étude. La poursuite de cet objectif m'a poussée à étudier les conséquences de ces nouvelles mutations sur la fixation à l'ADN, la dimérisation et la fonction de transactivation de ce facteur de transcription. Pour cela, des expériences de retard sur gel ainsi que des tests luciférase entre les cibles identifiées par ChIP-Seq et les protéines Rfx6 mutées ont été envisagés. Ce travail concernant des protéines RFX6 humaines a été réalisé tout au long de ma thèse en collaboration avec Mégane DENU (Stagiaire BTS 1^e année), Emélie LUDMANN (Stagiaire M1), Julia GRANS (Stagiaire M2) et Aline MEUNIER (Ingénieur).

B.8.a. Clonage des gènes Rfx6 humains mutés

Stratégie de clonage

Pour démarrer ces études, une stratégie de clonage a dans un premier temps permis d'obtenir les différents vecteurs d'expression pcDNA3-huRfx6 contenant ces mutations ponctuelles, et leur expression ont été testées dans les cellules humaines HEK293T.

Après avoir réceptionné un vecteur contenant la séquence codante de la protéine humaine RFX6, de la part de Vikash Chandra du laboratoire de Raphaël Scharfmann à Paris, le sous-clonage de huRfx6 dans le vecteur d'expression pcDNA3 a été réalisé (**Figure 58**).

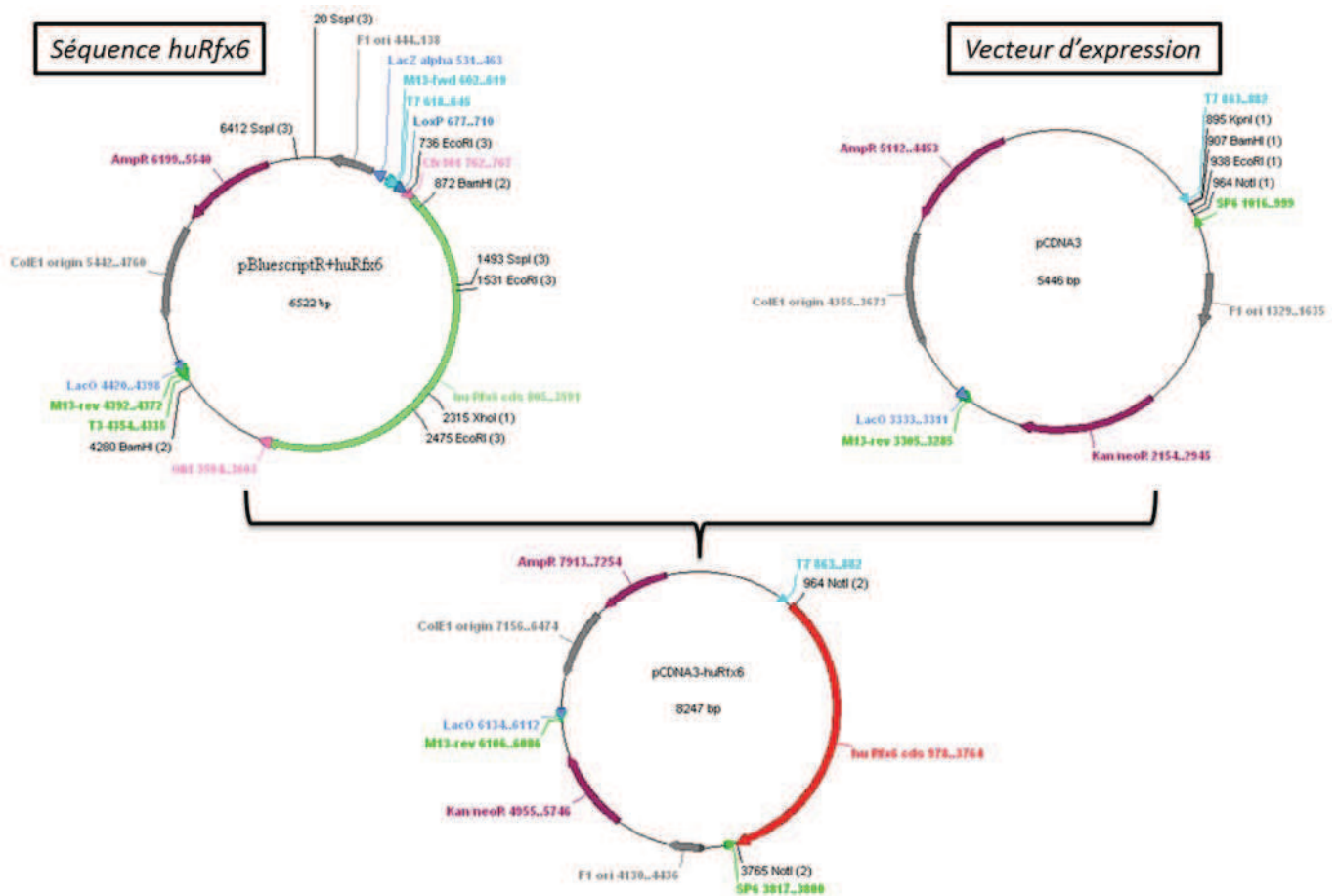


Figure 58 : Production du vecteur d'expression pcDNA3-huRfx6. Après insertion de sites de restriction KpnI et NotI par PCR (à l'aide des amorces indiquées dans le Tableau 6) de part et d'autre de la séquence de huRfx6 (pBSr-hurfx6), un premier sous-clonage de cette séquence a été effectué dans le plasmide PCR®-II-Topo par TA cloning (Life technologies). Enfin, la séquence huRfx6 est insérée au niveau du site NotI du vecteur d'expression pcDNA3, pour obtenir le pcDNA3-huRfx6.

Mutagenèses ponctuelles de la protéine RFX6 humaine

A partir de ce vecteur d'expression pcDNA3-huRfx6, différentes mutations ponctuelles ont par la suite été insérées dans la séquence codante de huRfx6 à l'aide du Kit Quick Change XL Site-directed mutagenesis Kit de chez Agilent, selon les indications du fabricant. Ce sont tout d'abord des mutations publiées et déjà connues qui ont été choisies : R181Q et S217P (Smith et al, 2010), mais également les trois nouvelles mutations identifiées dans le groupe de Martine Vaxillaire : F294V, V506G et R578P (**Figure 59**).



Figure 59 : Mutations ponctuelles de la protéine RFX6 humaine. Les cinq mutations ponctuelles étudiées sont localisées (étoiles roses) sur le schéma de la protéine RFX6 humaine. (DBD = domaine de liaison à l'ADN, B, C et DIM = domaines de dimérisation).

Pour cela, des oligonucléotides spécifiques (**Tableau 6**) ont été désignés à l'aide de l'interface web dédiée (<http://www.genomics.agilent.com/primerDesignProgram.jsp>) et les cinq plasmides pCDNA3-huRfx6 mutants ont été obtenus.

B.8.b. Expression des protéines RFX6 humaines mutées

L'expression des différents gènes a ensuite été testée en cellules de mammifère humaines HEK293T. Les protéines Rfx6 humaines sauvages et mutées sont produites dans les cellules HEK293T (**Figure 60**).

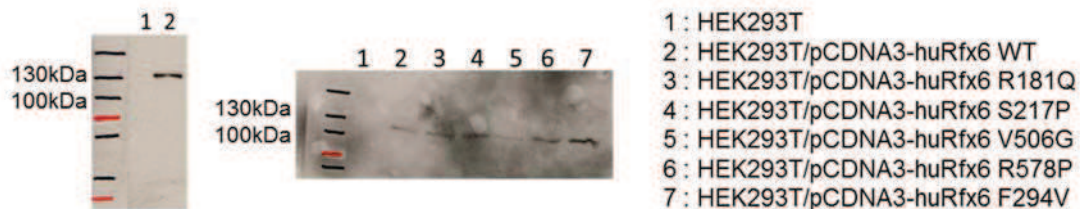


Figure 60 : Expression des protéines huRfx6 sauvage et mutées dans les cellules HEK293T. Western blot anti-huRfx6 sur 15 µg d'extraits protéiques totaux de cellules HEK293T transfectées ou non. Les protéines Rfx6 humaines sauvages et mutées sont produites dans les cellules HEK293T (105kDa).

Des tests luciférase mis en place ont également permis de montrer que la fonction de transactivation entre RFX6 et la Xbox2 dans le premier intron de Mlxipl est perdue pour quatre mutants de Rfx6 testés sur cinq (**Figure 55B**).

La perte de cette fonction de transactivation peut être expliquée par un défaut de liaison à l'ADN qui devra être testé par la suite. Pour cela, ce sont dans un premier temps les protéines RFX6 murine et humaine non mutées qui ont été soumises à des tests de liaison à l'ADN.

B.9. Test de liaison à l'ADN entre RFX6 et son site de liaison dans l'intron 1 de Mlxipl

B.9.a. Expression *in vitro* de la protéine RFX6 humaine

Transcription Traduction in vitro de la protéine RFX6 humaine

Pour réaliser des tests de liaison à l'ADN, une grande quantité de protéine d'intérêt pure doit être produite. Pour cela, des essais de Transcription Traduction *in vitro* (TTIV) (Kit TnT® Quick Coupled Transcription/Translation System Promega) de la protéine RFX6 humaine sauvage ont été réalisés dans un premier temps, et préalablement aux expériences de retard sur gel. Après un premier essai n'ayant pas permis de produire la protéine RFX6 humaine avec le plasmide pcDNA-huRfx6 matrice (**Figure 61A, puits 4**), nous avons choisi de contourner les éventuels problèmes liés à des structures secondaires hypothétiques qui pourraient être formées avec le vecteur pcDNA3 par utilisation d'un plasmide linéarisé en 3' du gène d'intérêt (**Figure 61A, puits 5**) ou d'un produit de PCR en tant que matrice (**amorces dans le Tableau 6**) (**Figure 61A, puits 6**). De plus, pour stimuler l'étape de traduction de la protéine d'intérêt et pour contourner un éventuel défaut de transcription *in vitro*, une matrice d'ARNm de huRfx6 possédant le promoteur T7 du vecteur pcDNA3 a également été testée (produite par transcription *in vitro* à l'aide de l'ARN polymérase T7) (**Figure 61A, puits 7**).

Etant donné que la production de la protéine RFX6 humaine n'a pu être observée dans aucune des conditions testées, nous avons choisi d'essayer de produire la protéine RFX6 de souris à partir du vecteur d'expression pcDNA3-HA-mRfx6 présent dans le laboratoire par la même technique. Cette fois-ci, quelle que soit la matrice utilisée, les protéines sont détectées en WB anti-HA (sauf pour l'ARN HA-mRfx6) et anti-RFX6. La production de la protéine HA-mRFX6 est donc efficace par TTIV (**Figure 61B**).

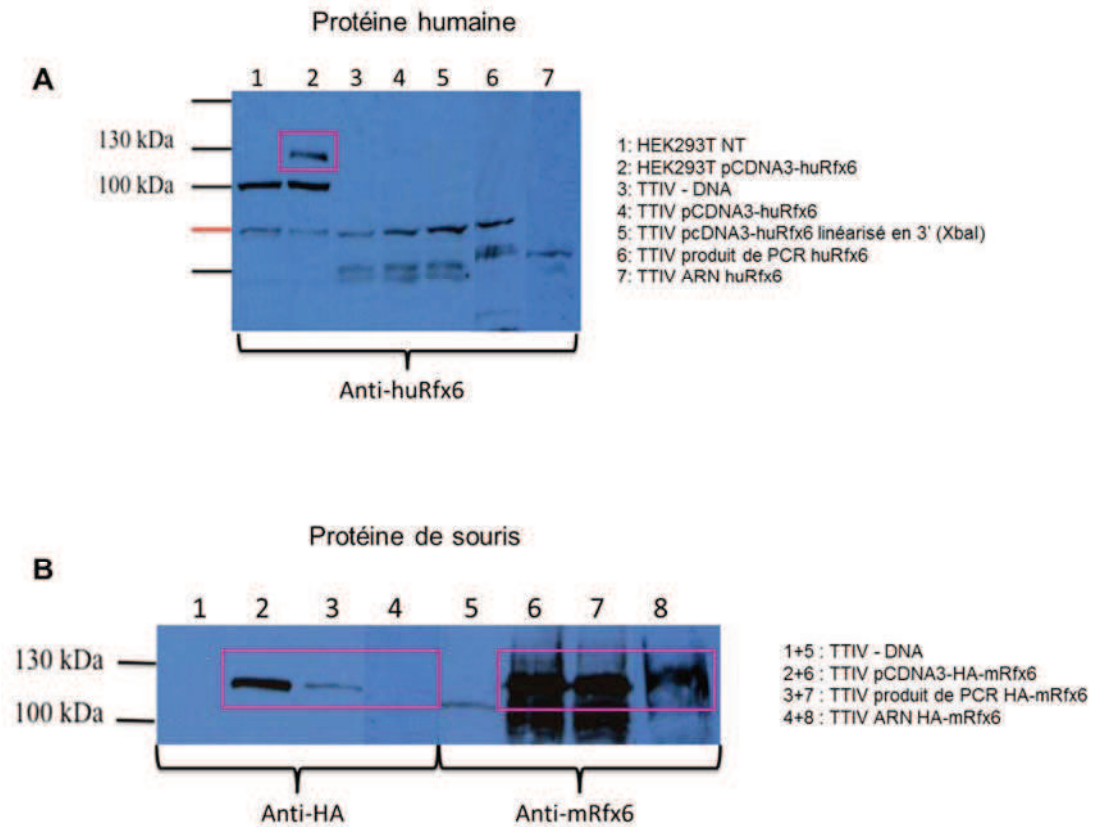


Figure 61 : Essai de production des protéines RFX6 humaine et murine par Transcription Traduction *in vitro*. (A) Western blot sur 15 µg d'extraits protéiques totaux de cellules HEK293T transfectées ou non (1 et 2 : contrôle d'expression de huRFX6 en cellules HEK293T) ou des aliquots de 5 µL de TTV (3 à 7). La protéine huRFX6 est produite en cellules de mammifères (encadrement rose) mais pas en TTV. (B) Western blot sur des aliquots de 5µL de TTV de la protéine HA-mRFX6 de souris (1 à 8). La protéine de fusion HA-mRFX6 est produite par TTV (encadrement rose).

La protéine humaine RFX6 n'est ainsi produite que dans les cellules HEK293T alors que la protéine de souris peut être produite par Transcription Traduction *in vitro*.

Clonage et Transcription Traduction in vitro de la protéine de fusion 3HA-RFX6 humaine

Finalement, une autre stratégie de clonage a été envisagée dans le but de changer l'environnement de démarrage de la traduction (séquence KOZAK) (Tableau 10) de notre protéine d'intérêt, et peut-être ainsi, favoriser sa production en grande quantité *in vitro*, qui nous permettra de mettre en place les études fonctionnelles souhaitées.

Plasmide	Environnement de démarrage de la traduction
pcDNA3-HA-mRfx6	CAAGCTTAC <u>ATGGG</u> GCTACC
pcDNA3-huRfx6	CGCAGGAGG <u>ATGGC</u> CAAGGTC
pcDNA3-3HA-huRfx6	CGCAGGAGG <u>ATGGC</u> TTACCAA

Tableau 10 : Comparaison des séquences entourant le codon d'initiation des protéines huRfx6 et 3HA-huRfx6 dans le plasmide pcDNA3. La séquence KOZAK est soulignée et le codon d'initiation de la traduction est en gras. Les nucléotides présents après la séquence KOZAK diffèrent (en rose) dans les plasmides pcDNA3-huRfx6 et pcDNA3-3HA-huRfx6.

Pour cela, la stratégie de clonage présentée en **Figure 62** permettant l'ajout d'une étiquette 3HA en 5' de huRfx6 dans le vecteur pcDNA3-huRfx6 a été réalisée.

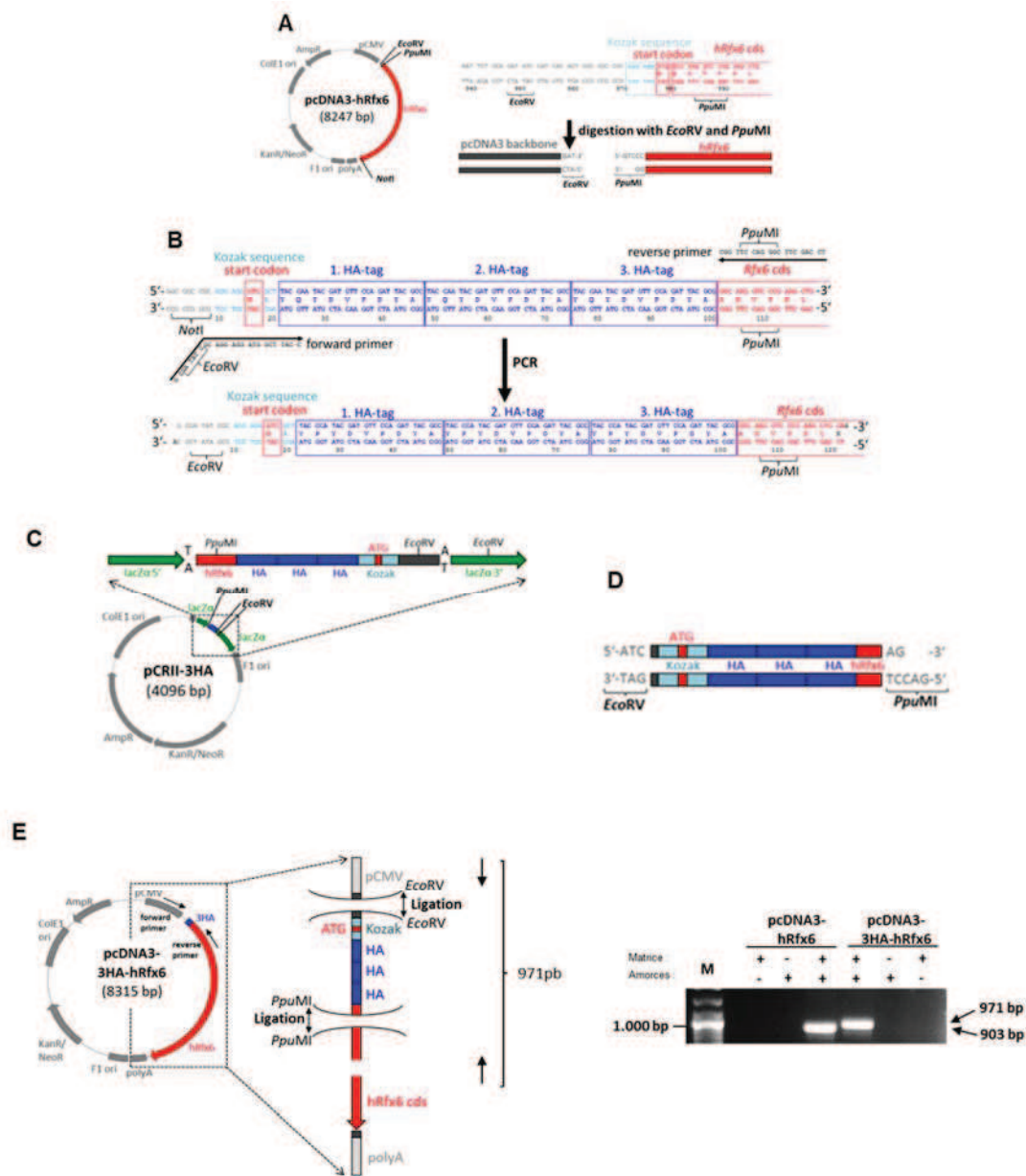


Figure 62 : Stratégie de production de la protéine 3HA-huRfx6. (A) Préparation du vecteur d'expression pcDNA3-huRfx6 par digestion de restriction EcoRVet PpuMI. **(B)** Production de l'insert 3HA par PCR. **(C)** Sous-clonage de l'insert 3HA dans le vecteur pCRII-TOPO. **(D)** Préparation de l'insert 3HA par digestion de restriction EcoRV et PpuMI du plasmide pCRII-3HA obtenu en (C) **(E)** Obtention du plasmide d'expression pcDNA3-3HA-huRfx6 et identification des clones positifs (produit de PCR de 971pb) par PCR sur colonie.

Après obtention du plasmide pcDNA-3HA-huRfx6 d'intérêt, des essais de transcription traduction *in vitro* ont finalement permis la production de la protéine 3HA-huRFX6 (Figure 63).

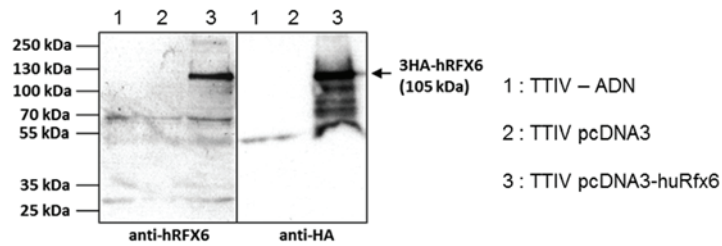


Figure 63 : Production de la protéine 3HA-RFX6 humaine par Transcription Traduction *in vitro*. Western blot sur des aliquots de 5 μ L de TTIV (1 à 3). La protéine humaine RFX6 étiquetée en 5' par 3HA est produite par TTIV.

Cette protéine de fusion 3HA-RFX6 humaine produite par TTIV permettra ainsi la mise en place d'expériences de retard sur gel, et de montrer la spécificité de nos résultats par l'utilisation d'un double retard sur gel à l'aide d'un anticorps anti-HA. Les vecteurs d'expression possédant les mutations ponctuelles produits dans le paragraphe Résultats B.8.a pourront également être soumis à la même stratégie de clonage afin de pouvoir les utiliser en TTIV puis les tester en gel retard.

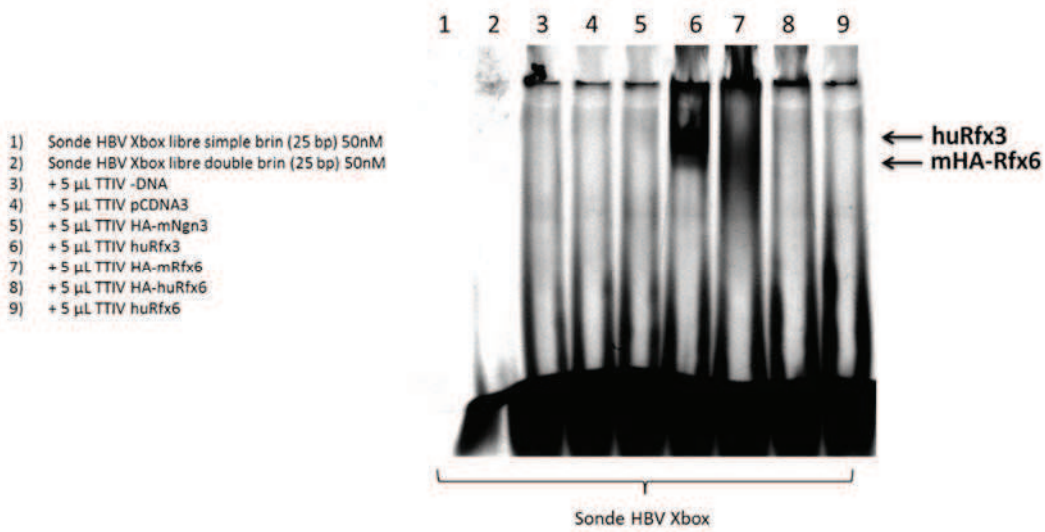
B.9.b. Mise au point des tests de gel retard

La fonction de fixation à l'ADN étant fondamentale pour que RFX6 puisse jouer son rôle de facteur de transcription, des tests de gel retard à l'aide de sondes ADN couplées à la Cyanine5.5 ont été mis en place.

Comme indiqué en **Figure 64**, deux sondes différentes ont été testées. D'une part, la sonde HBV a été désignée à partir des différentes publications ayant déjà montré une fixation entre des facteurs de transcription de type « RFX » et le motif spécifique retrouvé dans le virus de l'hépatite B (David et al, 1995; Reith et al, 1994; Smith et al, 2010). Celle-ci servira de contrôle positif pour cette expérience. D'autre part, la sonde Xbox2 a été désignée à partir de la séquence d'intérêt identifiée par ChIP Seq RFX6 dans les cellules Min6b1. La fixation du facteur de transcription humain RFX3 en tant que contrôle positif est observée pour les deux sondes testées (**Figure 64A piste 6** et **Figure 64B piste 5**). Les différents contrôles négatifs utilisés sont des aliquots de TTIV à partir de vecteur vide pcDNA3, sans ADN ou encore avec un vecteur pcDNA3-HA-mNgn3. Dans cette expérience les contrôles

négatifs ne montrent aucune liaison spécifique aux Xbox présentes dans les sondes ADN testées comme attendu (**Figure 64A piste 3 à 5 et Figure 64B piste 2 à 4**). Par ailleurs, la fixation du facteur de transcription murin HA-RFX6 est également observée mais uniquement avec la sonde précédemment publiée HBV (**Figure 64A piste 7**). Enfin, la fixation spécifique du facteur de transcription humain 3HA-RFX6 n'a pas encore pu être démontrée ici, quelle que soit la sonde utilisée en gel retard.

A



B

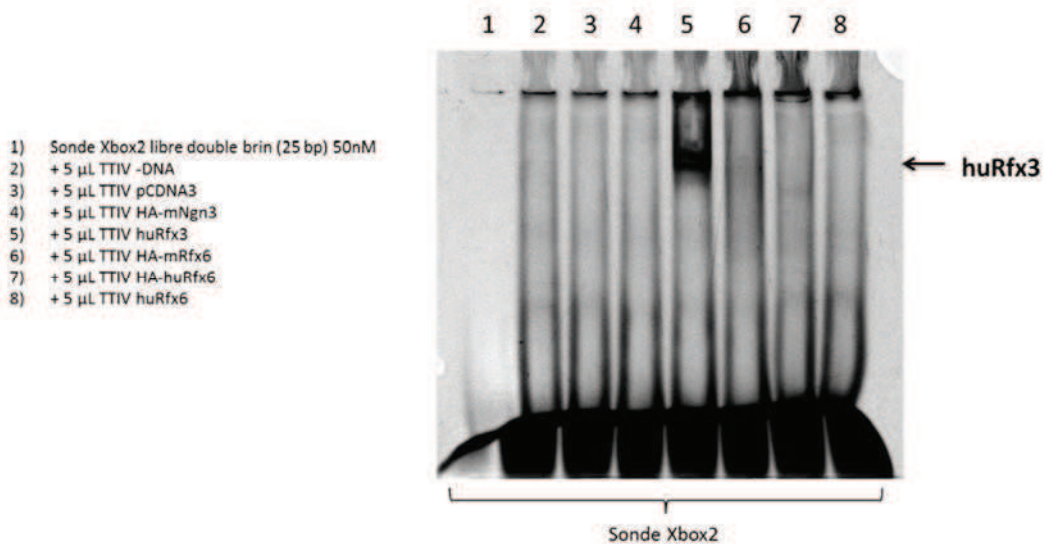


Figure 64 : Etude de la liaison du facteur RFX6 sur la Xbox2 de Mlxipl. (A) Gel retard (acrylamide) entre des protéines produites par Transcription Traduction *in vitro* et la sonde HBV Xbox. La liaison spécifique des protéines huRfx3-Flag et mHA-Rfx6 est montrée. **(B)** Gel retard entre des protéines produites par Transcription Traduction *in vitro* et la sonde Xbox2. Seule la protéine huRfx3-Flag semble se lier à la Xbox2 de Mlxipl.

Des mises au point pour cette technique sont ainsi encore nécessaires pour pouvoir conclure quant à la liaison spécifique de la protéine RFX6 de souris ou humaine sur la Xbox2 identifiée par ChIP Seq RFX6 dans les cellules Min6b1 et dans les ilots pancréatiques de souris au niveau du premier intron de Mlxipl. Un contrôle à l'aide d'une concentration croissante de sonde non marquée pourra être réalisé pour montrer la spécificité de fixation entre la protéine HA-mRFX6 de souris et sa cible Xbox2 lorsque celle-ci aura pu être

montrée. Enfin, l'étude des conséquences des mutations humaines sur la fixation à l'ADN de RFX6 devra être finalisée par l'utilisation de Transcription Traduction *in vitro* des protéines RFX6 humaines sauvage et mutantes et «double» retard sur gel à l'aide d'un anticorps dirigé spécifiquement contre la protéine de fusion 3HA-RFX6 humaine (**paragraphe Résultats B.9.a**).

B.10. Etude du lien entre la stimulation de la sécrétion de l'insuline en réponse au glucose et le couple Rfx6/Mlxipl

Dans la cellule beta adulte, plusieurs études montrent que Mlxipl est essentiel à la sécrétion d'insuline en réponse au glucose. En effet, la diminution d'expression de Mlxipl dans les cellules Min6 provoque une augmentation de la sécrétion d'insuline (da Silva Xavier et al, 2006). Un effet opposé est également observé lorsque Mlxipl est surexprimé dans les cellules INS-1 ou dans les ilots pancréatiques (Poungvarin et al, 2012). Quand Mlxipl est surexprimé, l'architecture des ilots pancréatiques est également perturbée et les souris sont diabétiques. Enfin, Noordeen et al. montrent que Mlxipl réprime le FT ARNT/HIF1beta, régulateur positif de la sécrétion d'insuline en réponse au glucose (Noordeen et al, 2010).

Parce que nous avons récemment montré que RFX6 est un acteur clé de la fonction de sécrétion de la cellule beta (Piccand et al, 2014) et qu'il régule directement Mlxipl (**paragraphe Résultats B.5 et B.6**), la perturbation d'expression du couple Rfx6/Mlxipl et ses conséquences sur la synthèse de l'insuline ont été recherchées.

Pour répondre à cette question, nous avons choisi de mettre en place des tests ELISA (Kit 80-INSMSU AlpcO) sur des cellules Min6b1 traitées par siRFX6 où l'expression de Mlxipl est également réduite de 50% (**Figure 54A et B**). De premiers tests montrent que la sécrétion d'insuline est diminuée de 25% lorsque l'expression de Rfx6 est réduite dans les cellules Min6b1 (**Figure 65**) ce qui confirme le rôle important de Rfx6 pour la sécrétion d'insuline dans ce système cellulaire. Ces résultats corréleront également avec ceux obtenus chez la souris delta beta (figure 3E, (Piccand et al, 2014)). Ainsi, les mécanismes de sécrétion d'insuline en rapport avec RFX6 peuvent être étudiés par l'utilisation de siRfx6 dans les cellules Min6b1.

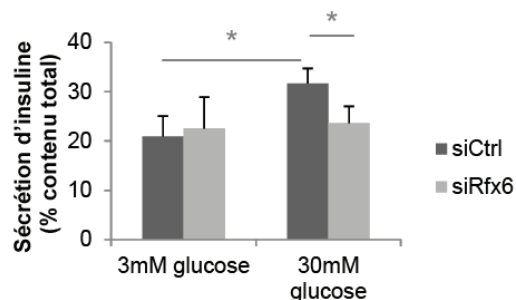


Figure 65 : La réduction d'expression de Rfx6 diminue la sécrétion d'insuline stimulée par le glucose dans les cellules Min6b1. La sécrétion d'insuline est diminuée de 25% lorsque l'expression de Rfx6 est réduite dans les cellules Min6b1 traitées par 75nM de siRfx6 comparées aux cellules traitées par des siCtrl (n=3). Les résultats sont exprimés en moyenne \pm SD *p<0,05.

Etonnamment, ces résultats *in vitro* et *in vivo* sont en contradiction avec un éventuel rôle de répresseur de la sécrétion d'insuline en réponse au glucose pour Mlxipl suggéré par les publications citées ci-dessus. La diminution d'expression de Mlxipl consécutive à la diminution d'expression de Rfx6 et la conséquence négative sur la sécrétion d'insuline observée résulte peut-être d'une combinaison d'effets opposés, ou de l'inactivation seulement partielle de Rfx6 et Mlxipl obtenue dans ces expériences. En perspective de ce travail, pour confirmer les données publiées et pour éclaircir le rôle exact de Rfx6 et Mlxipl dans ce système expérimental, il faudra réaliser un essai de diminution directe d'expression de Mlxipl par siMlxipl (ON-TARGETplus Mouse Mlxipl (58805) siRNA – SMARTpool L-04231000 Dharmacon) suivi de test ELISA. Un premier test réalisé avec 75nM de siMlxipl ne nous a pour l'instant pas permis de diminuer de manière significative son expression dans les cellules Min6b1.

Cette étude devrait ainsi permettre de définir plus précisément le rôle du couple Rfx6/Mlxipl dans la sécrétion d'insuline stimulée par le glucose dans la cellule beta adulte.

B.11. Étude du transport cytoplasme/noyau de MLXIPL et RFX6 en fonction de la concentration en glucose

Le transport nucléocytoplasmique de MLXIPL est régulé par la concentration en glucose (da Silva Xavier et al, 2006). De plus, Noordeen et al. (Noordeen et al, 2012)

montrent qu'en présence d'une forte concentration en glucose, Mlxipl est transloqué dans le noyau et que ce phénomène est contrôlé par la sorcine.

Pour sa part, la protéine RFX6 est observée dans le cytoplasme des cellules Min6b1 par IF (**Figure 66**). Pour savoir si le transport nucléocytoplasmique observé pour MLXIPL est également présent pour RFX6, et découvrir son éventuel rôle dans cette translocation, j'ai réalisé des IF contre les protéines RFX6 et MLXIPL endogènes des cellules Min6b1. Trois différents anticorps commerciaux dirigés contre MLXIPL et un anticorps anti-RFX6 ont été testés sur des cellules Min6b1. Aucune différence de localisation cellulaire n'a pu être révélée quelle que soit la concentration en glucose testée (3 ou 30 mM) (**Figure 66**) et les résultats précédemment publiés n'ont donc pas pu être confirmés pour les protéines exprimées de manière endogène dans ce modèle cellulaire.

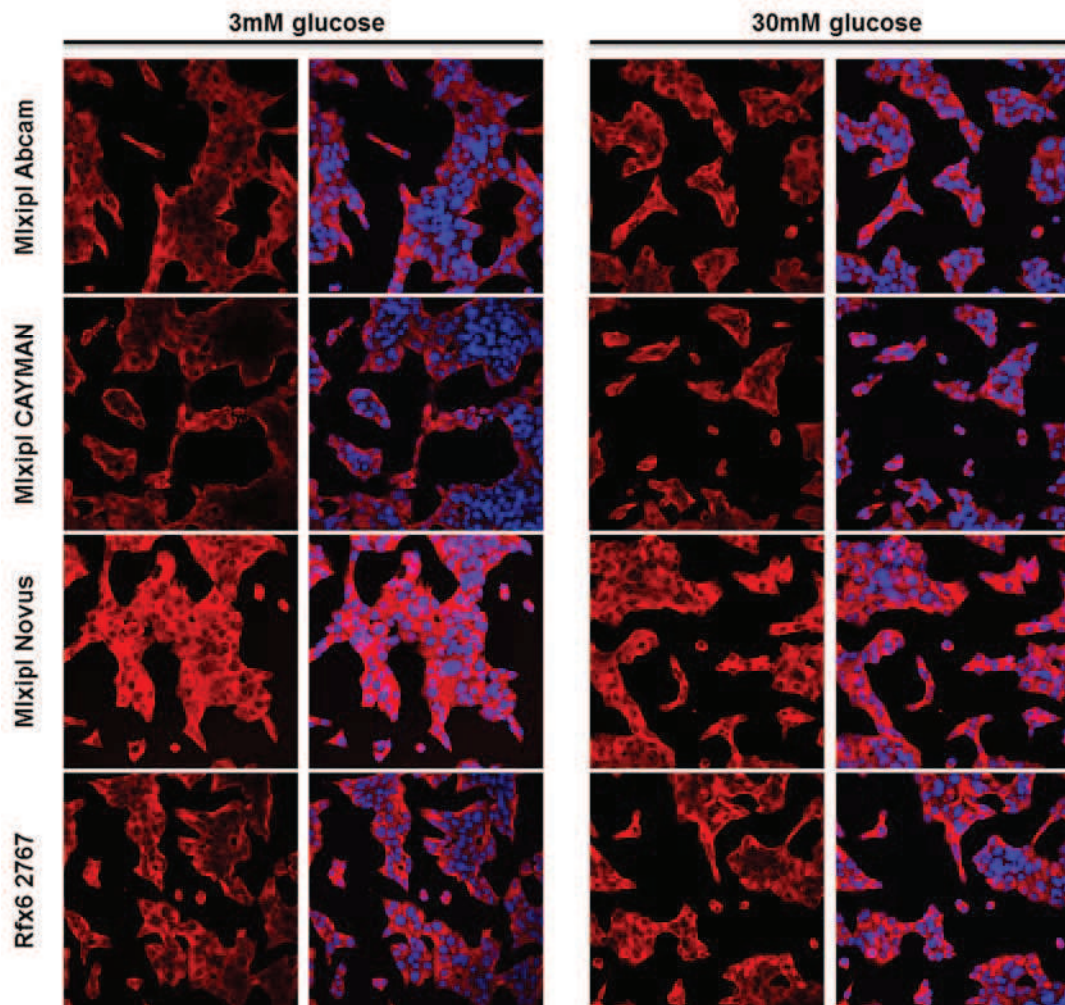


Figure 66 : Le transport nucléaire des protéines MLXIPL et RFX6 endogènes ne se produit pas sous stimuli de glucose dans les cellules Min6b1. Immunofluorescence contre MLXIPL et RFX6 (en rouge) après stimulation de 6h en présence de 3 ou 30mM de glucose précédée par une pénurie en glucose à 3mM pendant une nuit. Les acquisitions ont été réalisées au microscope confocal. Les noyaux cellulaires sont marqués au DAPI en bleu. Aucune différence de localisation cellulaire des protéines MLXIPL et RFX6 endogène n'est observée en présence de 3 ou 30mM de glucose.

Il est à noter que dans chaque cas, les études publiées concernant la localisation cellulaire sont réalisées dans des cellules Min6b1 transfectées entraînant une surexpression de Mlxipl et que les marquages ne sont pas dirigés contre Mlxipl mais contre une étiquette « myc ». D'autre part, la localisation nucléaire de la protéine de fusion 3HA-RFX6 surexprimée dans les cellules Min6b1 avait pu être montrée (**Figure 18**) et ne correspond pas au marquage observé pour la protéine RFX6 endogène observé ici. Ceci suggère qu'une étude plus approfondie de la localisation intracellulaire de la protéine RFX6 devra être

développée. Pour cela, différents essais, à l'aide d'autres anticorps contre Rfx6, d'une autre lignée cellulaire beta, telle que la lignée INS-1E ou directement au niveau des ilots pancréatiques *in vivo*, pourront encore être réalisés pour mieux comprendre l'éventuel rôle de RFX6 dans la translocation nucléocytoplasmique.

B.12. Étude de la prolifération des cellules Min6b1 en fonction de la concentration en glucose

En 2012, les travaux de Metukuri et al. montrent par études de perte et de gain de fonction dans des ilots pancréatiques dispersés de rongeur, mais également humains, que Mlxipl, par son effet sur les régulateurs du cycle cellulaire, stimule la prolifération des cellules beta (Metukuri et al, 2012) en réponse au glucose (**paragraphe Résultats B.4**). D'autre part, le métabolisme glucidique est fortement perturbé dans les cellules tumorales. En effet, afin de s'adapter à l'hypoxie et maintenir la production d'ATP, les cellules augmentent la glycolyse et donc la consommation de glucose. Ce processus est essentiel pour la transformation maligne. En 2009, il a été suggéré que MLXIPL, dont l'activité est induite par le glucose, contribuait à la fois au maintien de la glycolyse aérobie et à l'expansion des cellules tumorales (Tong et al, 2009b).

Ces études ainsi que mes résultats, montrant que RFX6 régule directement et de manière positive l'expression de Mlxipl, suggèrent donc que RFX6 pourrait avoir un rôle dans la prolifération des cellules beta. Pour explorer cette hypothèse, le modèle cellulaire beta Min6b1, qui exprime à la fois Rfx6 et Mlxipl, et dans lequel leur expression peut être perturbée par siRNA, pourra être utilisé et leur rôle dans la prolifération de ces cellules précisé.

Pour répondre à cette question, l'effet de la concentration de glucose sur la prolifération des cellules Min6b1 a tout d'abord été testé. Un marquage des cellules en division à l'iodure de propidium est effectué avant tri cellulaire au FACS. Si le glucose a un impact sur la prolifération des cellules Min6b1, l'expression de Mlxipl sera induite et devrait provoquer une augmentation du pourcentage de cellules en phase S. Comme indiqué en **Figure 67**, le pourcentage de cellules Min6b1 en phase S ne présentent pas de différence significative en présence de 3 ou 15 mM de glucose, que ce soit pour un nombre de passages faible (17) ou plus élevé (34).

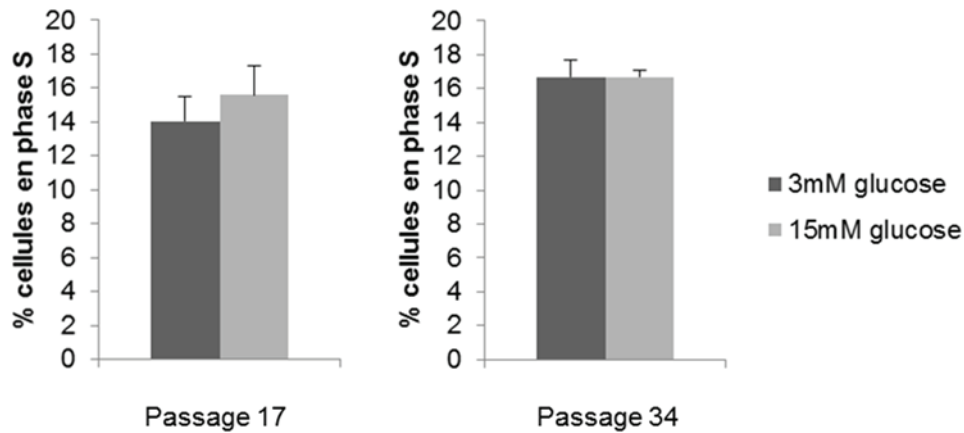


Figure 67 : Prolifération des cellules Min6b1 en fonction de la concentration en glucose.

Les cellules Min6b1 (passage 17 ou 34) en division sont marquées à l'iodure de propidium puis triées par FACS après avoir été préincubées pendant 6h à 3mM glucose puis exposées pendant 16h à 3 ou 15mM glucose. La prolifération des cellules Min6b1 n'est pas dépendante de la concentration en glucose.

La prolifération des cellules Min6b1 ne semble donc pas dépendante du glucose, suggérant que ce modèle n'est pas adapté à la question posée ici. Ces études préliminaires devront encore être testées dans d'autres lignées cellulaires (par exemple INS1-E) ou *in vivo* dans les îlots de souris, en présence de différentes concentrations en glucose. Enfin, des conditions dans lesquelles l'expression de Mlxipl et/ou de Rfx6 est dérégulée (siRNA – Rfx6 KO) pourront être évaluées pour préciser l'implication du couple Rfx6/Mlxipl dans la prolifération des cellules beta.

Les expériences proposées dans ce paragraphe visent à comprendre le rôle de Rfx6 et de sa cible Mlxipl dans la prolifération des cellules beta sous différents stimuli. Nous escomptons que les résultats qui seront obtenus dans le futur apporteront des éclaircissements quant aux relations entre le métabolisme et la prolifération cellulaire.

Pour conclure, ces travaux de thèse ont tout d'abord permis de démontrer l'importance de Rfx6 dans le maintien de l'identité et de la fonction de la cellule beta par une approche combinée de transcriptomique chez la souris et la recherche des sites de fixation de ce facteur de transcription dans une lignée cellulaire beta.

Pour la première fois, l'identification des cibles directes de Rfx6 *in vivo* au niveau d'ilots de Langerhans a été réalisée et a permis d'une part de valider *in vivo* les cibles identifiées dans la lignée cellulaire Min6b1 et d'autre part d'identifier l'ensemble du répertoire des gènes régulés directement par Rfx6 dans les cellules beta qui n'ont pas été révélés dans le système cellulaire.

Cette étude aura également permis d'identifier Mlxipl comme principale cible directement régulée par Rfx6 à la fois chez la souris et l'homme.

En conclusion les expériences réalisées ont permis de déterminer les gènes cibles directs de Rfx6 et ainsi contribuer à élucider le rôle de ce facteur de transcription dans la différenciation et la fonction des cellules beta.

Discussion et perspectives

Discussions et perspectives

A. Rfx6 est un modulateur clé de la sécrétion d'insuline et du maintien de l'homéostasie du glucose

A.1. RFX6 cible des gènes essentiels à la sécrétion d'insuline et à l'homéostasie du glucose dans les ilots pancréatiques

Lors de ce travail, les gènes cibles du facteur de transcription RFX6 ont été caractérisés dans la lignée cellulaire beta Min6b1 *in vitro* puis dans les ilots pancréatiques de souris *in vivo*. Par une approche combinée de transcriptomique dans la souris n'exprimant pas Rfx6 spécifiquement dans les cellules beta, ce sont les gènes directement régulés par RFX6 qui ont été déterminés.

En 2014 dans le journal Cell Reports, notre équipe a montré que la protéine RFX6 est essentielle au maintien des principales caractéristiques des cellules beta fonctionnellement matures chez la souris (Piccand et al, 2014). L'inactivation conditionnelle de Rfx6 spécifiquement dans les cellules beta chez l'adulte, mène à une sécrétion d'insuline altérée et à l'intolérance au glucose. Notre étude révèle que ce phénotype résulte d'une réduction combinée de l'expression de gènes nécessaires au mécanisme de sécrétion d'insuline dont la glucokinase, la sous-unité Abcc8 du canal potassium sensible à l'ATP ainsi que plusieurs canaux calciques voltage dépendant qui sont tous des cibles directes de RFX6 identifiés par ChIP Seq dans les cellules Min6b1. De plus, nous avons découvert que RFX6 contribue à la répression de la majorité des gènes dit «disallowed», habituellement spécifiquement réprimés dans les cellules beta adultes, et donc au maintien de l'identité de la cellule beta. En effet, ce sont plus de 58% de la totalité de ces gènes « disallowed » qui sont ciblés par RFX6 dans les cellules Min6b1. La plupart des cibles de RFX6 révélées par cette étude ont ensuite été confirmées par ChIP Seq RFX6 dans les ilots pancréatiques de souris. La confirmation de ces résultats *in vivo* (**Figure 46**) montre que la lignée cellulaire beta Min6b1 constitue un excellent modèle d'étude.

En 2010, Smith et al. (Smith et al, 2010) montrent que Rfx6 est requis pour la production d'insuline chez l'homme et la souris. De même en 2014, Chandra et al. (Chandra et al, 2014) montrent que RFX6 régule l'expression du gène de l'insuline ainsi que la sécrétion de cette hormone dans la lignée cellulaire beta humaine EndoC-betaH2. Rfx6 se lie à une Xbox située 288pb en amont du TSS du gène de l'insuline et active sa transcription. Contrairement à cette observation dans la lignée cellulaire beta humaine, chez la souris,

Rfx6 n'est pas nécessaire à la production d'insuline (Piccand et al, 2014) où seulement une diminution de 60% de l'expression du gène *Ins1* et aucune variation d'expression du gène *Ins2* sont observées en absence de Rfx6 dans la cellule beta adulte. De plus, par criblage des X-box présentes dans les 10 kb en amont du TSS du gène *Ins1*, aucune liaison de RFX6 n'a pu être mise en évidence par CHIP quantitatif dans les cellules Min6b1 lors de mon travail. De même, aucun site de liaison de Rfx6 n'a été révélé à proximité du gène *Ins1* lors des CHIP Seq effectués. La régulation du gène de l'insuline semble ainsi différente entre l'homme et la souris. Cette distinction de régulation du gène de l'insuline avait déjà été proposée par l'analyse des promoteurs de l'insuline chez différentes espèces et l'observation d'une grande dissimilitude entre ces derniers chez l'homme et la souris (Hay & Docherty, 2006). Par ailleurs, lorsque Rfx6 est inactivé chez l'embryon dans toutes les cellules endocrines (Piccand et al, 2014), l'expression des gènes *Ins1* et *Ins2*, cette fois ci, diminue fortement, suggérant que la régulation de ce gène lors de la différenciation terminale des cellules beta est également différente de celle observée chez l'adulte.

A.2. Les cibles de RFX6 sont déterminées *in vitro* et *in vivo*

D'autre part, la comparaison des jeux de données des CHIP Seq RFX6 dans les cellules Min6b1 et dans les ilots pancréatiques, a permis de révéler 670 gènes cibles communs (**Figure 39**) soit 45,6% des gènes cibles identifiés dans les cellules Min6b1 ce qui permet de valider ce modèle d'étude. Par contre, seulement 8,9% des gènes cibles dans les ilots de Langerhans sont communs à ceux identifiés dans les Min6b1. Cela montre qu'il existe de nombreuses cibles spécifiques des ilots. Celles-ci pourront donc être étudiées par l'utilisation du jeu de données CHIP RFX6 dans les ilots pancréatiques *in vivo*.

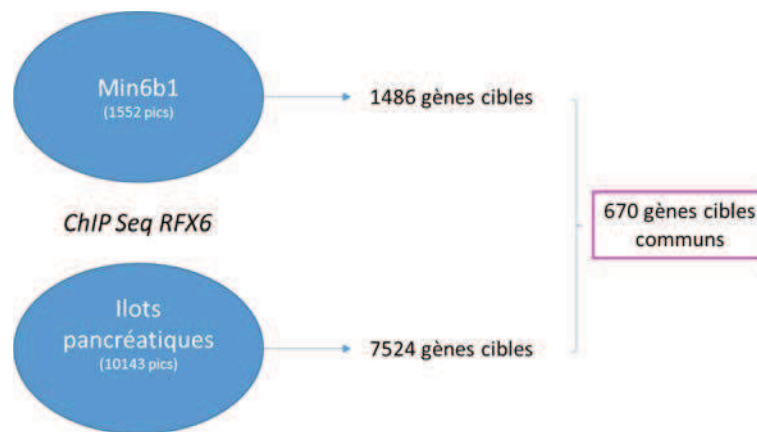


Figure 68 : Comparaison des gènes ciblés par RFX6 dans les cellules Min6b1 et les îlots pancréatiques. Le ChIP Seq RFX6 dans les Min6b1 permet l'identification de 1486 gènes cibles (correspondant à 1552 pics) alors que dans les îlots pancréatiques, RFX6 cible 7524 gènes (correspondant à 10143 pics). 670 gènes cibles sont communs aux deux jeux de données.

Une étude préliminaire indique que parmi les cibles retrouvées uniquement dans les îlots, 1869 gènes sont significativement (p -valeur $<0,05$) régulés par RFX6 (RNA Seq souris delta beta) et parmi eux 482 présentent une variation d'expression supérieure à 2. L'étude fonctionnelle de ces gènes directement régulés par RFX6 pourra être réalisée notamment pour permettre l'identification des cibles de RFX6 qui contribuent au défaut de sécrétion d'insuline observé chez la souris déficiente pour RFX6.

La distinction des cibles de RFX6 parmi les différents sous-types cellulaires qui composent les îlots pancréatiques pourrait également préciser l'implication de RFX6 dans les fonctions spécifiques des cellules alpha, delta, PP et epsilon. Par exemple, la recherche des cibles spécifiques de RFX6 dans les cellules alpha est prévue dans notre équipe. Pour cela, le développement d'une souris dans laquelle l'expression de Rfx6 sera conditionnellement inactivée dans les cellules alpha va être démarré et des études transcriptomiques par RNA Seq seront réalisées. Pour l'instant aucun gène cible spécifique des cellules alpha tel que Gcg, Arx ou Irx1 et 2, n'a pu être identifié dans notre ChIP Seq anti-RFX6 dans les îlots pancréatiques. Par l'intermédiaire de cette étude, le rôle précis de RFX6 dans les cellules alpha du pancréas pourra être déterminé.

Parmi les cibles « beta » de RFX6, celles citées dans ce travail (**paragraphe Résultats A.4.d**), constituent un point de départ à l'identification des cibles de RFX6 qui contribuent au défaut de sécrétion d'insuline observé chez la souris déficiente pour ce facteur de transcription. Le gène *Ffar1* par exemple est exprimé dans les cellules beta et permet une régulation positive de la sécrétion d'insuline en réponse à la présence d'acides gras (Itoh et al, 2003). Comme montré lors de ce travail, RFX6 cible et régule positivement son expression puisqu'une diminution d'expression de 4 fois est observée en absence de *Rfx6* dans les cellules beta. *Irs2* est également ciblé et régulé par RFX6 dans les ilots pancréatiques. En absence de *Rfx6* dans les cellules beta chez la souris, son expression chute de 1,5 fois. Les souris mutantes pour *Irs2* présentent une intolérance au glucose due à un défaut de sécrétion d'insuline et de mobilisation des ions calcium dans ce processus (Cantley et al, 2007). *Irs2* est ainsi essentiel à la fonction de sécrétion d'insuline de la cellule beta. C'est aussi le cas de *Hmgn3*, dont la perte chez la souris, entraîne une sécrétion d'insuline perturbée, notamment due à un défaut de transcription du transporteur de glucose *Glut2*, et un phénotype diabétique (Ueda et al, 2009). Ce gène *Glut2* est lui aussi sous-exprimé en l'absence de *Rfx6* dans les ilots pancréatiques (FC = -1,9) et directement ciblé par RFX6. En outre, le gène circadien *Per2*, dont les souris déficientes présentent une sécrétion d'insuline fortement stimulée en présence de glucose (Zhao et al, 2012), et qui est donc considéré comme régulateur négatif de la sécrétion d'insuline, est également ciblé et régulé (FC = -2) par RFX6.

La régulation de l'expression de l'ensemble de ces gènes par RFX6, le place ainsi comme un modulateur clé de la sécrétion d'insuline et du maintien de l'homéostasie du glucose dans la cellule beta.

A.3. RFX6 et ses partenaires

L'étude des facteurs agissant en collaboration avec RFX6 pour réguler l'expression de ses gènes cibles constitue une stratégie supplémentaire pour découvrir et préciser le rôle de ce facteur de transcription dans la cellule sécrétrice d'insuline.

Grâce aux études de ChIP Seq de ce travail, et à l'analyse des régions spécifiquement ciblées par RFX6 dans les ilots pancréatiques, d'autres motifs de fixation de facteur de transcription ont été révélés. C'est notamment le cas du motif lié par le facteur de transcription NRF1 qui est enrichi au niveau des sites ciblés par RFX6 (**Figure 29 et Figure 40**). Cette protéine a spécifiquement été identifiée comme régulatrice clé de nombreux gènes dérégulés chez des patients diabétiques et pré-diabétiques (Patti, 2004) ce qui suggère que RFX6 et NRF1 pourraient avoir une fonction commune dans la cellule beta sécrétrice

d'insuline. Pour répondre à cette question, l'étude des gènes cibles communs à ces deux protéines peut être envisagée. Les gènes ciblés par NRF1 ont déjà été analysés dans une lignée cellulaire issue de neuroblastome humain (Sato et al, 2013) à partir d'un jeu de données issue du Projet Encode (2004) et cette étude pourrait constituer un premier support pour répondre à cette question. Un second motif également enrichi au niveau des sites ciblés par RFX6 dans les cellules beta est le motif associé au facteur de transcription INSM1. Par analyse de souris mutantes pour *Insm1*, notre équipe a montré en 2006 que *Ngn3* induit son expression en amont des gènes de la différenciation endocrine tels que *NeuroD1*, *Pax4*, *Arx* et *Pax6* (Mellitzer et al, 2006). Il est identifié comme un régulateur négatif transitoire au cours du développement des cellules beta pancréatiques par répression de sa propre expression mais également de *NeuroD1* ou de l'insuline (Lan & Breslin, 2009). De plus, la perte d'*Insm1* ou de *Rfx6* chez la souris entraîne un défaut de différenciation terminale des cellules endocrines en cellules sécrétrices d'hormones, ce qui suggère que ces facteurs de transcription peuvent avoir une fonction similaire (Gierl et al, 2006; Mellitzer et al, 2006; Smith et al, 2010; Soyer et al, 2010). Le rôle commun de RFX6 et INSM1 au cours du développement du pancréas, et la place de RFX6 dans la cascade de régulation transcriptionnelle mise en jeu lors du développement, pourra de ce fait faire l'objet d'un nouvel objectif de recherche.

D'autre part, les motifs présents de manière conjointe avec des motifs RFX connus et référencés dans la base de données de motifs, Jaspar 2014, ont été analysés spécifiquement au niveau des régions ciblées par RFX6 dans les cellules *Min6b1* et les ilots pancréatiques par recherche de co-occurrence de motifs (**Figure 39** et **Figure 42**). Dans les deux analyses, c'est particulièrement la co-occurrence des motifs de type « RFX » qui a été remarquée et qui pourrait permettre une régulation transcriptionnelle fine des gènes cibles de RFX6 dans les ilots pancréatiques par différents membres de la famille RFX.

Dans cette famille de facteurs de transcription, *Rfx6* et *Rfx3* sont exprimés dans le pancréas. Les souris KO pour *Rfx3* présentent une diminution du nombre de cellules alpha, beta et delta, une augmentation du nombre de cellules PP dans les ilots pancréatique et une production d'hormone affectée (Ait-Lounis et al, 2007). En 2010, le rôle clé de *Rfx3* pour la différenciation et la fonction des cellules beta adultes a été démontré (Ait-Lounis et al, 2010) notamment par la régulation de l'expression de la glucokinase. Les gènes ciblés par RFX3 et identifiés dans ces études sont également *Ift88* et *Dync2li1* composants essentiels du transport intraflagellaire et de la formation des cils cellulaires. En absence de *Rfx3*, les cils des cellules beta sont plus courts et moins nombreux mais leur rôle dans ces cellules n'est

pas établi. Lors de ce travail, la glucokinase a également été identifiée comme une cible directement régulée par RFX6, il en est de même pour Ift88 dont l'expression chute d'un facteur 1,4 en absence de Rfx6 dans les cellules beta et Dync2li1 est ciblé mais pas régulé par RFX6. Ces deux protéines partagent ainsi des cibles communes et leur comparaison à l'échelle génomique pourra permettre d'éclaircir leur rôle dans le pancréas.

Enfin, un travail complémentaire à la détermination des gènes directement régulés par RFX6 dans la cellule beta adulte pourra être la définition des gènes « actif » des cellules beta. Pour cela, l'analyse des modifications d'histones par ChIP Seq (H3K4me3, H3K27me3) dans les cellules beta pourra être réalisée et comparée aux gènes directement régulés par RFX6 dans les îlots pancréatiques déterminés lors de ce travail.

B. RFX6 régule directement l'expression de Mxipl dans la cellule beta pancréatique

Une autre cible de RFX6 dans la cellule beta identifiée dans ce travail est le gène du facteur de transcription Mxipl. En effet, RFX6 régule positivement son expression dans les îlots pancréatique et sa liaison au niveau de trois X-box dans le premier intron du gène a été démontrée dans les cellules Min6b1 (**Figure 51 et Figure 54**). Les études de transactivation menées suggèrent que la X-box2, qui est conservée à 100% entre l'homme et la souris et localisée dans le premier intron de Mxipl, est essentielle pour l'activation de sa transcription par la protéine RFX6 chez la souris et l'homme (**Figure 55**).

Diverses études montrent que l'expression de Mxipl est régulée par des facteurs de transcription qui se lient au niveau de son promoteur ou de son premier intron. En effet, Sirek et al. (Sirek et al, 2009) montrent dans des cellules de carcinome de foie humain que OCT-1 lie une région conservée au niveau du promoteur de Mxipl pour réprimer son expression. De même, la région promotrice de Mxipl est liée par les hétérodimères LXR/RXR et THR/RXR pour activer sa transcription dans le foie (Cha & Repa, 2007; Hashimoto et al, 2009). Dans les îlots pancréatiques, ce sont FOXA1 et FOXA2 qui ont été identifiés pour leur liaison au niveau de l'intron 1 de Mxipl (Gao et al, 2010). De plus, l'inactivation simultanée de Foxa1 et 2 entraîne une diminution d'expression de Mxipl dans les îlots pancréatiques, ce qui n'est pas observé lorsque Foxa1 ou 2 uniquement est inactivé. Les motifs de liaison de FOXA1 et 2 identifiés dans cette publication sont situés en 5' de la X-box1 et entre les X-box2 et 3 ciblées par RFX6 et identifiées lors de ce travail (**Figure 51**). Ces résultats suggèrent que RFX6 et FOXA1 et 2 pourraient agir de concert pour réguler l'expression de Mxipl. Pour

vérifier cette hypothèse, des expériences de transactivation entre la région du premier intron de Mlxipl comprenant les motifs de liaison de RFX6 et de FOXA1 et 2 (à l'aide de la construction produite et décrite dans le paragraphe Matériels et méthodes D.1 et Figure 15) et les facteurs de transcription d'intérêt pourront être réalisées.

B.1. Contrôle de la fonctionnalité des protéines RFX6 humaines mutantes

Des mutations humaines dans le gène Rfx6 sont à l'origine du syndrome de Mitchell Riley (paragraphe Introduction D.3) caractérisé par un diabète néonatal et des affections du système digestif plus ou moins sévères. La fonctionnalité des protéines RFX6 mutantes peut être testée par l'utilisation des tests fonctionnels de transactivation à l'aide du site cible de RFX6 dans le premier intron de Mlxipl (Xbox2). En effet, notre système a été validé par essai avec les protéines mutées déjà connues R181Q (Martinovici et al, 2010) et S217P (Chappell et al, 2008) (Figure 55). La protéine RFX6 R181Q perd significativement sa capacité à transactiver l'expression de Mlxipl par sa liaison à l'activateur de transcription Xbox2 de son premier intron. Ceci est en accord avec la perte de la capacité de cette protéine mutée à lier l'ADN, comme l'ont montré Smith et al en 2010 (Smith et al, 2010). De même, pour la mutation S217P, la fonction de transactivation testée dans le même système est également diminuée significativement. Smith et al (Smith et al, 2010) montrent que la dimérisation n'est pas affectée pour ce mutant et que la perte de liaison à l'ADN est plus faible. Contrairement au patient atteint de la mutation R181Q, la mutation S217P cause un phénotype moins sévère, ce qui corrèle avec une fonction de transactivation toujours présente.

D'autre part, trois autres mutations humaines de Rfx6, identifiées chez des patients diabétiques dans le laboratoire de P. Froguel par le groupe de Martine Vaxillaire à Lille, ont été testées à l'aide de ce système (Figure 55). Les mutations F294V et V506G sont localisées dans les domaines de dimérisation du gène Rfx6 et diminuent significativement la transactivation, probablement par perturbation de la dimérisation de la protéine. Il a aussi été montré par l'équipe de Raphaël Scharfmann, que la protéine RFX6 mutante V506G ne peut pas transactiver l'insuline, qui a été identifié comme gène cible dans la lignée cellulaire beta humaine EndoC-betaH2 (Chandra et al, 2014). Enfin, la mutation R578P, localisée en dehors des domaines de liaison à l'ADN ou de dimérisation de la protéine RFX6, ne présente aucune perturbation de sa fonction de transactivation.

La fonctionnalité des protéines RFX6 mutantes peut ainsi être testée par l'utilisation de ces tests fonctionnels de transactivation à l'aide du site cible de RFX6 dans le premier intron de Mlxipl (Xbox2). Enfin, la fonction de liaison à l'ADN de ces protéines RFX6 mutées devra encore être vérifiée par expérience de retard sur gel.

B.2. Les gènes régulés par Mlxipl contribuent-ils au phénotype de la souris delta beta ?

L'expression de Mlxipl est très fortement diminuée lorsque Rfx6 est spécifiquement inactivé dans les cellules beta sécrétrices d'insuline chez la souris (**Figure 54**), sa contribution au phénotype de ces souris peut ainsi être discutée. En raison de la perte induite de Mlxipl, causée par l'absence de Rfx6 dans la cellule beta chez la souris, des gènes dépendant de Mlxipl peuvent présenter une variation d'expression dans les analyses de RNA Seq chez la souris delta beta. Quels sont ces gènes et contribuent-ils au phénotype observé chez cette souris ?

La souris n'exprimant pas RFX6 spécifiquement dans les cellules beta, présente une sécrétion d'insuline affectée. Les gènes cibles de RFX6 responsables de ce phénotype ont été identifiés dans la publication de 2014 de notre équipe dans le journal Cell Report (Piccand et al, 2014). La contribution de Mlxipl pourra ici être démontrée par l'identification de ses gènes cibles dans les cellules beta pancréatiques. Les rôles spécifiques de RFX6 et MLXIPL dans le mécanisme de sécrétion d'insuline pourront être caractérisés dans le modèle cellulaire Min6b1 où de premiers résultats montrent que lorsque l'expression de Rfx6 est réduite par ARN interférence, l'expression de Mlxipl est affectée, tout comme la sécrétion d'insuline en réponse au glucose (**Figure 65**). Ces résultats récapitulent le phénotype observé chez la souris delta beta. De plus, par surexpression de Mlxipl dans ce modèle, un essai de rétablissement de la sécrétion d'insuline en réponse au glucose pourra être tenté.

D'autre part, nous avons également montré que RFX6 réprime l'expression de la majorité des gènes « disallowed » des cellules beta pancréatiques définis par plusieurs études (Pullen et al, 2010; Thorrez et al, 2011). RFX6 cible spécifiquement 58% de ces gènes « disallowed » dans les cellules Min6b1 et seulement 31% dans les îlots pancréatiques. La fonction d'activateur et de répresseur de la transcription mise en évidence pour Mlxipl dans différentes études (Boergesen et al, 2011; Noordeen et al, 2010) évoque qu'il pourrait être le candidat approprié pour réprimer les gènes disallowed qui ne sont pas ciblés par RFX6 mais dont l'expression est tout de même affectée dans la souris delta beta.

Par ailleurs, l'expression de gènes régulés par MLXIPL dans le foie tels que Lpk, Acc1, ou Fas n'est pas affectée par la perte de Rfx6 dans les cellules beta, ce qui évoque que MLXIPL régule un répertoire de gènes spécifique aux cellules beta. Cependant, l'expression du gène régulé par MLXIPL dans les cellules beta, Txnip, ne varie pas non plus de manière significative dans les souris delta beta. Cela suggère que l'inhibition d'expression de Mlxipl observée quand Rfx6 est inactivé n'est peut-être pas suffisante à l'identification de l'ensemble de ses gènes cibles spécifiques des cellules beta. La régulation de Txnip par

Mlxipl pourrait également ne pas exister dans les ilots pancréatiques de souris saines étudiés ici et n'avoir lieu que chez des souris diabétiques dans lesquelles son expression est augmentée (Minn et al, 2005a). Une compensation de la diminution d'expression de Mlxipl pourrait également avoir lieu par Foxo1 par exemple. Ceci a en effet déjà été montré pour la régulation de Txnip ou Lpk (Kibbe et al, 2013).

La détermination des gènes régulés par RFX6 de manière dépendante de MLXIPL pourra être réalisée par inactivation de l'expression de Rfx6 et/ou Mlxipl dans les cellules Min6b1 à l'aide de la technologie CrispR-CAS9, récemment mise en œuvre dans l'équipe pour d'autres projets, puis par analyse transcriptomique par RNA Seq.

C. RFX6 et ses perspectives

C.1. RFX6 et cancer

La régulation de Mlxipl par RFX6 indique également un éventuel rôle de ce facteur de transcription dans la prolifération des cellules beta puisque Mlxipl a été identifié comme stimulateur de la prolifération de ces cellules (Metukuri et al, 2012). D'autre part, Mlxipl est également impliqué dans le maintien de la glycolyse aérobie en réponse au glucose et dans l'expansion des cellules tumorales (Tong et al, 2009b). De plus, Rfx6 a été associé très récemment au cancer de la prostate et à la prolifération des cellules cancéreuses (Huang et al, 2014).

La régulation de gènes spécifiques par RFX6 pourrait ainsi avoir un rôle clé dans la perturbation du métabolisme glucidique et la prolifération observées dans les cellules tumorales. Entre autres, les gènes Pim-3 (Pim-3 proto-oncogene, serine/threonine kinase) et Bad (BCL2-associated agonist of cell death) sont ciblés et régulés par RFX6 dans la cellule beta. Les sites de fixation de RFX6, que ce soit dans les cellules Min6b1 ou dans les ilots de Langerhans, sont situés au niveau des régions promotrices de ces gènes et leur expression est augmentée de 1,7 et 1,3 fois respectivement pour Pim3 et Bad. Il a été montré que Pim-3 stimule l'angiogenèse de tissus cancéreux humains et est impliqué dans la régulation de l'apoptose et du cycle cellulaire (Li & Mukaida, 2014; Xu et al, 2014). Par ailleurs, Pim-3 est réprimé dans les cellules beta pancréatiques et son expression est induite en réponse au glucose. Lorsque Pim-3 est inactivé, la sécrétion d'insuline en réponse au glucose est stimulée (Vlacich et al, 2010). Dans les cancers du pancréas et du colon, Pim-3 inactive l'expression de Bad, et semble ainsi prévenir l'apoptose des cellules cancéreuses (Li et al,

2006b; Popivanova et al, 2007). RFX6 pourrait ainsi avoir un rôle crucial quant à la stimulation de la prolifération et de la survie des cellules cancéreuses.

C.2. RFX6 dans l'intestin

Enfin, dans le contexte du domaine de recherche de notre équipe, focalisé sur les cellules endocrines pancréatiques mais également intestinales, la recherche des gènes cibles de RFX6 dans l'intestin devrait permettre de comparer les jeux de données de ChIP Seq dans les ilots pancréatiques et dans l'intestin, et en conséquence de définir les gènes spécifiquement ou conjointement régulés dans les deux organes. Le rôle de Rfx6 dans l'intestin a déjà été proposé par Suzuki et al. (Suzuki et al, 2012) qui montrent que Rfx6 est exprimé dans le sous-type cellulaire K de l'intestin et que ce facteur de transcription cible et régule positivement GIP par liaison au niveau d'une région située entre 5 et 6kb en amont du TSS du gène. Dans notre équipe, la liaison de Rfx6 au niveau de 3 Xbox localisées 4kb en amont du TSS de Gip a également été montrée. D'autre part, une souris permettant l'inactivation conditionnelle de Rfx6 dans l'intestin a été développée. Des études (non publiées à ce jour) réalisées principalement par Julie Piccand, ancienne membre de notre équipe, montrent que Rfx6 est exprimé dans les cellules entéroendocrines exprimant Ngn3 en développement mais également dans la plupart des cellules intestinales sécrétrices d'hormones. La délétion constitutive de Rfx6 dans l'intestin est létale quelques jours après la naissance. Lorsque Rfx6 est inactivé spécifiquement dans les cellules intestinales adultes, la production de la plupart des hormones intestinales est altérée et des problèmes d'absorption des lipides et d'assimilation de la nourriture ainsi qu'une homéostasie du glucose perturbée sont observés. Ngn3 contrôle la destinée entéroendocrine des cellules intestinales lors du développement et nous avons montré que RFX6 réprime directement la transcription de Ngn3 par liaison à des Xbox situées dans les séquences promotrices du gène. Rfx6 semble ainsi essentiel à la spécification entéroendocrine des cellules intestinales. Une première tentative de ChIP Seq dans la lignée cellulaire intestinale neuroendocrine STC-1 avec l'anticorps anti-RFX6 a été réalisée et devra être analysée en regard des résultats de RNA Seq chez la souris n'exprimant pas RFX6 spécifiquement dans l'intestin.

Annexes

Perrine STRASSER – Curriculum vitae

FORMATION

2011-2015	Doctorat en Biologie moléculaire et cellulaire et biologie du développement financé par une bourse FRM (3 ans) et ARC (1 an) et dirigé par le Docteur Gérard Gradwohl à l'IGBMC : <i>Rôle du facteur de transcription RFX6 dans la différenciation et la fonction des cellules β sécrétrices d'insuline : identification et étude de gènes cibles</i>
2009-2011	Master Biochimie, Biologie Moléculaire et Cellulaire, Parcours Biologie et Génétique Moléculaires à l'Université de Strasbourg (13.67/20 rang 4/22)
2008-2009	Troisième année de licence Sciences du Vivant spécialité Biochimie et Biologie Moléculaire à l'Université de Strasbourg, options Connaissance du médicament, Biologie cellulaire, Techniques de biologie cellulaire, Virologie et pathogenèse, Structure des protéines et des acides nucléiques (14/20)
2006-2008	DUT Génie biologique à l'IUT Louis Pasteur de Schiltigheim (14,2/20 Major de promotion)
2004-2006	Première année de faculté de médecine à l'Université Louis Pasteur de Strasbourg
2003-2004	Baccalauréat série scientifique, option physique-chimie

EXPERIENCES

Janvier à Juin 2011 (6 mois)	Stage de M2 à l'Institut de Génétique et de Biologie Moléculaire et Cellulaire de Strasbourg dans le programme de Biologie du Développement au sein l'équipe de Gérard Gradwohl : Etudes préliminaires en vue de l'identification des gènes cibles du facteur de transcription Rfx6 dans la lignée cellulaire Min6b1. Tuteur : Dr. Gérard GRADWOHL
Juin 2010 (1 mois)	Stage à l'Institut de Sciences et d'Ingénierie Supramoléculaire de Strasbourg dans le laboratoire de Biologie Chimique (CNRS UMR 7006) au sein de l'équipe d'Andrew Griffiths : Mise en place de l'outil microfluidique dédié à l'évolution dirigée. Tuteur : Dr. Michaël RYCKELYNCK
Juin 2009 (1 mois)	Stage à l'Institut de Biologie Moléculaire et Cellulaire de Strasbourg dans l'unité de recherche Architecture et Réactivité de l'ARN (CNRS UPR 9002) au sein de l'équipe de Philippe Dumas : Etude d'interactions aminoglycosides-ARN. Tuteur : Dr. Eric ENNIFAR (travaux ayant donné lieu à publication)
Mai à Juillet 2008 (3 mois)	Stage de deuxième année d'IUT Génie biologique au Centre européen d'étude du Diabète de Strasbourg : Rôle des chémokines dans les réactions inflammatoires et la perte d'îlots pancréatiques au cours de la transplantation. Tuteur : Dr. Séverine SIGRIST
Juillet 2007 (1 mois)	Stage de première année d'IUT Génie biologique dans l'entreprise Café SATI de Strasbourg : Granulométrie et contrôle qualité du café. Tuteur : Sébastien MAURER
Mai-Juillet 2006 (1 mois et demi)	Préparation de commandes chez Alliance santé, distributeur pharmaceutique à Strasbourg.
2004-2005-2006 2007-2008-2009	Co-animation du stand Café SATI à la Foire Européenne de Strasbourg (10 jours / an)

COMPETENCES

Langues : Français (langue maternelle), Anglais, Allemand, Italien, Alsacien
Informatique : Utilisatrice avertie d'Excel, Word, Powerpoint, Recherche documentaire
Diplômée du Certificat Informatique et Internet (C2i), 2007

INFORMATIONS COMPLEMENTAIRES

Date de naissance : 17/05/1986

Attestation de formation aux premiers secours (AFPS), 2004

Participation à des évènements de communication et sensibilisation :
Fête de la Science 2011 et 2012

Participation active à l'association des Doctorants et Post-Doctorants de l'IGBMC (SPB : spb.igbmc.fr) depuis 2011 : Organisation puis Responsable des activités Career lunch en 2011 et 2012, Responsable local de l'organisation du Forum Biotechno 2013, Présidente de l'association en 2013/2014

Encadrement de stagiaires :

Stage de Mégane DENU - BTS Biotechnologie première année de 8 semaines : Mutagenèse dirigée du facteur de transcription humain Rfx6 et étude de l'expression de ce facteur.

Stage de Emélie LUDMANN - Master 1 de Biologie Moléculaire et Cellulaire Intégrée de 7 semaines : Mise en place d'outils pour l'étude fonctionnelle du facteur de transcription Rfx6 humain.

Stage de Julia GRANS – Master 2 franco-allemand de Biologie Moléculaire et Cellulaire Intégrée de 6 mois :
Role of Rfx6 transcription factor in pancreatic beta cell function: Preliminary studies on the transcriptional regulation of Mlxipl/Chrebp a novel direct target of Rfx6

PUBLICATIONS ET COMMUNICATIONS SCIENTIFIQUES

Présentation d'un poster lors de la retraite de Département Biologie du Développement et Cellules Souches de l'IGBMC en avril 2015 "Genome-wide binding and transcriptome analysis revealed Mlxipl as a novel regulated direct target of RFX6 transcription factor in beta cells" P Strasser, J Grans, J Piccand, A Meunier and G Gradwohl.

Rfx6 maintains the functional identity of adult pancreatic β cells. Piccand J, Strasser P, Hodson DJ, Meunier A, Ye T, Keime C, Birling MC, Rutter GA, Gradwohl G. Cell Rep. 2014

Structure-guided discovery of a novel aminoglycoside conjugate targeting HIV-1 RNA viral genome. Ennifar E, Aslam MW, Strasser P, Hoffmann G, Dumas P, van Delft FL. ACS Chem Biol. 2013

Participation au congrès international TriRhena 1st Metabolism Club à Freiburg en Allemagne en avril 2014 : "The winged helix transcription factor Rfx6 controls beta cell identity and function in adult mice". J Piccand, P Strasser, A Beucher, M Poulet, A Meunier, G Rutter, G Gradwohl.

Participation au congrès international de l'European Association for the Study of Diabetes à Sitges en Espagne en septembre 2013 : présentation d'un poster "Targets of the winged-helix transcription factor Rfx6 in beta-cells" P Strasser, J Piccand, M Poulet, A Meunier and G Gradwohl.

Présentation orale du Séminaire de Département Biologie du Développement et Cellules Souches de l'IGBMC en septembre 2013 "Identification of Rfx6 target genes, a winged-helix transcription factor controlling beta cells development and causing monogenic diabetes"

Présentation d'un poster lors de la retraite de Département Biologie du Développement et Cellules Souches de l'IGBMC en avril 2013 "Targets of the winged-helix transcription factor Rfx6 in beta cell line" P Strasser, J Piccand, M Poulet, A Meunier and G Gradwohl.

Structure-Guided Discovery of a Novel Aminoglycoside Conjugate Targeting HIV-1 RNA Viral Genome

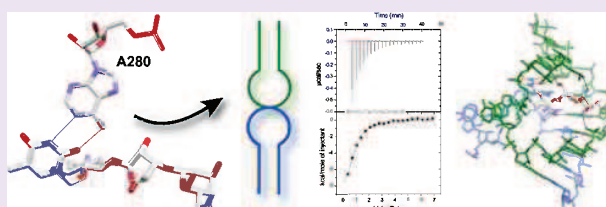
Eric Ennifar,^{*,†} Muhammad Waqar Aslam,[‡] Perrine Strasser,[†] Guillaume Hoffmann,[†] Philippe Dumas,[†] and Floris L. van Delft^{*,‡}

[†]Architecture et Réactivité de l'ARN, Institut de Biologie Moléculaire et Cellulaire, CNRS, Université Louis Pasteur, 15 rue René Descartes, 67084 Strasbourg, France

[‡]Institute for Molecules and Materials, Radboud University Nijmegen, Heijendaalseweg 135, 6525 AJ, Nijmegen, The Netherlands

Supporting Information

ABSTRACT: The dimerization initiation site (DIS) of the HIV-1 genomic RNA is a conserved stem-loop that promotes viral genome dimerization by forming a loop-loop complex. The DIS constitutes a potentially interesting target because it is crucial for several key steps of the viral replication. In this work we describe the synthesis of a rationally designed aminoglycoside conjugate that binds the HIV-1 DIS viral RNA with high specificity, as shown by an extensive *in vitro* binding characterization. We propose a three-dimensional model of the drug–RNA interaction that perfectly fits with binding data. Our results show the feasibility of targeting the HIV DIS viral RNA dimer and open the way to the rationale design of a new class of antiviral drugs. In addition, due to similarities between the HIV-1 DIS RNA and the bacterial aminoacyl decoding site (A site) RNA, we show that this novel aminoglycoside conjugate also binds the bacterial A site with a similar affinity as natural aminoglycoside antibiotics.



In the absence of a vaccine and in the context of increasing prevalence of resistant strains, identifying and validating new therapeutic targets and finding new inhibition strategies remain major goals of HIV research. Beside the viral proteins and cellular cofactors of HIV-1 replication, the genomic RNA itself has been proposed as an interesting target for antiviral agents. All retroviruses, including HIV-1, encapsidate their genome as a dimer of two single-stranded RNAs that are noncovalently linked close to their 5'-ends.¹ The HIV-1 dimerization initiation site (DIS) is a highly conserved sequence in the 5' noncoding region of the viral genomic RNA.^{2,3} It was shown that alteration of the DIS affects several key steps of the viral replication such as RNA dimerization, packaging, and reverse transcription and dramatically reduces viral infectivity, thus making the DIS a potentially interesting RNA target.^{4,5} The DIS loop initiates genome dimerization by forming a loop-loop complex that is further stabilized into an extended duplex form upon interaction with the viral NCp7 nucleocapsid protein (Figure 1a).^{6–9}

X-ray structures of the DIS kissing-loop complex and extended duplex forms^{10–13} revealed surprising structural and sequence similarities with the bacterial 16 S rRNA aminoacyl-tRNA decoding site (A-site), which is the target of aminoglycoside antibiotics.^{14,15} As a result, it was shown that 4,5-disubstituted 2-deoxystreptamine (2-DOS) aminoglycosides bind the “A site motif” within HIV-1 DIS *in vitro* with a high specificity and with a significantly higher affinity than their natural target, the bacterial A site.^{16–20} Aminoglycoside binding to the viral RNA induces a strong stabilization of the DIS

kissing-loop, thus interfering with its conversion into the extended duplex form.¹⁶ Importantly, binding of aminoglycosides to the DIS was also observed on the complete viral genome *ex vivo* in HIV-infected human cells and in viral particles.¹⁸ Crystal structures of DIS kissing-loop complex and extended duplex bound to several aminoglycosides were solved,^{18,21} disclosing the molecular requirements for the drug/RNA recognition and opening the way to structure-based design of novel potential drugs.

Aminoglycoside antibiotics were among the first antibacterial agents active against both Gram-negative and Gram-positive pathogens and are currently still important for the treatment of some pathogens. Their interactions with the bacterial A-site RNA induce mRNA decoding errors, block mRNA and tRNA translocation, and inhibit ribosome recycling.^{22–24} However, their relative toxicity and the rapid development of resistant strains prompted extensive research efforts in order to discover aminoglycoside derivatives with improved potency against resistant pathogens and reduced side effects.²⁵ In addition, intense research also focused on the development of novel aminoglycosides binding to new RNA targets, especially conserved and structured HIV-1 RNA sequences. Aminoglycoside dimers were thus developed to confer selectivity for the HIV-1 RRE (rev response element),²⁶ the TAR (trans-activation responsive element),^{27,28} or the DIS.²⁹ The develop-

Received: July 5, 2013

Accepted: September 9, 2013

Published: September 9, 2013



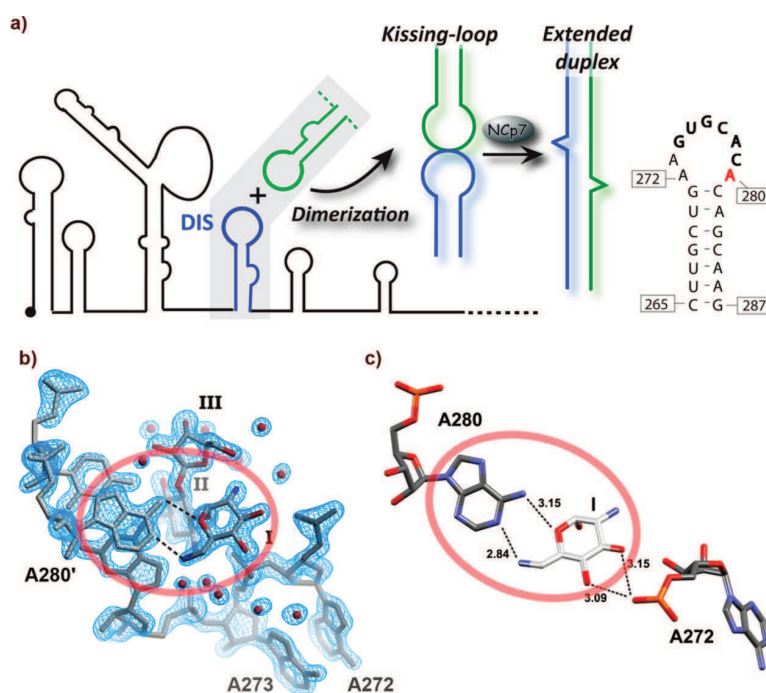


Figure 1. (a) HIV-1 genomic RNA dimerization mechanism and RNA sequence corresponding to the HIV-1 subtypes A and F used in this study. The self-complementary sequence is bolded, and the adenine 280 is highlighted in red. (b and c) Detailed views of the HIV-1 DIS-ribostamycin complexes showing the pseudo Watson–Crick base-pair. (b) DIS extended duplex crystal structure (PDB ID 3C3Z) superimposed with the electron density map. (c) DIS kissing-loop complex crystal structure (PDB ID 2FCZ).

ment of nucleobase-aminoglycoside conjugates was also explored to target the HIV-1 TAR.³⁰

In this work, by combining previous structural and thermodynamic data about DIS/aminoglycoside complexes, we could rationally design and synthesize a nucleobase-aminoglycoside conjugate that targets the HIV-1 DIS RNA dimer. The specificity of the interaction was assessed by an extensive binding study based on isothermal titration calorimetry (ITC). Finally, we could show that our synthetic analogue also recognize *in vitro* the bacterial A-site with a similar affinity to that of natural aminoglycosides.

RESULTS AND DISCUSSION

Design and Synthesis of the Thymine-Neomycin Aminoglycoside Conjugate. Examination of all DIS/aminoglycoside crystal structures revealed that the highly conserved adenine 280 of the viral RNA genome is involved in a key interaction with ring I of aminoglycosides (Figure 1b,c). We therefore hypothesized that replacing ring I by a thymine nucleobase would lead to similar interactions, forming a Watson–Crick base-pairing between A280 and the drug, without significantly affecting the specificity and affinity of binding. Because neomycin was found to be the best DIS binder,¹⁶ a neomycine-thymine conjugate was designed (compound **1**, Figure 2).

The synthesis of the envisaged thymine-neomycin conjugate **1** started with the commercially available neomycin B sulfate, which was transformed to fully protected peracetylated hexaazidoneomycin **2**³⁸ via transition metal catalyzed diazotransfer, followed by global O-acetylation. We argued that Lewis acid promoted fragmentation of the peracetylated hexaazidoneomycin **2**, as reported by Swayze,³⁹ followed by glycosylation of the resulting fragments after protective group

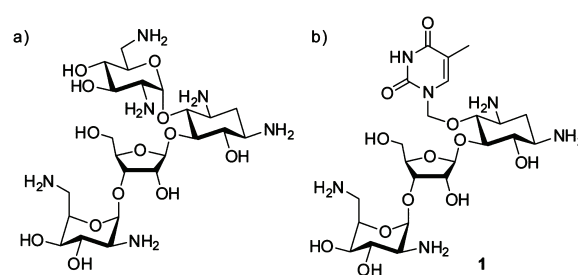
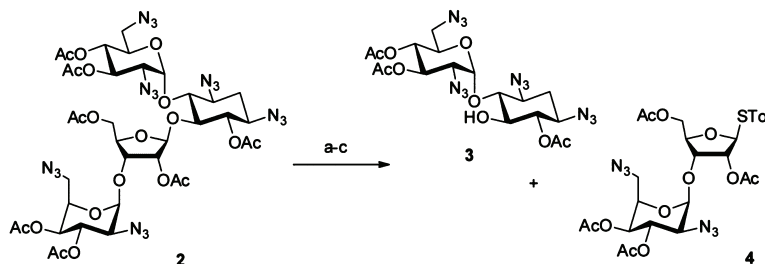


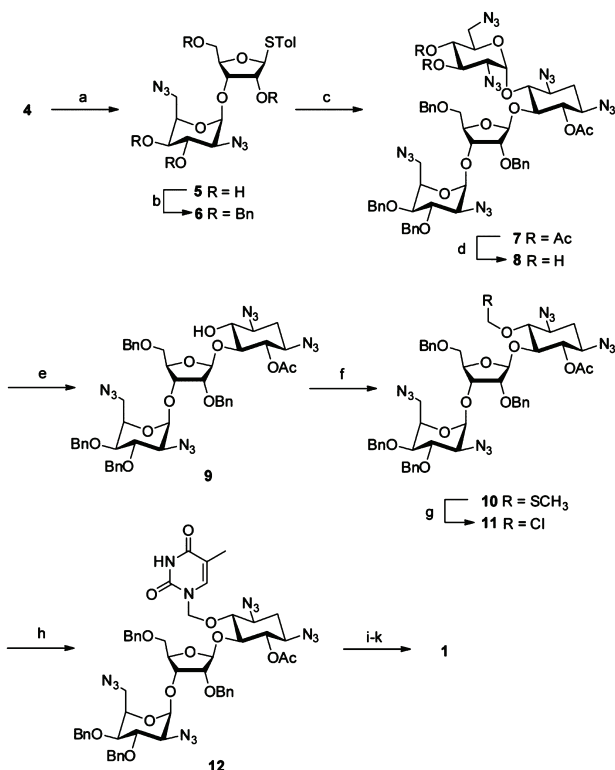
Figure 2. (a) Structure of neomycin. (b) Structure of thymine-neomycin conjugate **1**.

interconversion, would differentiate the two similar rings I and IV. Thus, compound **2** was treated with 3 equiv of $\text{BF}_3 \cdot \text{OEt}_2$ and 1.1 equiv of *p*- CH_3PhSH (TolSH) in CH_2Cl_2 to yield 1,3,2',6'-tetraazido-tri-O-acetyl-D-neamine **3** and per-O-acetyl-neobiosamine **4** (Scheme 1). Unfortunately, the R_f difference between compound **3** and **4** was too small to allow chromatographic separation, and therefore the crude mixture was treated with hexamethyldisilazane and catalytic TMSCl in CH_3CN to convert the free 5-OH of **3** to the more lipophilic TMS ether, followed by a facile separation step and acid hydrolysis of the TMS ether to afford the pure compound **3** in good yield. The resulting disaccharide **4** has the proper functionality at the ribose anomeric center for activation with thiophilic promoters and to allow glycosylation after protective group interconversion.

In the next step, per-O-acetyl-neobiosamine **4** was treated with KO^tBu in MeOH to remove the acetyl groups, followed by conversion of the unprotected compound **5** into benzyl derivative **6** upon treatment with NaH and BnBr (Scheme 2). Next, glycosylation of **3** with thioglycoside donor **6** was

Scheme 1^a

^aReagents and conditions: (a) TolSH (1.1 equiv), $\text{BF}_3 \cdot \text{OEt}_2$ (3.0 equiv), CH_2Cl_2 , rt, 15 h. (b) HMDS, TMSCl, CH_3CN , 3 h. (c) 1 M HCl, THF, rt, 30 min. (3, 82%; 4, 74%).

Scheme 2^a

^aReagents and conditions: (a) KO^tBu , MeOH, rt, 2 h (91%). (b) NaH, BnBr, DMF, 0 °C \rightarrow rt, 3 h (83%). (c) 3, NIS, AgOTf , CH_2Cl_2 , MS 4 Å, $-40 \rightarrow -10$ °C, 1 h (51%). (d) KO^tBu , MeOH, rt, 2 h (65%). (e) (i) NaIO_4 , MeOH, rt, 16 h. (ii) *n*-BuNH₂, MeOH, rt, 16 h (70% over two steps). (f) DMSO, Ac_2O , AcOH, 16 h (60%). (g) SO_2Cl_2 , CH_2Cl_2 , 1 h (70%). (h) 5-Methyl-2,4-bis(trimethylsilyloxy)-pyrimidine, *n*-Bu₄NI, $\text{C}_2\text{H}_4\text{Cl}_2$, 40 h, reflux (80%). (i) LiOH, MeOH, 1,4-dioxane, H₂O, 48 h (79%). (j) PMe_3 , THF, 0.1 M NaOH, 4 h (84%). (k) $\text{Pd}(\text{OH})_2$, H₂, AcOH, H₂O, 4 h (94%); TFA, MeOH.

attempted. After several attempts, it was found that NIS-promoted coupling of neamine acceptor 3 with thioglycoside donor 6 in CH_2Cl_2 gave the desired compound 7 in reasonable yield and with a favorable α/β ratio of approximately 1:6.5. Since it was found impossible to separate the anomers, we continued with the mixture of anomers with the objective to separate them at a later stage of the synthesis. The next target was to remove ring I of the resulting neomycin derivative 7 so thymine could be introduced at 4-OH. As anticipated, the selective deacetylation⁴⁰ of compound 7 with KO^tBu in MeOH

afforded monoacetate 8 in high yield. Oxidative cleavage of the unprotected diol functionality 8 with NaIO_4 in MeOH and subsequent elimination of the resulting dialdehyde proceeded uneventfully to yield pseudotrisaccharide 9 with the 4-OH free.

Finally, compound 9 was treated with DMSO, Ac_2O , and AcOH to produce the methyl thiomethyl ether 10.⁴¹ Reaction of compound 10 with SO_2Cl_2 in CH_2Cl_2 gave chlorinated compound 11, which was refluxed overnight with bis-(trimethylsilyl)thymine and *n*-Bu₄NI in dichloroethane to afford the desired thymine-substituted derivative 12. We then proceeded to remove all protecting groups. First, saponification of the acetyl ester of compound 12 was attempted under Zemplén conditions, but the acetyl group was found unaffected. After exploring a range of conditions, we succeeded in removing the acetyl group by treating compound 12 with LiOH in a mixture of H₂O and dioxane for 48 h. Next, azides were reduced by treating with PMe_3 in a mixture of THF and aqueous NaOH, and finally benzyl ethers were removed by hydrogenolysis in the presence of Pearlman's catalyst⁴² in aqueous AcOH/MeOH, to afford fully deprotected thymine-substituted neomycin analogue 1 as the tetraacetic acid salt. For analytical purposes, acetate counterions were exchanged for trifluoroacetates by lyophilization from 2% TFA in MeOH.

Molecular Modeling of the Drug–RNA Interaction.

Based on known HIV DIS RNA/neomycin structures, a 3D model of the DIS kissing-loop/compound 1 complex was built. Two molecules of compound 1 were placed in the DIS kissing-loop homodimer, similarly to the parent compound neomycin. This model shows that with a small accommodation of ring II, which was slightly rotated compared to the DIS/neomycin complex, all direct RNA/neomycin interactions involving rings II (the 2-DOS ring), III, and IV can be preserved by forming a Watson–Crick base-pairing between the thymine moiety of compound 1 and adenine 280 (Figures 3 and 4, Supplementary Figure 7). However, four direct contacts involving ring I of neomycin with phosphates 272 and 273 are lost due to the substitution into thymine (Figure 3). Interestingly, these four drug-phosphate RNA interactions are not present in the neomycin/ribosomal A site complex due to the difference in RNA topology between the A site and the HIV DIS kissing-loop. In addition, following accommodation of ring II, distances between N1, N3, and O6 and their respective RNA ligands slightly increased. Possible consequences of this distance change and loss of direct interactions will be discussed later. Regarding stacking interactions, an optimal stacking was observed in crystal structure between ring I of natural aminoglycosides and guanosine 271, but no stacking was observed with guanosine 274. The situation is different in the

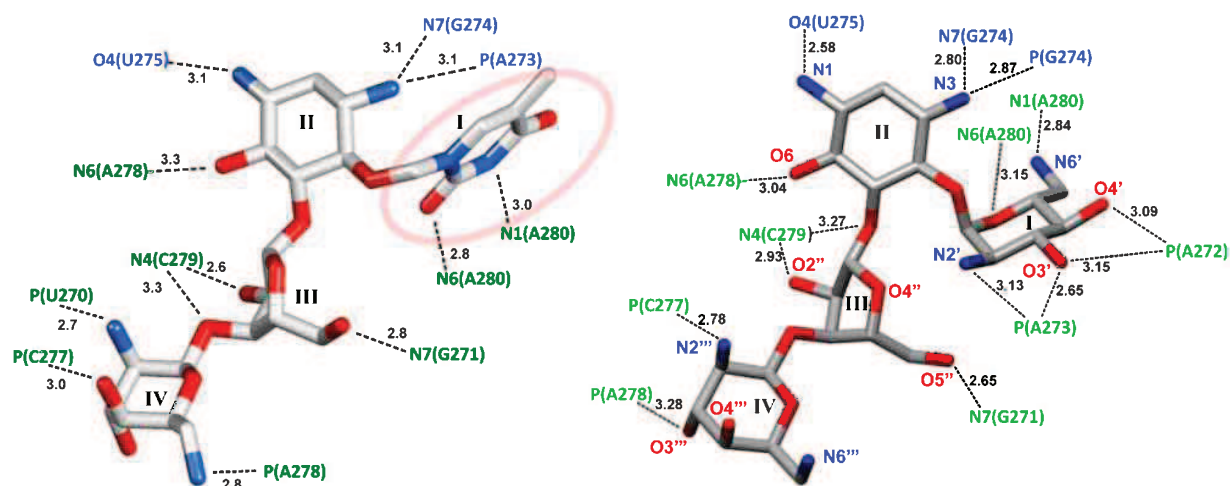


Figure 3. Comparison of predicted RNA–drug contacts for compound **1** (left) with observed RNA–neomycin interactions (PDB ID 2FCY, right). The thymine moiety is red-circled. Distances are indicated in angstroms. RNA ligands are indicated in green or blue depending on the strand to which they belong.

present model with compound **1**, likely because of the aforementioned obligatory accommodation and the rigidity of the thymine base in comparison of the flexible ring I. In the model, partial stacking interactions are expected both with guanosine 271 and guanosine 274 (Figure 4). These differences might also affect the affinity of compound **1** compared to that of neomycin.

In Vitro Binding Assay of Compound 1 to the HIV-1 DIS RNA Loop–Loop Complex. The binding of compound **1** to the HIV-1 DIS kissing-loop RNA was then evaluated *in vitro* using isothermal titration calorimetry (ITC) on a 23-nucleotide fragment containing the minimal DIS hairpin loop (Figure 1a, Supplementary Figure 1). ITC is a true in-solution technique that directly measures the heat released or absorbed during a reaction, providing in one single experiment the complete binding profile (ΔH , ΔS , affinity constant, and stoichiometry) of this reaction. ITC experiments performed at 25 °C in the presence of 25 mM sodium cacodylate pH 7.0, 2 mM $MgCl_2$, and 100 mM KCl (“low salt” conditions) confirmed that, as predicted, compound **1** binds the DIS kissing-loop (Table 1) with a 2:1 stoichiometry (two ligands per DIS kissing-loop dimer). However, as anticipated from the model suggesting the loss of four direct drug/RNA hydrogen bonds and a nonoptimal stacking of the thymine moiety with the guanosine 271, the affinity is decreased compared to that of natural aminoglycosides (Supplementary Figures 2, 3, and 4). This decrease in affinity is also partly due to the loss of two amino groups (compared to the parent compound neomycin), which are mostly positively charged at pH 7.0 and therefore strongly contribute to affinity through electrostatic interactions.

In order to investigate the specificity of the binding, ITC experiments were then performed at higher salt concentration. In presence of 200 mM KCl (“high salt” conditions), unspecific interactions of paromomycin and lividomycin aminoglycosides with the DIS RNA are prevented (Supplementary Figure 5) but the specific binding is still preserved. Due to additional positively charged amino groups, unspecific interactions are still observed (but strongly reduced) with neomycin in these conditions (Supplementary Figures 4 and 5). In “high salts” conditions, compound **1** binds the DIS kissing-loop, and no unspecific interaction was detected (Figure 5). However, the

binding occurs with a 7-fold loss in affinity ($K_d = 5.3 \mu M$) compared to “low salt” conditions. The specificity of binding was further assessed using various RNA sequences. No significant binding to compound **1** was detected by ITC on the ribosomal 23 S sarcin-ricin hairpin loop (SRL), which does not bind any aminoglycoside specifically, or on the HIV-1 TAR hairpin loop, which binds neomycin but in a different geometry not involving Watson–Crick-like base pairing with ring I,⁴³ or on a 22-base-pair duplex corresponding to the DIS sequence deprived of “A site motif” (Figure 5). Finally, the binding of compound **1** was also evaluated on a DIS A280U mutant sequence. This adenine to uridine mutation should disrupt the Watson–Crick hydrogen bonds between residue 280 and the thymine moiety of compound **1**. As expected, this sequence does not bind the neomycin-thymine conjugate (Figure 5). This clearly assesses the specificity of compound **1** binding to the HIV DIS dimer and the formation of the expected A–T base-pair between the drug of the viral RNA.

We have previously shown that DIS sequences with a $G_{274}CGCGC_{279}$ self-complementary sequence found in HIV subtype B are not able to interact specifically with aminoglycosides due to a steric hindrance between the amino group of C275 and N1 of ring II.^{18,19} The binding of compound **1** to HIV-1 subtype B DIS kissing-loop was therefore also investigated. Surprisingly, ITC data revealed that compound **1** also interacts specifically with this DIS sequence both in low (Table 1) and high salt conditions ($K_d = 5.6 \mu M$), with an affinity very similar to the one observed for the subtype F DIS (Figure 5). Very likely, the binding is possible due to the slight obligatory accommodation of ring II in compound **1**, as previously mentioned, thus avoiding any steric hindrance between the RNA and the ligand. This observation is important since compound **1** is the first ligand that is able to bind indifferently any HIV DIS RNA dimer, thus circumventing the serious limitation of natural aminoglycosides, which could only bind DIS RNA having the $G_{274}UGCAC_{279}$ self-complementary sequence.

Interestingly, though binding affinities of compound **1** to HIV-1 DIS subtype F and B are similar, enthalpy and entropy changes are different in low salt conditions (Table 1). The origin of this enthalpy/entropy compensation is not obvious

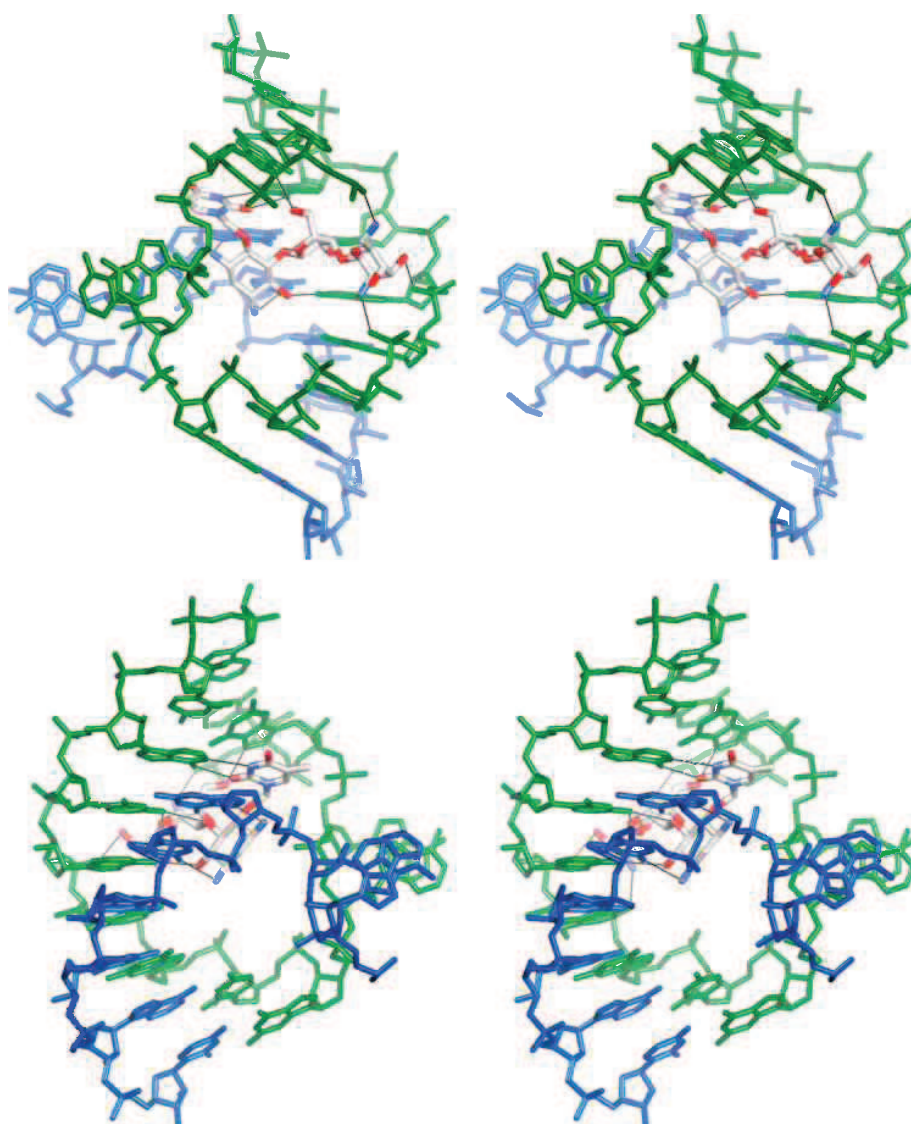


Figure 4. Two different stereo views (rotated by 180° along the RNA axis) of the 3D model of the HIV-1 DIS kissing-loop bound to compound 1. The two RNA strands are shown in blue and green. Direct hydrogen bonds are depicted in dark gray.

Table 1. Binding Parameters Obtained by ITC Measurements for Compound 1 in Buffer Containing 100 mM KCl

RNA sequence	N^a	K_d (nM)	ΔH (kcal mol ⁻¹)	$-T\Delta S$ (kcal mol ⁻¹)
HIV DIS-F	0.96	730	-8.0	-0.5
HIV DIS-B	1.01	752	-13.6	5.0
bacterial A site	1.12	507	-7.2	-1.5

^aStoichiometry for one RNA strand.

from the model showing final bound state of drug/RNA interactions, since both complexes are expected to be almost identical. However, a major difference between both DIS subtypes is found in unbound structures: whereas a specific hexahydrated Mg^{2+} was found in crystal structures of the DIS subtype F loop-loop helix,¹⁰ no cation was found in subtype B.¹² Consequently, displacement of a specific hexahydrated Mg^{2+} is a prerequisite for the binding of compound 1 to the subtype F DIS, but not to the subtype B DIS (disordered

monovalent ions, instead of an ordered divalent cation, are likely displaced following ligand binding to this RNA sequence) and this could be the origin of the observed differences in ΔH and ΔS for these two RNA. Supporting this hypothesis, almost no difference in ΔH and ΔS are observed for ITC experiments performed in higher monovalent salt concentrations where Mg^{2+} cations might be partially screened by K^+ ions (Figure 5).

Binding of Compound 1 to the Bacterial Ribosomal A Site. Interactions between the bacterial A site and natural aminoglycosides have been extensively investigated using biophysical approaches,^{15,44–52} resulting in the identification of modified aminoglycosides with improved binding properties.^{53–58} Since the pseudo base-pairing interaction between ring I and adenine 1408 of bacterial rRNA is also conserved in A site/aminoglycoside interactions,^{15,45,51,52} we anticipated that compound 1 could also bind the bacterial A site *in vitro*. This interaction was therefore investigated by ITC on a 27-nucleotide RNA fragment containing the bacterial A site.⁵⁹ As expected, a specific 1:1 binding was observed (Table 1), but

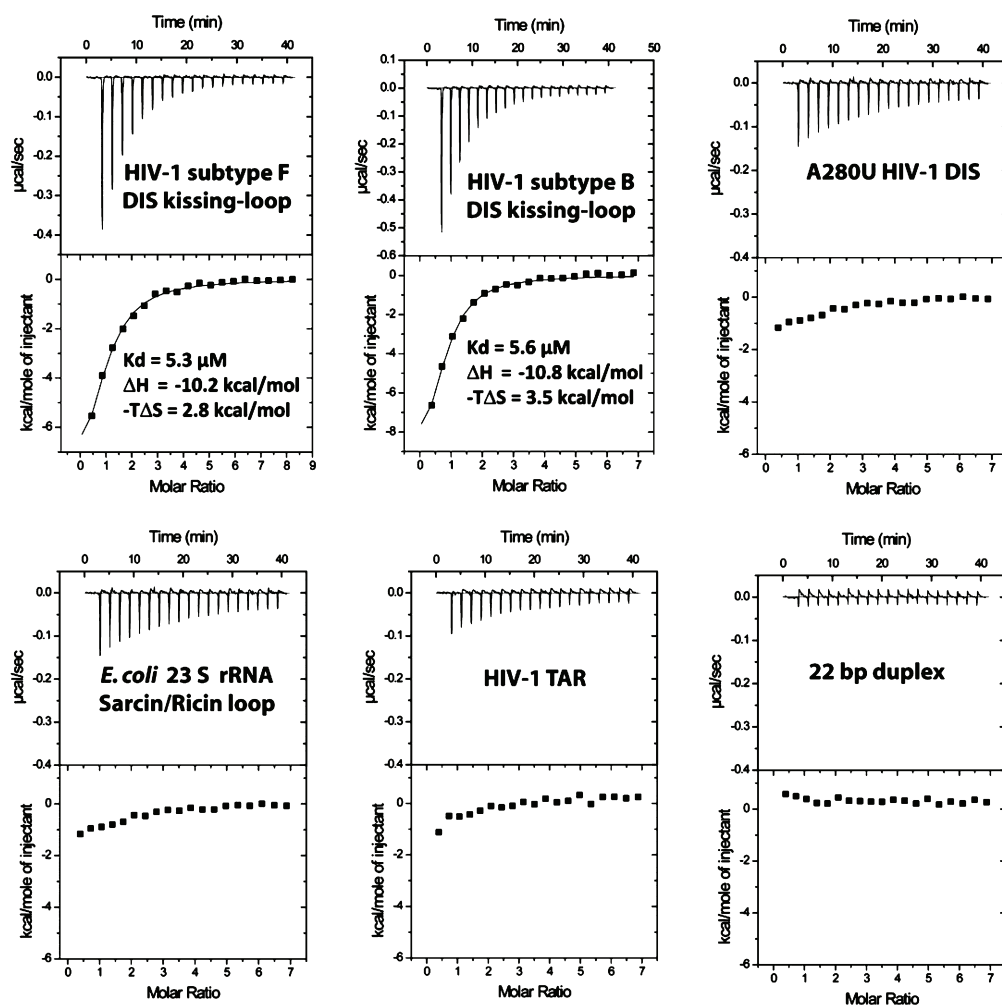


Figure 5. Isothermal titration calorimetry experiments showing the specific binding of compound **1** to the DIS kissing-loops from HIV-1 subtypes F and B. The binding is lost by mutating the adenine 280, which should interact with the thymine moiety, into uridine. No binding is observed on a 22 base-pair duplex, the ribosomal SRL hairpin loop or on the HIV-1 TAR loop.

again with a slight decrease in affinity compared to natural aminoglycosides (Figure 6, Supplementary Figure 6 and Supplementary Table 1). This loss is however rather limited (2-fold loss compared to paromomycin and neomycin) since, in contrast with the previous situation with the HIV DIS, no direct interaction should be lost according to known A site/neomycin crystal structure.⁵⁸

Conclusions. We rationally designed and synthesized an aminoglycoside conjugate by replacing ring I with a thymine moiety, in order to form Watson–Crick hydrogen bonds with adenine 280 of the viral RNA. We showed that this conjugate **1** is able to bind DIS kissing-loops with an improved specificity compared to that of the parent compound neomycin, but only partial loss of affinity. In addition, conjugate **1** binds DIS from HIV-1 subtypes A, B, and F with a similar affinity, thus extending the specificity of binding to HIV-1 subtypes B RNA compared to neomycin. It was previously reported that neomycin inhibits up to 85% of HIV-1 production.⁶⁰ This effect might be due, at least partly, to the observed stabilization of the RNA dimer following drug binding to the DIS sequence.¹⁸ The latter stabilization will induce excessive pauses and/or dissociation of the viral reverse transcriptase during proviral synthesis, thus interfering with virus production. Finally

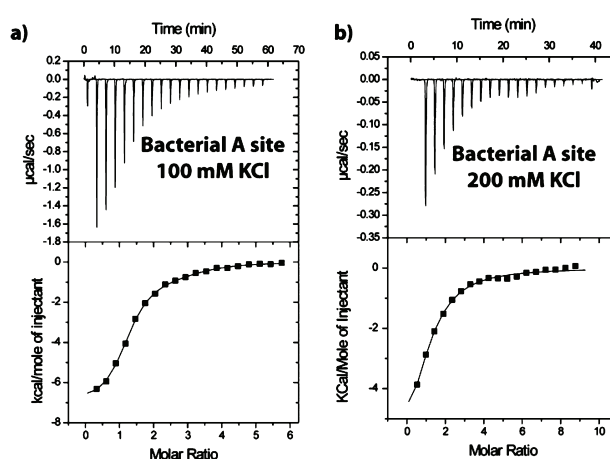


Figure 6. ITC profile for the titration of compound **1** in the bacterial A site in a solution containing either 100 mM KCl (a) or 200 mM KCl (b). Binding parameters are reported in Supplementary Table 1.

we showed that neomycin-thymine conjugate **1** binds the bacterial A site *in vitro* with a similar affinity than neomycin and

paromomycin. It is therefore a potential new antibiotic with improved resistance against aminoglycoside-modifying enzymes, especially those affecting ring I like the neomycin 3'-phosphotransferase.⁶¹ Work is in progress to evaluate the potential antiviral and antibacterial activity of conjugate 1.

METHODS

Chemistry. Preparation of conjugate 1 is described in the Supporting Information.

RNA Sample Preparation. Chemically synthesized RNA sequences were purchased from Integrated DNA technologies or from Dharmacon. RNA were purified as described.³¹ In short, RNA was loaded on a Nucleopac PA-100 column (Dionex) heated at 70 °C equilibrated in 4 M urea, 20 mM MES [2-(*N*-morpholino)ethanesulfonic acid], pH 6.2 and eluted using a NaClO₄ gradient.

DIS kissing-loop complexes were obtained as follows: RNA was diluted to 2 μM concentration in water, heated at 90 °C for 5 min, and then cooled on ice for 5 min. As shown previously, this protocol essentially leads to the formation of DIS loop-loop dimers and prevents the formation of extended duplex dimers.^{32,33} A similar protocol was performed for folding of TAR and SRL hairpins. The 22 base-pair RNA duplex was diluted to 60 μM in water, heated at 90 °C for 5 min, and then cooled at RT. This protocol was used in order to favor the formation of a duplex instead of hairpin monomers. All RNA samples were dialyzed against ITC buffer (25 mM Na cacodylate pH 7.0, 100 or 200 mM KCl, 2 mM MgCl₂) and adjusted to a final concentration of 10 to 60 μM.

Natural Aminoglycosides. Neomycin, paromomycin and lividomycin were purchased from Sigma-Aldrich and used without further purification. Aminoglycosides were dissolved into ITC buffer prior experiments.

Isothermal Titration Calorimetry (ITC). ITC measurements were performed on a MicroCal ITC₂₀₀ (GE Healthcare) in ITC buffer containing 100 or 200 mM KCl. In a typical experiment, 20 injections of 2.0 μL aliquots of drug at 400 μM were injected (at 0.5 μL s⁻¹) into 203 μL of RNA at 12 μM in the sample cell. The delay between injections was 180 s. All ITC curves were analyzed using the software Origin (OriginLab) using either a one-site model (in absence of unspecific binding) or a two sets of sites model (in presence of unspecific binding). Standard free energies of binding (ΔG) were obtained using the equation

$$\Delta G = -RT \ln K_a$$

Entropic contributions were obtained from the relationship

$$\Delta G = \Delta H - T\Delta S$$

using standard free energy of binding determined previously and the binding enthalpies derived from fitted ITC data. In our experimental conditions, the product $K_a \times [\text{RNA}] \times N$, where N is the number of binding sites, lies in the 0.1–1000 range, allowing an accurate and simultaneous determination of binding parameters by ITC.³⁴ Because of a fast ligand-RNA binding, it was not possible to use our recently developed *kinITC* approach to derive kinetic parameters of binding in addition to thermodynamic data.³⁵

Model Building. Modeling was carried out starting from the DIS kissing-loop/neomycin crystal structure¹⁸ (PDB ID 2FCY) using the program O³⁶ followed by energy minimization using CNS.³⁷ The RNA was not fixed during the energy minimization.

ASSOCIATED CONTENT

Supporting Information

Supplementary methods, one table, and seven figures as described in the text. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: e.ennifar@ibmc-cnrs.unistra.fr.

*E-mail: f.vandelft@science.ru.nl.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank T. Langer (Prestwick Chemical), V. Vidal (InPheno AG), and T. Klimkait (University of Basel) for fruitful and stimulating discussions. This work was supported by the Agence Nationale pour la Recherche (grant ANR-12-BS07-0007-03 "ClickEnARN").

REFERENCES

- (1) Paillart, J. C., Shehu-Xhilaga, M., Marquet, R., and Mak, J. (2004) Dimerization of retroviral RNA genomes: an inseparable pair. *Nat. Rev. Microbiol.* 2, 461–472.
- (2) Laughrea, M., and Jetté, L. (1994) A 19-nucleotide sequence upstream of the 5' major splice donor site is part of the dimerization domain of human immunodeficiency virus 1 genomic RNA. *Biochemistry* 33, 13464–13474.
- (3) Skripkin, E., Paillart, J. C., Marquet, R., Ehresmann, B., and Ehresmann, C. (1994) Identification of the primary site of the Human Immunodeficiency Virus Type I RNA dimerization *in vitro*. *Proc. Natl. Acad. Sci. U.S.A.* 91, 4945–4949.
- (4) Paillart, J.-C., Berthou, L., Ottmann, M., Darlix, J.-L., Marquet, R., Ehresmann, C., and Ehresmann, B. (1996) A dual role of the dimerization initiation site of HIV-1 in genomic RNA packaging and proviral DNA synthesis. *J. Virol.* 70, 8348–8354.
- (5) Shen, N., Jette, L., Liang, C., Wainberg, M. A., and Laughrea, M. (2000) Impact of human immunodeficiency virus type 1 RNA dimerization on viral infectivity and of stem-loop B on RNA dimerization and reverse transcription and dissociation of dimerization from packaging. *J. Virol.* 74, 5729–5735.
- (6) Laughrea, M., and Jetté, L. (1996) Kissing-loop model of HIV-1 genome dimerization: HIV-1 RNA can assume alternative dimeric forms, and all sequences upstream or downstream of hairpin 248–271 are dispensable for dimer formation. *Biochemistry* 35, 1589–1598.
- (7) Muriaux, D., Fossé, P., and Paoletti, J. (1996) A kissing complex together with a stable dimer is involved in the HIV-1_{Lai} RNA dimerization process *in vitro*. *Biochemistry* 35, 5075–5082.
- (8) Paillart, J. C., Skripkin, E., Ehresmann, B., Ehresmann, C., and Marquet, R. (1996) A loop-loop "kissing" complex is the essential part of the dimer linkage of genomic HIV-1 RNA. *Proc. Natl. Acad. Sci. U.S.A.* 93, 5572–5577.
- (9) Takahashi, K. I., Baba, S., Chattopadhyay, P., Koyanagi, Y., Yamamoto, N., Takaku, H., and Kawai, G. (2000) Structural requirement for the two-step dimerization of human immunodeficiency virus type 1 genome. *RNA* 6, 96–102.
- (10) Ennifar, E., and Dumas, P. (2006) Polymorphism of bulged-out residues in HIV-1 RNA DIS kissing complex and structure comparison with solution studies. *J. Mol. Biol.* 356, 771–782.
- (11) Ennifar, E., Walter, P., and Dumas, P. (2010) Cation-dependent cleavage of the duplex form of the subtype-B HIV-1 RNA dimerization initiation site. *Nucleic Acids Res.* 38, 5807–5816.
- (12) Ennifar, E., Walter, P., Ehresmann, B., Ehresmann, C., and Dumas, P. (2001) Crystal structures of coaxially stacked kissing complexes of the HIV-1 RNA dimerization initiation site. *Nat. Struct. Biol.* 8, 1064–1068.
- (13) Ennifar, E., Yusupov, M., Walter, P., Marquet, R., Ehresmann, B., Ehresmann, C., and Dumas, P. (1999) The crystal structure of the dimerization initiation site of genomic HIV-1 RNA reveals an extended duplex with two adenine bulges. *Structure* 7, 1439–1449.
- (14) Carter, A. P., Clemons, W. M., Brodersen, D. E., Morgan-Warren, R. J., Wimberly, B. T., and Ramakrishnan, V. (2000) Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* 407, 340–348.

- (15) Vicens, Q., and Westhof, E. (2001) Crystal structure of paromomycin docked into the eubacterial ribosomal decoding A site. *Structure* 9, 647–658.
- (16) Bernacchi, S., Freisz, S., Maechling, C., Spiess, B., Marquet, R., Dumas, P., and Ennifar, E. (2007) Aminoglycoside binding to the HIV-1 RNA dimerization initiation site: thermodynamics and effect on the kissing-loop to duplex conversion. *Nucleic Acids Res.* 35, 7128–7139.
- (17) Ennifar, E., Paillart, J. C., Bernacchi, S., Walter, P., Pale, P., Decout, J. L., Marquet, R., and Dumas, P. (2007) A structure-based approach for targeting the HIV-1 genomic RNA dimerization initiation site. *Biochimie* 89, 1195–1203.
- (18) Ennifar, E., Paillart, J. C., Bodlener, A., Walter, P., Weibel, J. M., Aubertin, A. M., Pale, P., Dumas, P., and Marquet, R. (2006) Targeting the dimerization initiation site of HIV-1 RNA with aminoglycosides: from crystal to cell. *Nucleic Acids Res.* 34, 2328–2339.
- (19) Ennifar, E., Paillart, J. C., Marquet, R., Ehresmann, B., Ehresmann, C., Dumas, P., and Walter, P. (2003) HIV-1 RNA dimerization initiation site is structurally similar to the ribosomal A site and binds aminoglycoside antibiotics. *J. Biol. Chem.* 278, 2723–2730.
- (20) Tam, V. K., Kwong, D., and Tor, Y. (2007) Fluorescent HIV-1 Dimerization Initiation Site: design, properties, and use for ligand discovery. *J. Am. Chem. Soc.* 129, 3257–3266.
- (21) Freisz, S., Lang, K., Micura, R., Dumas, P., and Ennifar, E. (2008) Binding of aminoglycoside antibiotics to the duplex form of the HIV-1 genomic RNA dimerization initiation site. *Angew. Chem., Int. Ed.* 47, 4110–4113.
- (22) Borovinskaya, M. A., Pai, R. D., Zhang, W., Schuwirth, B. S., Holton, J. M., Hirokawa, G., Kaji, H., Kaji, A., and Cate, J. H. (2007) Structural basis for aminoglycoside inhibition of bacterial ribosome recycling. *Nat. Struct. Mol. Biol.* 14, 727–732.
- (23) Hermann, T. (2007) Aminoglycoside antibiotics: old drugs and new therapeutic approaches. *Cell. Mol. Life Sci.* 64, 1841–1852.
- (24) Mingeot-Leclercq, M. P., Glupczynski, Y., and Tulkens, P. M. (1999) Aminoglycosides: activity and resistance. *Antimicrob. Agents Chemother.* 43, 727–737.
- (25) Houghton, J. L., Green, K. D., Chen, W., and Garneau-Tsodikova, S. (2010) The future of aminoglycosides: the end or renaissance? *ChemBioChem* 11, 880–902.
- (26) Luedtke, N. W., Liu, Q., and Tor, Y. (2003) RNA-ligand interactions: affinity and specificity of aminoglycoside dimers and acridine conjugates to the HIV-1 Rev response element. *Biochemistry* 42, 11391–11403.
- (27) Kumar, S., Kellish, P., Robinson, W. E., Jr., Wang, D., Appella, D. H., and Arya, D. P. (2012) Click dimers to target HIV TAR RNA conformation. *Biochemistry* 51, 2331–2347.
- (28) Riguet, E., Desire, J., Boden, O., Ludwig, V., Gobel, M., Bailly, C., and Decout, J. L. (2005) Neamine dimers targeting the HIV-1 TAR RNA. *Bioorg. Med. Chem. Lett.* 15, 4651–4655.
- (29) Bodlener, A., Alix, A., Weibel, J. M., Pale, P., Ennifar, E., Paillart, J. C., Walter, P., Marquet, R., and Dumas, P. (2007) Synthesis of a neamine dimer targeting the dimerization initiation site of HIV-1 RNA. *Org. Lett.* 9, 4415–4418.
- (30) Blount, K. F., and Tor, Y. (2006) A tale of two targets: differential RNA selectivity of nucleobase-aminoglycoside conjugates. *ChemBioChem* 7, 1612–1621.
- (31) Ennifar, E., Walter, P., and Dumas, P. (2003) A crystallographic study of the binding of 13 metal ions to two related RNA duplexes. *Nucleic Acids Res.* 31, 2671–2682.
- (32) Bernacchi, S., Ennifar, E., Toth, K., Walter, P., Langowski, J., and Dumas, P. (2005) Mechanism of hairpin-duplex conversion for the HIV-1 dimerization initiation site. *J. Biol. Chem.* 280, 40112–40121.
- (33) Ennifar, E., Bernacchi, S., Wolff, P., and Dumas, P. (2007) Influence of C-5 halogenation of uridines on hairpin versus duplex RNA folding. *RNA* 13, 1445–1452.
- (34) Velazquez Campoy, A., and Freire, E. (2005) ITC in the post-genomic era...? Priceless. *Biophys Chem* 115, 115–124.
- (35) Burnouf, D., Ennifar, E., Guedich, S., Puffer, B., Hoffmann, G., Bec, G., Disdier, F., Baltzinger, M., and Dumas, P. (2011) kinITC: A new method for obtaining joint thermodynamic and kinetic data by isothermal titration calorimetry. *J. Am. Chem. Soc.* 134, 559–565.
- (36) Jones, T. A. (1978) A graphic model building and refinement system for macromolecules. *J. Appl. Crystallogr.* 11, 268–272.
- (37) Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 54, 905–921.
- (38) Greenberg, W. A., Priestley, E. S., Sears, P. S., Alper, P. B., Rosenbohm, C., Hendrix, M., Hung, S. C., and Wong, C. H. (1999) Design and synthesis of new aminoglycoside antibiotics containing neamine as an optimal core structure: correlation of antibiotic activity with in vitro inhibition of translation. *J. Am. Chem. Soc.* 121, 6527–6541.
- (39) Wu, B., Yang, J., He, Y., and Swayze, E. E. (2002) Reexamination of neomycin B degradation: efficient preparation of its CD and D rings as protected glycosyl donors. *Org. Lett.* 4, 3455–3458.
- (40) van den Broek, S. B., Gruijters, B. W., Rutjes, F. P., van Delft, F. L., and Blaauw, R. H. (2007) A short and scalable route to orthogonally O-protected 2-deoxystreptomine. *J. Org. Chem.* 72, 3577–3580.
- (41) Corey, E. J., Hua, P. H., Pan, B. C., and Seitz, S. P. (1982) Total synthesis of aplasmomycin. *J. Am. Chem. Soc.* 104, 6818–6820.
- (42) Pearlman, W. M. (1967) Noble metal hydroxides on carbon nonpyrophoric dry catalysts. *Tetrahedron Lett.* 8, 1663–1664.
- (43) Faber, C., Sticht, H., Schweimer, K., and Rosch, P. (2000) Structural rearrangements of HIV-1 Tat-responsive RNA upon binding of neomycin B. *J. Biol. Chem.* 275, 20660–20666.
- (44) Barbieri, C. M., and Pilch, D. S. (2006) Complete thermodynamic characterization of the multiple protonation equilibria of the aminoglycoside antibiotic paromomycin: a calorimetric and natural abundance ¹⁵N NMR study. *Biophys. J.* 90, 1338–1349.
- (45) Francois, B., Russell, R. J., Murray, J. B., Aboul-ela, F., Masquida, B., Vicens, Q., and Westhof, E. (2005) Crystal structures of complexes between aminoglycosides and decoding A site oligonucleotides: role of the number of rings and positive charges in the specific binding leading to miscoding. *Nucleic Acids Res.* 33, 5677–5690.
- (46) Han, Q., Zhao, Q., Fish, S., Simonsen, K. B., Vourloumis, D., Froelich, J. M., Wall, D., and Hermann, T. (2005) Molecular recognition by glycoside pseudo base pairs and triples in an apramycin-RNA complex. *Angew. Chem., Int. Ed.* 44, 2694–2700.
- (47) Kaul, M., Barbieri, C. M., and Pilch, D. S. (2004) Fluorescence-based approach for detecting and characterizing antibiotic-induced conformational changes in ribosomal RNA: comparing aminoglycoside binding to prokaryotic and eukaryotic ribosomal RNA sequences. *J. Am. Chem. Soc.* 126, 3447–3453.
- (48) Kaul, M., Barbieri, C. M., and Pilch, D. S. (2006) Aminoglycoside-induced reduction in nucleotide mobility at the ribosomal RNA A-site as a potentially key determinant of antibacterial activity. *J. Am. Chem. Soc.* 128, 1261–1271.
- (49) Pilch, D. S., Kaul, M., Barbieri, C. M., and Kerrigan, J. E. (2003) Thermodynamics of aminoglycoside-rRNA recognition. *Biopolymers* 70, 58–79.
- (50) Shandrick, S., Zhao, Q., Han, Q., Ayida, B. K., Takahashi, M., Winters, G. C., Simonsen, K. B., Vourloumis, D., and Hermann, T. (2004) Monitoring molecular recognition of the ribosomal decoding site. *Angew. Chem., Int. Ed.* 43, 3177–3182.
- (51) Vicens, Q., and Westhof, E. (2002) Crystal structure of a complex between the aminoglycoside tobramycin and an oligonucleotide containing the ribosomal decoding site. *Chem. Biol.* 9, 747–755.
- (52) Vicens, Q., and Westhof, E. (2003) Crystal structure of Geneticin bound to a bacterial 16S ribosomal RNA A site oligonucleotide. *J. Mol. Biol.* 326, 1175–1188.

- (53) Blount, K. F., Zhao, F., Hermann, T., and Tor, Y. (2005) Conformational constraint as a means for understanding RNA-aminoglycoside specificity. *J. Am. Chem. Soc.* *127*, 9818–9829.
- (54) Francois, B., Szychowski, J., Adhikari, S. S., Pachamuthu, K., Swayze, E. E., Griffey, R. H., Migawa, M. T., Westhof, E., and Hanessian, S. (2004) Antibacterial aminoglycosides with a modified mode of binding to the ribosomal-RNA decoding site. *Angew. Chem., Int. Ed.* *43*, 6735–6738.
- (55) Hanessian, S., Szychowski, J., Adhikari, S. S., Vasquez, G., Kandasamy, P., Swayze, E. E., Migawa, M. T., Ranken, R., Francois, B., Wirmer-Bartoschek, J., Kondo, J., and Westhof, E. (2007) Structure-based design, synthesis, and A-site rRNA cocrystal complexes of functionally novel aminoglycoside antibiotics: C2' ether analogues of paromomycin. *J. Med. Chem.* *50*, 2352–2369.
- (56) Simonsen, K. B., Ayida, B. K., Vourloumis, D., Takahashi, M., Winters, G. C., Barluenga, S., Qamar, S., Shandrick, S., Zhao, Q., and Hermann, T. (2002) Novel paromamine derivatives exploring shallow-groove recognition of ribosomal-decoding-site RNA. *ChemBioChem* *3*, 1223–1228.
- (57) Vourloumis, D., Winters, G. C., Simonsen, K. B., Takahashi, M., Ayida, B. K., Shandrick, S., Zhao, Q., Han, Q., and Hermann, T. (2005) Aminoglycoside-hybrid ligands targeting the ribosomal decoding site. *ChemBioChem* *6*, 58–65.
- (58) Zhao, F., Zhao, Q., Blount, K. F., Han, Q., Tor, Y., and Hermann, T. (2005) Molecular recognition of RNA by neomycin and a restricted neomycin derivative. *Angew. Chem., Int. Ed.* *44*, 5329–5334.
- (59) Recht, M. I., Fourmy, D., Blanchard, S. C., Dahlquist, K. D., and Puglisi, J. D. (1996) RNA sequence determinants for aminoglycoside binding to an A-site rRNA model oligonucleotide. *J. Mol. Biol.* *262*, 421–436.
- (60) Zapp, M. L., Stern, S., and Green, M. R. (1993) Small molecules that selectively block RNA binding of HIV-1 Rev protein inhibit Rev function and viral production. *Cell* *74*, 969–978.
- (61) Pokrovskaya, V., Nudelman, I., Kandasamy, J., and Baasov, T. (2010) Aminoglycosides redesign strategies for improved antibiotics and compounds for treatment of human genetic diseases. *Methods Enzymol.* *478*, 437–462.

Bibliographie

Bibliographie

(2004) The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* **306**: 636-640

Abel ED, Peroni O, Kim JK, Kim YB, Boss O, Hadro E, Minnemann T, Shulman GI, Kahn BB (2001) Adipose-selective targeting of the GLUT4 gene impairs insulin action in muscle and liver. *Nature* **409**: 729-733

Aftab S, Semene L, Chu JS, Chen N (2008) Identification and characterization of novel human tissue-specific RFX transcription factors. *BMC Evol Biol* **8**: 226

Ait-Lounis A, Baas D, Barras E, Benadiba C, Charollais A, Nlend Nlend R, Liegeois D, Meda P, Durand B, Reith W (2007) Novel function of the ciliogenic transcription factor RFX3 in development of the endocrine pancreas. *Diabetes* **56**: 950-959

Ait-Lounis A, Bonal C, Seguin-Estevez Q, Schmid CD, Bucher P, Herrera PL, Durand B, Meda P, Reith W (2010) The transcription factor Rfx3 regulates beta-cell differentiation, function, and glucokinase expression. *Diabetes* **59**: 1674-1685

Annicotte JS, Blanchet E, Chavey C, Iankova I, Costes S, Assou S, Teyssier J, Dalle S, Sardet C, Fajas L (2009) The CDK4-pRB-E2F1 pathway controls insulin secretion. *Nat Cell Biol* **11**: 1017-1023

Aramata S, Han SI, Kataoka K (2007) Roles and regulation of transcription factor MafA in islet beta-cells. *Endocr J* **54**: 659-666

Ashcroft FM, Rorsman P (2012) Diabetes mellitus and the beta cell: the last ten years. *Cell* **148**: 1160-1171

Ashique AM, Choe Y, Karlen M, May SR, Phamluong K, Solloway MJ, Ericson J, Peterson AS (2009) The Rfx4 transcription factor modulates Shh signaling by regional control of ciliogenesis. *Sci Signal* **2**: ra70

Bailey TL, Boden M, Whittington T, Machanick P (2010) The value of position-specific priors in motif discovery using MEME. *BMC Bioinformatics* **11**: 179

Bailey TL, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* **2**: 28-36

Bailey TL, Machanick P (2012) Inferring direct DNA binding from ChIP-seq. *Nucleic Acids Res* **40**: e128

Balakrishnan S, Sadasivam M, Kannan A, Panneerselvam A, Prahalthan C (2014) Glucose modulates Pax6 expression through the JNK/p38 MAP kinase pathway in pancreatic beta-cells. *Life Sci* **109**: 1-7

Barton FB, Rickels MR, Alejandro R, Hering BJ, Wease S, Naziruddin B, Oberholzer J, Odorico JS, Garfinkel MR, Levy M, Pattou F, Berney T, Secchi A, Messinger S, Senior PA, Maffi P, Posselt A, Stock PG, Kaufman DB, Luo X, Kandeel F, Cagliero E, Turgeon NA, Witkowski P, Naji A, O'Connell PJ, Greenbaum C, Kudva YC, Brayman KL, Aull MJ, Larsen C, Kay TW, Fernandez LA, Vantyghem MC, Bellin M, Shapiro AM (2012) Improvement in outcomes of clinical islet transplantation: 1999-2010. *Diabetes Care* **35**: 1436-1445

Baur F, Nau K, Sadic D, Allweiss L, Elsasser HP, Gillemans N, de Wit T, Kruger I, Vollmer M, Philipsen S, Suske G (2010) Specificity protein 2 (Sp2) is essential for mouse development and autonomous proliferation of mouse embryonic fibroblasts. *PLoS One* **5**: e9587

Bellin MD, Barton FB, Heitman A, Harmon JV, Kandaswamy R, Balamurugan AN, Sutherland DE, Alejandro R, Hering BJ (2012) Potent induction immunotherapy promotes long-term insulin independence after islet transplantation in type 1 diabetes. *Am J Transplant* **12**: 1576-1583

Benitez CM, Qu K, Sugiyama T, Pauerstein PT, Liu Y, Tsai J, Gu X, Ghodasara A, Arda HE, Zhang J, Dekker JD, Tucker HO, Chang HY, Kim SK (2014) An integrated cell purification and genomics strategy reveals multiple regulators of pancreas development. *PLoS Genet* **10**: e1004645

Beucher A, Martin M, Spenle C, Poulet M, Collin C, Gradwohl G (2012) Competence of failed endocrine progenitors to give rise to acinar but not ductal cells is restricted to early pancreas development. *Dev Biol* **361**: 277-285

Bi HC, Pan YZ, Qiu JX, Krausz KW, Li F, Johnson CH, Jiang CT, Gonzalez FJ, Yu AM (2014) N-methylnicotinamide and nicotinamide N-methyltransferase are associated with microRNA-1291-altered pancreatic carcinoma cell metabolome and suppressed tumorigenesis. *Carcinogenesis* **35**: 2264-2272

Bisgrove BW, Makova S, Yost HJ, Brueckner M (2012) RFX2 is essential in the ciliated organ of asymmetry and an RFX2 transgene identifies a population of ciliated cells sufficient for fluid flow. *Dev Biol* **363**: 166-178

- Biswas M, Chan JY (2010) Role of Nrf1 in antioxidant response element-mediated gene expression and beyond. *Toxicol Appl Pharmacol* **244**: 16-20
- Boergesen M, Poulsen L, Schmidt SF, Frigerio F, Maechler P, Mandrup S (2011) ChREBP mediates glucose repression of peroxisome proliferator-activated receptor alpha expression in pancreatic beta-cells. *J Biol Chem* **286**: 13214-13225
- Bonnafe E, Touka M, AitLounis A, Baas D, Barras E, Ucla C, Moreau A, Flamant F, Dubruille R, Couble P, Collignon J, Durand B, Reith W (2004) The transcription factor RFX3 directs nodal cilium development and left-right asymmetry specification. *Mol Cell Biol* **24**: 4417-4427
- Bonnefond A, Froguel P, Vaxillaire M (2010) The emerging genetics of type 2 diabetes. *Trends Mol Med* **16**: 407-416
- Bosco D, Armanet M, Morel P, Niclauss N, Sgroi A, Muller YD, Giovannoni L, Parnaud G, Berney T (2010) Unique arrangement of alpha- and beta-cells in human islets of Langerhans. *Diabetes* **59**: 1202-1210
- Brennan RG (1993) The winged-helix DNA-binding motif: another helix-turn-helix takeoff. *Cell* **74**: 773-776
- Breslin MB, Wang HW, Pierce A, Aucoin R, Lan MS (2007) Neurogenin 3 recruits CBP co-activator to facilitate histone H3/H4 acetylation in the target gene INSM1. *FEBS Lett* **581**: 949-954
- Burghoorn J, Piasecki BP, Crona F, Phirke P, Jeppsson KE, Swoboda P (2012) The in vivo dissection of direct RFX-target gene promoters in *C. elegans* reveals a novel cis-regulatory element, the C-box. *Dev Biol* **368**: 415-426
- Cairo S, Merla G, Urbinati F, Ballabio A, Reymond A (2001) WBSCR14, a gene mapping to the Williams--Beuren syndrome deleted region, is a new member of the Mlx transcription factor network. *Hum Mol Genet* **10**: 617-627
- Cantley J, Choudhury AI, Asare-Anane H, Selman C, Lingard S, Heffron H, Herrera P, Persaud SJ, Withers DJ (2007) Pancreatic deletion of insulin receptor substrate 2 reduces beta and alpha cell mass and impairs glucose homeostasis in mice. *Diabetologia* **50**: 1248-1256
- Cerf ME (2006) Transcription factors regulating beta-cell function. *Eur J Endocrinol* **155**: 671-679

Cha-Molstad H, Saxena G, Chen J, Shalev A (2009) Glucose-stimulated expression of Txnip is mediated by carbohydrate response element-binding protein, p300, and histone H4 acetylation in pancreatic beta cells. *J Biol Chem* **284**: 16898-16905

Cha JY, Repa JJ (2007) The liver X receptor (LXR) and hepatic lipogenesis. The carbohydrate-response element-binding protein is a target gene of LXR. *J Biol Chem* **282**: 743-751

Chandra R, Liddle RA (2014a) Modulation of pancreatic exocrine and endocrine secretion. *Curr Opin Gastroenterol* **29**: 517-522

Chandra R, Liddle RA (2014b) Recent advances in the regulation of pancreatic secretion. *Curr Opin Gastroenterol* **30**: 490-494

Chandra V, Albagli-Curiel O, Hastoy B, Piccand J, Randriamampita C, Vaillant E, Cave H, Busiah K, Froguel P, Vaxillaire M, Rorsman P, Polak M, Scharfmann R (2014) RFX6 regulates insulin secretion by modulating Ca²⁺ homeostasis in human beta cells. *Cell Rep* **9**: 2206-2218

Chappell L, Gorman S, Campbell F, Ellard S, Rice G, Dobbie A, Crow Y (2008) A further example of a distinctive autosomal recessive syndrome comprising neonatal diabetes mellitus, intestinal atresias and gall bladder agenesis. *Am J Med Genet A* **146A**: 1713-1717

Chen H, Gu X, Liu Y, Wang J, Wirt SE, Bottino R, Schorle H, Sage J, Kim SK (2011) PDGF signalling controls age-dependent proliferation in pancreatic beta-cells. *Nature* **478**: 349-355

Chen J, Hui ST, Couto FM, Mungrue IN, Davis DB, Attie AD, Lusic AJ, Davis RA, Shalev A (2008a) Thioredoxin-interacting protein deficiency induces Akt/Bcl-xL signaling and pancreatic beta-cell mass and protects against diabetes. *FASEB J* **22**: 3581-3594

Chen J, Saxena G, Mungrue IN, Lusic AJ, Shalev A (2008b) Thioredoxin-interacting protein: a critical link between glucose toxicity and beta-cell apoptosis. *Diabetes* **57**: 938-944

Choksi SP, Lauter G, Swoboda P, Roy S (2014) Switching on cilia: transcriptional networks regulating ciliogenesis. *Development* **141**: 1427-1441

Chung MI, Peyrot SM, LeBoeuf S, Park TJ, McGary KL, Marcotte EM, Wallingford JB (2012) RFX2 is broadly required for ciliogenesis during vertebrate development. *Dev Biol* **363**: 155-165

Cleveland MH, Sawyer JM, Afelik S, Jensen J, Leach SD (2012) Exocrine ontogenies: on the development of pancreatic acinar, ductal and centroacinar cells. *Semin Cell Dev Biol* **23**: 711-719

Collombat P, Hecksher-Sorensen J, Krull J, Berger J, Riedel D, Herrera PL, Serup P, 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, Mansouri A (2009) Turning on the beta-cell identity in the pancreas. *Cell Cycle* **8**: 3450-3451

Collombat P, Mansouri A, Hecksher-Sorensen J, Serup P, Krull J, Gradwohl G, Gruss P (2003) Opposing actions of Arx and Pax4 in endocrine pancreas development. *Genes Dev* **17**: 2591-2603

Concepcion JP, Reh CS, Daniels M, Liu X, Paz VP, Ye H, Highland HM, Hanis CL, Greeley SA (2014) Neonatal diabetes, gallbladder agenesis, duodenal atresia, and intestinal malrotation caused by a novel homozygous mutation in RFX6. *Pediatr Diabetes* **15**: 67-72

Courtney M, Gjernes E, Druelle N, Ravaud C, Vieira A, Ben-Othman N, Pfeifer A, Avolio F, Leuckx G, Lacas-Gervais S, Burel-Vandenbos F, Ambrosetti D, Hecksher-Sorensen J, Ravassard P, Heimberg H, Mansouri A, Collombat P (2013) The inactivation of Arx in pancreatic alpha-cells triggers their neogenesis and conversion into functional beta-like cells. *PLoS Genet* **9**: e1003934

D'Amour KA, Bang AG, Eliazer S, Kelly OG, Agulnick AD, Smart NG, Moorman MA, Kroon E, Carpenter MK, Baetge EE (2006) Production of pancreatic hormone-expressing endocrine cells from human embryonic stem cells. *Nat Biotechnol* **24**: 1392-1401

da Silva Xavier G, Rutter GA, Diraison F, Andreolas C, Leclerc I (2006) ChREBP binding to fatty acid synthase and L-type pyruvate kinase genes is stimulated by glucose in pancreatic beta-cells. *J Lipid Res* **47**: 2482-2491

da Silva Xavier G, Sun G, Qian Q, Rutter GA, Leclerc I (2010) ChREBP regulates Pdx-1 and other glucose-sensitive genes in pancreatic beta-cells. *Biochem Biophys Res Commun* **402**: 252-257

David E, Garcia AD, Hearing P (1995) Interaction of EF-C/RFX-1 with the inverted repeat of viral enhancer regions is required for transactivation. *J Biol Chem* **270**: 8353-8360

Davies MN, O'Callaghan BL, Towle HC (2008) Glucose activates ChREBP by increasing its rate of nuclear entry and relieving repression of its transcriptional activity. *J Biol Chem* **283**: 24029-24038

de Luis O, Valero MC, Jurado LA (2000) WBSCR14, a putative transcription factor gene deleted in Williams-Beuren syndrome: complete characterisation of the human gene and the mouse ortholog. *Eur J Hum Genet* **8**: 215-222

Dentin R, Tomas-Cobos L, Foufelle F, Leopold J, Girard J, Postic C, Ferre P (2012) Glucose 6-phosphate, rather than xylulose 5-phosphate, is required for the activation of ChREBP in response to glucose in the liver. *J Hepatol* **56**: 199-209

Dorajoo R, Liu J, Boehm BO (2015) Genetics of Type 2 Diabetes and Clinical Utility. *Genes (Basel)* **6**: 372-384

Drucker DJ (2013) Incretin action in the pancreas: potential promise, possible perils, and pathological pitfalls. *Diabetes* **62**: 3316-3323

Dubruille R, Laurencon A, Vandaele C, Shishido E, Coulon-Bublex M, Swoboda P, Couble P, Kernan M, 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

El Zein L, Ait-Lounis A, Morle L, Thomas J, Chhin B, Spassky N, Reith W, Durand B (2009) RFX3 governs growth and beating efficiency of motile cilia in mouse and controls the expression of genes involved in human ciliopathies. *J Cell Sci* **122**: 3180-3189

Emery P, Durand B, Mach B, Reith W (1996a) RFX proteins, a novel family of DNA binding proteins conserved in the eukaryotic kingdom. *Nucleic Acids Res* **24**: 803-807

Emery P, Strubin M, Hofmann K, Bucher P, Mach B, Reith W (1996b) A consensus motif in the RFX DNA binding domain and binding domain mutants with altered specificity. *Mol Cell Biol* **16**: 4486-4494

Ezer D, Zabet NR, Adryan B (2014) Homotypic clusters of transcription factor binding sites: A model system for understanding the physical mechanics of gene expression. *Comput Struct Biotechnol J* **10**: 63-69

Fajas L, Annicotte JS, Miard S, Sarruf D, Watanabe M, Auwerx J (2004) Impaired pancreatic growth, beta cell mass, and beta cell function in E2F1 (-/-) mice. *J Clin Invest* **113**: 1288-1295

Fajas L, Blanchet E, Annicotte JS (2010) The CDK4-pRB-E2F1 pathway : A new modulator of insulin secretion. *Islets* **2**: 51-53

Fajas L, Landsberg RL, Huss-Garcia Y, Sardet C, Lees JA, Auwerx J (2002) E2Fs regulate adipocyte differentiation. *Dev Cell* **3**: 39-49

Fang Y, Gao J, Qi L, Li N (2014) CTCF-regulating endocrine function of pancreatic islet cells in transgenic mice. *Horm Metab Res* **46**: 419-423

Feng C, Xu W, Zuo Z (2009) Knockout of the regulatory factor X1 gene leads to early embryonic lethality. *Biochem Biophys Res Commun* **386**: 715-717

Feng J, Liu T, Qin B, Zhang Y, Liu XS (2012) Identifying ChIP-seq enrichment using MACS. *Nat Protoc* **7**: 1728-1740

Feng J, Liu T, Zhang Y (2011) Using MACS to identify peaks from ChIP-Seq data. *Curr Protoc Bioinformatics* **Chapter 2**: Unit 2 14

Flannick J, Thorleifsson G, Beer NL, Jacobs SB, Grarup N, Burt NP, Mahajan A, Fuchsberger C, Atzmon G, Benediktsson R, Blangero J, Bowden DW, Brandslund I, Brosnan J, Burslem F, Chambers J, Cho YS, Christensen C, Douglas DA, Duggirala R, Dymek Z, Farjoun Y, Fennell T, Fontanillas P, Forsen T, Gabriel S, Glaser B, Gudbjartsson DF, Hanis C, Hansen T, Hreidarsson AB, Hveem K, Ingelsson E, Isomaa B, Johansson S, Jorgensen T, Jorgensen ME, Kathiresan S, Kong A, Kooner J, Kravic J, Laakso M, Lee JY, Lind L, Lindgren CM, Linneberg A, Masson G, Meitinger T, Mohlke KL, Molven A, Morris AP, Potluri S, Rauramaa R, Ribel-Madsen R, Richard AM, Rolph T, Salomaa V, Segre AV, Skarstrand H, Steinthorsdottir V, Stringham HM, Sulem P, Tai ES, Teo YY, Teslovich T, Thorsteinsdottir U, Trimmer JK, Tuomi T, Tuomilehto J, Vaziri-Sani F, Voight BF, Wilson JG, Boehnke M, McCarthy MI, Njolstad PR, Pedersen O, Groop L, Cox DR, Stefansson K, Altshuler D (2014) Loss-of-function mutations in SLC30A8 protect against type 2 diabetes. *Nat Genet* **46**: 357-363

Fuhlbrigge R, Yip L (2014) Self-antigen expression in the peripheral immune system: roles in self-tolerance and type 1 diabetes pathogenesis. *Curr Diab Rep* **14**: 525

Gao N, Le Lay J, Qin W, Doliba N, Schug J, Fox AJ, Smirnova O, Matschinsky FM, Kaestner KH (2010) Foxa1 and Foxa2 maintain the metabolic and secretory features of the mature beta-cell. *Mol Endocrinol* **24**: 1594-1604

Gao N, LeLay J, Vatamaniuk MZ, Rieck S, Friedman JR, Kaestner KH (2008) Dynamic regulation of Pdx1 enhancers by Foxa1 and Foxa2 is essential for pancreas development. *Genes Dev* **22**: 3435-3448

Geloneze B, Lamounier RN, Coelho OR (2006) Postprandial hyperglycemia: treating its atherogenic potential. *Arq Bras Cardiol* **87**: 660-670

Gierl MS, Karoulias N, Wende H, Strehle M, 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

Giuliante R, Sartini D, Bacchetti T, Rocchetti R, Kloting I, Polidori C, Ferretti G, Emanuelli M (2015) Potential involvement of nicotinamide N-methyltransferase in the pathogenesis of metabolic syndrome. *Metab Syndr Relat Disord* **13**: 165-170

Gradwohl G, Dierich A, LeMeur M, 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

Graham FL, Smiley J, Russell WC, Nairn R (1977) Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J Gen Virol* **36**: 59-74

Grandori C, Cowley SM, James LP, Eisenman RN (2000) The Myc/Max/Mad network and the transcriptional control of cell behavior. *Annu Rev Cell Dev Biol* **16**: 653-699

Grant CE, Bailey TL, Noble WS (2011) FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**: 1017-1018

Grouwels G, Cai Y, Hoebeke I, Leuckx G, Heremans Y, Ziebold U, Stange G, Chintinne M, Ling Z, Pipeleers D, Heimberg H, Van de Casteele M (2010) Ectopic expression of E2F1 stimulates beta-cell proliferation and function. *Diabetes* **59**: 1435-1444

Gu C, Stein GH, Pan N, Goebbels S, Hornberg H, Nave KA, Herrera P, White P, Kaestner KH, Sussel L, Lee JE (2010) Pancreatic beta cells require NeuroD to achieve and maintain functional maturity. *Cell Metab* **11**: 298-310

Gu G, Dubauskaite J, Melton DA (2002) Direct evidence for the pancreatic lineage: NGN3+ cells are islet progenitors and are distinct from duct progenitors. *Development* **129**: 2447-2457

Gunton JE, Kulkarni RN, Yim S, Okada T, Hawthorne WJ, Tseng YH, Roberson RS, Ricordi C, O'Connell PJ, Gonzalez FJ, Kahn CR (2005) Loss of ARNT/HIF1beta mediates altered gene expression and pancreatic-islet dysfunction in human type 2 diabetes. *Cell* **122**: 337-349

Guo XR, Wang XL, Li MC, Yuan YH, Chen Y, Zou DD, Bian LJ, Li DS (2014) PDX-1 mRNA-induced reprogramming of mouse pancreas-derived mesenchymal stem cells into insulin-producing cells in vitro. *Clin Exp Med*

Hanna S, Etzioni A (2014) MHC class I and II deficiencies. *J Allergy Clin Immunol* **134**: 269-275

Hashimoto K, Ishida E, Matsumoto S, Okada S, Yamada M, Satoh T, Monden T, Mori M (2009) Carbohydrate response element binding protein gene expression is positively regulated by thyroid hormone. *Endocrinology* **150**: 3417-3424

Hay CW, Docherty K (2006) Comparative analysis of insulin gene promoters: implications for diabetes research. *Diabetes* **55**: 3201-3213

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**: 576-589

Herman MA, Peroni OD, Villoria J, Schon MR, Abumrad NA, Bluher M, Klein S, Kahn BB (2012) A novel ChREBP isoform in adipose tissue regulates systemic glucose metabolism. *Nature* **484**: 333-338

Hogh KL, Craig MN, Uy CE, Nygren H, Asadi A, Speck M, Fraser JD, Rudecki AP, Baker RK, Oresic M, Gray SL (2014) Overexpression of PPARgamma specifically in pancreatic beta-cells exacerbates obesity-induced glucose intolerance, reduces beta-cell mass, and alters islet lipid metabolism in male mice. *Endocrinology* **155**: 3843-3852

Huang HP, Liu M, El-Hodiri HM, Chu K, Jamrich M, Tsai MJ (2000) Regulation of the pancreatic islet-specific gene BETA2 (neuroD) by neurogenin 3. *Mol Cell Biol* **20**: 3292-3307

Huang Q, Whittington T, Gao P, Lindberg JF, Yang Y, Sun J, Vaisanen MR, Szulkin R, Annala M, Yan J, Egevad LA, Zhang K, Lin R, Jolma A, Nykter M, Manninen A, Wiklund F, Vaarala MH, Visakorpi T, Xu J, Taipale J, Wei GH (2014) A prostate cancer susceptibility allele at 6q22 increases RFX6 expression by modulating HOXB13 chromatin binding. *Nat Genet* **46**: 126-135

Hurtado del Pozo C, Vesperinas-Garcia G, Rubio MA, Corripio-Sanchez R, Torres-Garcia AJ, Obregon MJ, Calvo RM (2011) ChREBP expression in the liver, adipose tissue and differentiated preadipocytes in human obesity. *Biochim Biophys Acta* **1811**: 1194-1200

Ichimura A, Hasegawa S, Kasubuchi M, Kimura I (2014) Free fatty acid receptors as therapeutic targets for the treatment of diabetes. *Front Pharmacol* **5**: 236

Iizuka K, Bruick RK, Liang G, Horton JD, Uyeda K (2004) Deficiency of carbohydrate response element-binding protein (ChREBP) reduces lipogenesis as well as glycolysis. *Proc Natl Acad Sci U S A* **101**: 7281-7286

Iizuka K, Miller B, Uyeda K (2006) Deficiency of carbohydrate-activated transcription factor ChREBP prevents obesity and improves plasma glucose control in leptin-deficient (*ob/ob*) mice. *Am J Physiol Endocrinol Metab* **291**: E358-364

Iizuka K, Wu W, Horikawa Y, Takeda J (2013) Role of glucose-6-phosphate and xylulose-5-phosphate in the regulation of glucose-stimulated gene expression in the pancreatic beta cell line, INS-1E. *Endocr J* **60**: 473-482

Itoh Y, Kawamata Y, Harada M, Kobayashi M, Fujii R, Fukusumi S, Ogi K, Hosoya M, Tanaka Y, Uejima H, Tanaka H, Maruyama M, Satoh R, Okubo S, Kizawa H, Komatsu H, Matsumura F, Noguchi Y, Shinohara T, Hinuma S, Fujisawa Y, Fujino M (2003) Free fatty acids regulate insulin secretion from pancreatic beta cells through GPR40. *Nature* **422**: 173-176

Jay MA, Ren J (2007) Peroxisome proliferator-activated receptor (PPAR) in metabolic syndrome and type 2 diabetes mellitus. *Curr Diabetes Rev* **3**: 33-39

Jeong YS, Kim D, Lee YS, Kim HJ, Han JY, Im SS, Chong HK, Kwon JK, Cho YH, Kim WK, Osborne TF, Horton JD, Jun HS, Ahn YH, Ahn SM, Cha JY (2011) Integrated expression profiling and genome-wide analysis of ChREBP targets reveals the dual role for ChREBP in glucose-regulated gene expression. *PLoS One* **6**: e22544

Jia S, Ivanov A, Blasevic D, Muller T, Purfurst B, Sun W, Chen W, Poy MN, Rajewsky N, Birchmeier C (2015) *Insm1* cooperates with *Neurod1* and *Foxa2* to maintain mature pancreatic beta-cell function. *EMBO J* **34**: 1417-1433

Jonsson J, Carlsson L, Edlund T, Edlund H (1994) Insulin-promoter-factor 1 is required for pancreas development in mice. *Nature* **371**: 606-609

Jorgensen MC, Ahnfelt-Ronne J, Hald J, Madsen OD, Serup P, Hecksher-Sorensen J (2007) An illustrated review of early pancreas development in the mouse. *Endocr Rev* **28**: 685-705

Kabashima T, Kawaguchi T, Wadzinski BE, Uyeda K (2003) Xylulose 5-phosphate mediates glucose-induced lipogenesis by xylulose 5-phosphate-activated protein phosphatase in rat liver. *Proc Natl Acad Sci U S A* **100**: 5107-5112

Katan-Khaykovich Y, Shaul Y (1998) RFX1, a single DNA-binding protein with a split dimerization domain, generates alternative complexes. *J Biol Chem* **273**: 24504-24512

Kawaguchi T, Osatomi K, Yamashita H, Kabashima T, Uyeda K (2002a) Mechanism for fatty acid "sparing" effect on glucose-induced transcription: regulation of carbohydrate-responsive element-binding protein by AMP-activated protein kinase. *J Biol Chem* **277**: 3829-3835

Kawaguchi T, Takenoshita M, Kabashima T, Uyeda K (2001) Glucose and cAMP regulate the L-type pyruvate kinase gene by phosphorylation/dephosphorylation of the carbohydrate response element binding protein. *Proc Natl Acad Sci U S A* **98**: 13710-13715

Kawaguchi Y, Cooper B, Gannon M, Ray M, MacDonald RJ, Wright CV (2002b) The role of the transcriptional regulator Ptf1a in converting intestinal to pancreatic progenitors. *Nat Genet* **32**: 128-134

Kibbe C, Chen J, Xu G, Jing G, Shalev A (2013) FOXO1 competes with carbohydrate response element-binding protein (ChREBP) and inhibits thioredoxin-interacting protein (TXNIP) transcription in pancreatic beta cells. *J Biol Chem* **288**: 23194-23202

Kidder BL, Hu G, Zhao K (2011) ChIP-Seq: technical considerations for obtaining high-quality data. *Nat Immunol* **12**: 918-922

Kim JH, Lee SR, Li LH, Park HJ, Park JH, Lee KY, Kim MK, Shin BA, Choi SY (2011) High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS One* **6**: e18556

Kim S, Yu NK, Kaang BK (2015) CTCF as a multifunctional protein in genome regulation and gene expression. *Exp Mol Med* **47**: e166

Kim SY, Rane SG (2011) The Cdk4-E2f1 pathway regulates early pancreas development by targeting Pdx1+ progenitors and Ngn3+ endocrine precursors. *Development* **138**: 1903-1912

Kovac J, Husse J, Oster H (2009) A time to fast, a time to feast: the crosstalk between metabolism and the circadian clock. *Mol Cells* **28**: 75-80

Krapp A, Knofler M, Ledermann B, Burki K, Berney C, Zoerkler N, Hagenbuchle O, Wellauer PK (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

Krebs A, Frontini M, Tora L (2008) GPAT: retrieval of genomic annotation from large genomic position datasets. *BMC Bioinformatics* **9**: 533

Kroon E, Martinson LA, Kadoya K, Bang AG, Kelly OG, Eliazer S, Young H, Richardson M, Smart NG, Cunningham J, Agulnick AD, D'Amour KA, Carpenter MK, Baetge EE (2008) Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells in vivo. *Nat Biotechnol* **26**: 443-452

Kulkarni RN, Mizrahi EB, Ocana AG, Stewart AF (2012) Human beta-cell proliferation and intracellular signaling: driving in the dark without a road map. *Diabetes* **61**: 2205-2213

Kurahashi T, Furusawa T, Ueda T, Bustin M (2010) The nucleosome binding protein HMGN3 is expressed in pancreatic alpha-cells and affects plasma glucagon levels in mice. *J Cell Biochem* **109**: 49-57

Lan MS, Breslin MB (2009) Structure, expression, and biological function of INSM1 transcription factor in neuroendocrine differentiation. *FASEB J* **23**: 2024-2033

Langmead B (2010) Aligning short sequencing reads with Bowtie. *Curr Protoc Bioinformatics* **Chapter 11**: Unit 11 17

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25

Laurencon A, Dubruille R, Efimenko E, Grenier G, Bissett R, Cortier E, Rolland V, Swoboda P, Durand B (2007) Identification of novel regulatory factor X (RFX) target genes by comparative genomics in *Drosophila* species. *Genome Biol* **8**: R195

Lee EC, Yu D, Martinez de Velasco J, Tessarollo L, Swing DA, Court DL, Jenkins NA, Copeland NG (2001) A highly efficient *Escherichia coli*-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. *Genomics* **73**: 56-65

Lenting PJ, Casari C, Christophe OD, Denis CV (2012) von Willebrand factor: the old, the new and the unknown. *J Thromb Haemost* **10**: 2428-2437

Lenting PJ, Christophe OD, Denis CV (2015) von Willebrand factor biosynthesis, secretion, and clearance: connecting the far ends. *Blood* **125**: 2019-2028

Li MV, Chang B, Imamura M, Pongvarin N, Chan L (2006a) Glucose-dependent transcriptional regulation by an evolutionarily conserved glucose-sensing module. *Diabetes* **55**: 1179-1189

Li MV, Chen W, Harmancey RN, Nuotio-Antar AM, Imamura M, Saha P, Taegtmeyer H, Chan L (2010) Glucose-6-phosphate mediates activation of the carbohydrate responsive binding protein (ChREBP). *Biochem Biophys Res Commun* **395**: 395-400

Li YY, Mukaida N (2014) Pathophysiological roles of Pim-3 kinase in pancreatic cancer development and progression. *World J Gastroenterol* **20**: 9392-9404

Li YY, Popivanova BK, Nagai Y, Ishikura H, Fujii C, Mukaida N (2006b) Pim-3, a proto-oncogene with serine/threonine kinase activity, is aberrantly expressed in human pancreatic cancer and phosphorylates bad to block bad-mediated apoptosis in human pancreatic cancer cell lines. *Cancer Res* **66**: 6741-6747

Liang H, Xiao G, Yin H, Hippenmeyer S, Horowitz JM, Ghashghaei HT (2013) Neural development is dependent on the function of specificity protein 2 in cell cycle progression. *Development* **140**: 552-561

Liao X, Xue H, Wang YC, Nazor KL, Guo S, Trivedi N, Peterson SE, Liu Y, Loring JF, Laurent LC (2013) Matched miRNA and mRNA signatures from an hESC-based in vitro model of pancreatic differentiation reveal novel regulatory interactions. *J Cell Sci* **126**: 3848-3861

Lilla V, Webb G, Rickenbach K, Maturana A, Steiner DF, Halban PA, Irminger JC (2003) Differential gene expression in well-regulated and dysregulated pancreatic beta-cell (MIN6) sublines. *Endocrinology* **144**: 1368-1379

Lubelsky Y, Reuven N, 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

Ma L, Robinson LN, Towle HC (2006) ChREBP**Mlx* is the principal mediator of glucose-induced gene expression in the liver. *J Biol Chem* **281**: 28721-28730

Ma W, Noble WS, Bailey TL (2014) Motif-based analysis of large nucleotide data sets using MEME-ChIP. *Nat Protoc* **9**: 1428-1450

Machanick P, Bailey TL (2011) MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics* **27**: 1696-1697

Madsen OD, Serup P (2006) Towards cell therapy for diabetes. *Nat Biotechnol* **24**: 1481-1483

Marquez-Aguirre AL, Canales-Aguirre AA, Padilla-Camberos E, Esquivel-Solis H, Diaz-Martinez NE (2015) Development of the endocrine pancreas and novel strategies for beta-cell mass restoration and diabetes therapy. *Braz J Med Biol Res*: 0

Martinez AI, Perez-Arellano I, Pekkala S, Barcelona B, Cervera J (2010) Genetic, structural and biochemical basis of carbamoyl phosphate synthetase 1 deficiency. *Mol Genet Metab* **101**: 311-323

Martinovici D, Ransy V, Vanden Eijnden S, Ridremont C, Pardou A, Cassart M, Avni F, Donner C, Lingier P, Mathieu A, Gulbis B, De Brouckere V, Cnop M, Abramowicz M, Desir J (2010) Neonatal hemochromatosis and Martinez-Frias syndrome of intestinal atresia and diabetes mellitus in a consanguineous newborn. *Eur J Med Genet* **53**: 25-28

Matsubara K, Aoyama K, Suno M, Awaya T (2002) N-methylation underlying Parkinson's disease. *Neurotoxicol Teratol* **24**: 593-598

McGrath PS, Watson CL, Ingram C, Helmrath MA, Wells JM (2015) The basic helix-loop-helix transcription factor NEUROG3 is required for development of the human endocrine pancreas. *Diabetes*

Mellitzer G, Bonne S, Luco RF, Van De Castele M, Lenne-Samuel N, Collombat P, Mansouri A, Lee J, Lan M, Pipeleers D, Nielsen FC, Ferrer J, Gradwohl G, 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 PA, Sander M, Gradwohl G (2004) Pancreatic islet progenitor cells in neurogenin 3-yellow fluorescent protein knock-add-on mice. *Mol Endocrinol* **18**: 2765-2776

Melloul D, Marshak S, Cerasi E (2002) Regulation of insulin gene transcription. *Diabetologia* **45**: 309-326

Merla G, Howald C, Antonarakis SE, Reymond A (2004) The subcellular localization of the ChoRE-binding protein, encoded by the Williams-Beuren syndrome critical region gene 14, is regulated by 14-3-3. *Hum Mol Genet* **13**: 1505-1514

Metukuri MR, Zhang P, Basantani MK, Chin C, Stamateris RE, Alonso LC, Takane KK, Gramignoli R, Strom SC, O'Doherty RM, Stewart AF, Vasavada RC, Garcia-Ocana A, Scott DK (2012) ChREBP mediates glucose-stimulated pancreatic beta-cell proliferation. *Diabetes* **61**: 2004-2015

Minn AH, Hafele C, Shalev A (2005a) Thioredoxin-interacting protein is stimulated by glucose through a carbohydrate response element and induces beta-cell apoptosis. *Endocrinology* **146**: 2397-2405

Minn AH, Pise-Masison CA, Radonovich M, Brady JN, Wang P, Kendziorski C, Shalev A (2005b) Gene expression profiling in INS-1 cells overexpressing thioredoxin-interacting protein. *Biochem Biophys Res Commun* **336**: 770-778

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

Miyazaki J, Araki K, Yamato E, Ikegami H, Asano T, Shibasaki Y, Oka Y, Yamamura K (1990) Establishment of a pancreatic beta cell line that retains glucose-inducible insulin secretion: special reference to expression of glucose transporter isoforms. *Endocrinology* **127**: 126-132

Molven A, Njolstad PR (2011) Role of molecular genetics in transforming diagnosis of diabetes mellitus. *Expert Rev Mol Diagn* **11**: 313-320

Muller G, Ruppert S, Schmid E, Schutz G (1988) Functional analysis of alternatively spliced tyrosinase gene transcripts. *EMBO J* **7**: 2723-2730

Murphy KG, Bloom SR (2006) Gut hormones and the regulation of energy homeostasis. *Nature* **444**: 854-859

Murtaugh LC (2007) Pancreas and beta-cell development: from the actual to the possible. *Development* **134**: 427-438

Nekrep N, Jabrane-Ferrat N, Wolf HM, Eibl MM, Geyer M, Peterlin BM (2002) Mutation in a winged-helix DNA-binding motif causes atypical bare lymphocyte syndrome. *Nat Immunol* **3**: 1075-1081

Noordeen NA, Khera TK, Sun G, Longbottom ER, Pullen TJ, da Silva Xavier G, Rutter GA, Leclerc I (2010) Carbohydrate-responsive element-binding protein (ChREBP) is a negative regulator of ARNT/HIF-1 β gene expression in pancreatic islet beta-cells. *Diabetes* **59**: 153-160

Noordeen NA, Meur G, Rutter GA, Leclerc I (2012) Glucose-induced nuclear shuttling of ChREBP is mediated by sorcin and Ca²⁺ ions in pancreatic beta-cells. *Diabetes* **61**: 574-585

Okita K, Mizuguchi T, Shigenori O, Ishii M, Nishidate T, Ueki T, Meguro M, Kimura Y, Tanimizu N, Ichinohe N, Torigoe T, Kojima T, Mitaka T, Sato N, Sawada N, Hirata K (2015) Pancreatic regeneration: basic research and gene regulation. *Surg Today*

Oropez D, Horb M (2012) Transient expression of Ngn3 in *Xenopus* endoderm promotes early and ectopic development of pancreatic beta and delta cells. *Genesis* **50**: 271-285

Osipovich AB, Long Q, Manduchi E, Gangula R, Hipkens SB, Schneider J, Okubo T, Stoeckert CJ, Jr., Takada S, Magnuson MA (2014) Insm1 promotes endocrine cell differentiation by modulating the expression of a network of genes that includes Neurog3 and Ripply3. *Development* **141**: 2939-2949

Pagliuca FW, Melton DA (2013) How to make a functional beta-cell. *Development* **140**: 2472-2483

Pagliuca FW, Millman JR, Gurtler M, Segel M, Van Dervort A, Ryu JH, Peterson QP, Greiner D, Melton DA (2014) Generation of Functional Human Pancreatic beta Cells In Vitro. *Cell* **159**: 428-439

Pan FC, Wright C (2011) Pancreas organogenesis: from bud to plexus to gland. *Dev Dyn* **240**: 530-565

Patti ME (2004) Gene expression in humans with diabetes and prediabetes: what have we learned about diabetes pathophysiology? *Curr Opin Clin Nutr Metab Care* **7**: 383-390

Pearl EJ, Jarikji Z, Horb ME (2011) Functional analysis of Rfx6 and mutant variants associated with neonatal diabetes. *Dev Biol* **351**: 135-145

Perilhou A, Turrel-Cuzin C, Kharroubi I, Henique C, Fauveau V, Kitamura T, Magnan C, Postic C, Prip-Buus C, Vasseur-Cognet M (2008a) The transcription

factor COUP-TFII is negatively regulated by insulin and glucose via Foxo1- and ChREBP-controlled pathways. *Mol Cell Biol* **28**: 6568-6579

Perilhou A, Turrel-Cuzin C, Zhang P, Kharroubi I, Wang H, Fauveau V, Scott DK, Wollheim CB, Vasseur-Cognet M (2008b) The MODY1 gene for hepatocyte nuclear factor 4alpha and a feedback loop control COUP-TFII expression in pancreatic beta cells. *Mol Cell Biol* **28**: 4588-4597

Piccand J, Strasser P, Hodson DJ, Meunier A, Ye T, Keime C, Birling MC, Rutter GA, Gradwohl G (2014) Rfx6 Maintains the Functional Identity of Adult Pancreatic beta Cells. *Cell Rep* **9**: 2219-2232

Pober BR (2010) Williams-Beuren syndrome. *N Engl J Med* **362**: 239-252

Pober BR, Wang E, Caprio S, Petersen KF, Brandt C, Stanley T, Osborne LR, Dzuria J, Gulanski B (2010) High prevalence of diabetes and pre-diabetes in adults with Williams syndrome. *Am J Med Genet C Semin Med Genet* **154C**: 291-298

Popivanova BK, Li YY, Zheng H, Omura K, Fujii C, Tsuneyama K, Mukaida N (2007) Proto-oncogene, Pim-3 with serine/threonine kinase activity, is aberrantly expressed in human colon cancer cells and can prevent Bad-mediated apoptosis. *Cancer Sci* **98**: 321-328

Postic C, Dentin R, Denechaud PD, Girard J (2007) ChREBP, a transcriptional regulator of glucose and lipid metabolism. *Annu Rev Nutr* **27**: 179-192

Poungvarin N, Lee JK, Yechoor VK, Li MV, Assavapokee T, Suksaranjit P, Thepsongwajja JJ, Saha PK, Oka K, Chan L (2012) Carbohydrate response element-binding protein (ChREBP) plays a pivotal role in beta cell glucotoxicity. *Diabetologia* **55**: 1783-1796

Pullen TJ, Khan AM, Barton G, Butcher SA, Sun G, Rutter GA (2010) Identification of genes selectively disallowed in the pancreatic islet. *Islets* **2**: 89-95

Quintens R, Hendrickx N, Lemaire K, Schuit F (2008) Why expression of some genes is disallowed in beta-cells. *Biochem Soc Trans* **36**: 300-305

Reith W, Herrero-Sanchez C, Kobr M, Silacci P, Berte C, Barras E, Fey S, 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, Mach B (2001) The bare lymphocyte syndrome and the regulation of MHC expression. *Annu Rev Immunol* **19**: 331-373

Reith W, Ucla C, Barras E, Gaud A, Durand B, Herrero-Sanchez C, Kobr M, 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

Rezania A, Bruin JE, Riedel MJ, Mojibian M, Asadi A, Xu J, Gauvin R, Narayan K, Karanu F, O'Neil JJ, Ao Z, Warnock GL, Kieffer TJ (2012) Maturation of human embryonic stem cell-derived pancreatic progenitors into functional islets capable of treating pre-existing diabetes in mice. *Diabetes* **61**: 2016-2029

Rutter GA, Chimienti F (2015) SLC30A8 mutations in type 2 diabetes. *Diabetologia* **58**: 31-36

Sander M, Neubuser A, Kalamaras J, Ee HC, Martin GR, German MS (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, 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

Sansbury FH, Kirel B, Caswell R, Allen Lango H, Flanagan SE, Hattersley AT, Ellard S, Shaw-Smith CJ (2015) Biallelic RFX6 mutations can cause childhood as well as neonatal onset diabetes mellitus. *Eur J Hum Genet*

Satoh J, Kawana N, Yamamoto Y (2013) Pathway Analysis of ChIP-Seq-Based NRF1 Target Genes Suggests a Logical Hypothesis of their Involvement in the Pathogenesis of Neurodegenerative Diseases. *Gene Regul Syst Bio* **7**: 139-152

Schaffer AE, Taylor BL, Benthuisen JR, Liu J, Thorel F, Yuan W, Jiao Y, Kaestner KH, Herrera PL, Magnuson MA, May CL, Sander M (2013) Nkx6.1 controls a gene regulatory network required for establishing and maintaining pancreatic Beta cell identity. *PLoS Genet* **9**: e1003274

Selman C, Lingard S, Choudhury AI, Batterham RL, Claret M, Clements M, Ramadani F, Okkenhaug K, Schuster E, Blanc E, Piper MD, Al-Qassab H, Speakman JR, Carmignac D, Robinson IC, Thornton JM, Gems D, Partridge L, Withers DJ (2008) Evidence for lifespan extension and delayed age-related biomarkers in insulin receptor substrate 1 null mice. *FASEB J* **22**: 807-818

Shaner NC, Campbell RE, Steinbach PA, Giepmans BN, Palmer AE, Tsien RY (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat Biotechnol* **22**: 1567-1572

Shapiro AM, Lakey JR, Ryan EA, Korbutt GS, Toth E, Warnock GL, Kneteman NM, Rajotte RV (2000) Islet transplantation in seven patients with type 1 diabetes mellitus using a glucocorticoid-free immunosuppressive regimen. *N Engl J Med* **343**: 230-238

Shawlot W, Vazquez-Chantada M, Wallingford JB, Finnell RH (2015) Rfx2 is required for spermatogenesis in the mouse. *Genesis*

Shih HM, Liu Z, Towle HC (1995) Two CACGTG motifs with proper spacing dictate the carbohydrate regulation of hepatic gene transcription. *J Biol Chem* **270**: 21991-21997

Sirek AS, Liu L, Naples M, Adeli K, Ng DS, Jin T (2009) Insulin stimulates the expression of carbohydrate response element binding protein (ChREBP) by attenuating the repressive effect of Pit-1, Oct-1/Oct-2, and Unc-86 homeodomain protein octamer transcription factor-1. *Endocrinology* **150**: 3483-3492

Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S, Balkau B, Heude B, Charpentier G, Hudson TJ, Montpetit A, Pshzhetsky AV, Prentki M, Posner BI, Balding DJ, Meyre D, Polychronakos C, Froguel P (2007) A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* **445**: 881-885

Smith SB, Qu HQ, Taleb N, Kishimoto NY, Scheel DW, Lu Y, Patch AM, Grabs R, Wang J, Lynn FC, Miyatsuka T, Mitchell J, Seerke R, Desir J, Vanden Eijnden S, Abramowicz M, Kacet N, Weill J, Renard ME, Gentile M, Hansen I, Dewar K, Hattersley AT, Wang R, Wilson ME, Johnson JD, Polychronakos C, German MS (2010) Rfx6 directs islet formation and insulin production in mice and humans. *Nature* **463**: 775-780

Soggia A, Flosseau K, Ravassard P, Szinnai G, Scharfmann R, Guillemain G (2012) Activation of the transcription factor carbohydrate-responsive element-binding protein by glucose leads to increased pancreatic beta cell differentiation in rats. *Diabetologia* **55**: 2713-2722

Sosa-Pineda B, Chowdhury K, Torres M, Oliver G, Gruss P (1997) The Pax4 gene is essential for differentiation of insulin-producing beta cells in the mammalian pancreas. *Nature* **386**: 399-402

Soyer J, Flasse L, Raffelsberger W, Beucher A, Orvain C, Peers B, Ravassard P, Vermot J, Voz ML, Mellitzer G, Gradwohl G (2010) Rfx6 is an Ngn3-dependent winged helix transcription factor required for pancreatic islet cell development. *Development* **137**: 203-212

Spiegel R, Dobbie A, Hartman C, de Vries L, Ellard S, Shalev SA (2011) Clinical characterization of a newly described neonatal diabetes syndrome caused by RFX6 mutations. *Am J Med Genet A* **155A**: 2821-2825

Steneberg P, Rubins N, Bartoov-Shifman R, Walker MD, Edlund H (2005) The FFA receptor GPR40 links hyperinsulinemia, hepatic steatosis, and impaired glucose homeostasis in mouse. *Cell Metab* **1**: 245-258

Stephen L. Aronoff KB, Barb Shreiner and Laura Want (2004) Glucose Metabolism and Regulation: Beyond Insulin and Glucagon. *Diabetes Spectrum* **17**: 8

Stoeckman AK, Ma L, Towle HC (2004) Mlx is the functional heteromeric partner of the carbohydrate response element-binding protein in glucose regulation of lipogenic enzyme genes. *J Biol Chem* **279**: 15662-15669

Strub T, Giuliano S, Ye T, Bonet C, Keime C, Kobi D, Le Gras S, Cormont M, Ballotti R, Bertolotto C, Davidson I (2011) Essential role of microphthalmia transcription factor for DNA replication, mitosis and genomic stability in melanoma. *Oncogene* **30**: 2319-2332

Sund NJ, Vatamaniuk MZ, Casey M, Ang SL, Magnuson MA, Stoffers DA, Matschinsky FM, Kaestner KH (2001) Tissue-specific deletion of Foxa2 in pancreatic beta cells results in hyperinsulinemic hypoglycemia. *Genes Dev* **15**: 1706-1715

Sussel L, Kalamaras J, Hartigan-O'Connor DJ, Meneses JJ, Pedersen RA, Rubenstein JL, German MS (1998) Mice lacking the homeodomain transcription factor Nkx2.2 have diabetes due to arrested differentiation of pancreatic beta cells. *Development* **125**: 2213-2221

Suzuki K, Harada N, Yamane S, Nakamura Y, Sasaki K, Nasteska D, Joo E, Shibue K, Harada T, Hamasaki A, Toyoda K, Nagashima K, Inagaki N (2012) Transcriptional factor regulatory factor X 6 (Rfx6) increases gastric inhibitory polypeptide (GIP) expression in enteroendocrine K-cells and is involved in GIP hypersecretion in high-fat diet-induced obesity. *J Biol Chem*

Swoboda P, Adler HT, Thomas JH (2000) The RFX-type transcription factor DAF-19 regulates sensory neuron cilium formation in *C. elegans*. *Mol Cell* **5**: 411-421

Tanaka H, Yoshida S, Minoura H, Negoro K, Shimaya A, Shimokawa T, Shibasaki M (2014) Novel GPR40 agonist AS2575959 exhibits glucose metabolism improvement and synergistic effect with sitagliptin on insulin and incretin secretion. *Life Sci* **94**: 115-121

Taylor BL, Liu FF, Sander M (2013) Nkx6.1 is essential for maintaining the functional state of pancreatic beta cells. *Cell Rep* **4**: 1262-1275

Thomas J, Morle L, Soulavie F, Laurencon A, Sagnol S, Durand B (2010) Transcriptional control of genes involved in ciliogenesis: a first step in making cilia. *Biol Cell* **102**: 499-513

Thompson N, Gesina E, Scheinert P, Bucher P, Grapin-Botton A (2012) RNA profiling and chromatin immunoprecipitation-sequencing reveal that PTF1a stabilizes pancreas progenitor identity via the control of MNX1/HLXB9 and a network of other transcription factors. *Mol Cell Biol* **32**: 1189-1199

Thorrez L, Laudadio I, Van Deun K, Quintens R, Hendrickx N, Granvik M, Lemaire K, Schraenen A, Van Lommel L, Lehnert S, Aguayo-Mazzucato C, Cheng-Xue R, Gilon P, Van Mechelen I, Bonner-Weir S, Lemaigre F, Schuit F (2011) Tissue-specific disallowance of housekeeping genes: the other face of cell differentiation. *Genome Res* **21**: 95-105

Thumser AE, Moore JB, Plant NJ (2014) Fatty acid binding proteins: tissue-specific functions in health and disease. *Curr Opin Clin Nutr Metab Care* **17**: 124-129

Tong X, Zhao F, Mancuso A, Gruber JJ, Thompson CB (2009a) The glucose-responsive transcription factor ChREBP contributes to glucose-dependent anabolic synthesis and cell proliferation. *Proc Natl Acad Sci U S A* **106**: 21660-21665

Tong X, Zhao F, Mancuso A, Gruber JJ, Thompson CB. (2009b) The glucose-responsive transcription factor ChREBP contributes to glucose-dependent anabolic synthesis and cell proliferation. *Proceedings of the National Academy of Sciences*, Vol. 106, pp. 21660-21665.

Tsatsos NG, Towle HC (2006) Glucose activation of ChREBP in hepatocytes occurs via a two-step mechanism. *Biochem Biophys Res Commun* **340**: 449-456

Ueda T, Furusawa T, Kurahashi T, Tessarollo L, Bustin M (2009) The nucleosome binding protein HMG3 modulates the transcription profile of pancreatic beta cells and affects insulin secretion. *Mol Cell Biol* **29**: 5264-5276

Villasenor A, Chong DC, Henkemeyer M, Cleaver O (2010) Epithelial dynamics of pancreatic branching morphogenesis. *Development* **137**: 4295-4305

Vizcaino C, Mansilla S, Portugal J (2015) Sp1 transcription factor: A long-standing target in cancer chemotherapy. *Pharmacol Ther* **152**: 111-124

Vlacich G, Nawijn MC, Webb GC, Steiner DF (2010) Pim3 negatively regulates glucose-stimulated insulin secretion. *Islets* **2**: 308-317

Wang H, Wollheim CB (2002) ChREBP rather than USF2 regulates glucose stimulation of endogenous L-pyruvate kinase expression in insulin-secreting cells. *J Biol Chem* **277**: 32746-32752

Wang J, Elghazi L, Parker SE, Kizilocak H, Asano M, Sussel L, Sosa-Pineda B (2004) The concerted activities of Pax4 and Nkx2.2 are essential to initiate pancreatic beta-cell differentiation. *Dev Biol* **266**: 178-189

Wang S, Hecksher-Sorensen J, Xu Y, Zhao A, Dor Y, Rosenberg L, Serup P, Gu G (2008) Myt1 and Ngn3 form a feed-forward expression loop to promote endocrine islet cell differentiation. *Dev Biol* **317**: 531-540

Wang S, Jensen JN, Seymour PA, Hsu W, Dor Y, Sander M, Magnuson MA, Serup P, Gu G (2009) 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

Wang S, Zhang J, Zhao A, Hipkens S, Magnuson MA, Gu G (2007) Loss of Myt1 function partially compromises endocrine islet cell differentiation and pancreatic physiological function in the mouse. *Mech Dev* **124**: 898-910

Wen JH, Chen YY, Song SJ, Ding J, Gao Y, Hu QK, Feng RP, Liu YZ, Ren GC, Zhang CY, Hong TP, Gao X, Li LS (2009) Paired box 6 (PAX6) regulates glucose metabolism via proinsulin processing mediated by prohormone convertase 1/3 (PC1/3). *Diabetologia* **52**: 504-513

Wierup N, Sundler F, Heller RS (2013) The islet ghrelin cell. *J Mol Endocrinol* **52**: R35-49

Wolf G, Hessabi B, Karkour A, Henrion U, Dahlhaus M, Ostmann A, Giese B, Fraunholz M, Grabarczyk P, Jack R, Walther R (2010) The activation of the rat insulin gene II by BETA2 and PDX-1 in rat insulinoma cells is repressed by Pax6. *Mol Endocrinol* **24**: 2331-2342

Xie J, Yin H, Nichols TD, Yoder JA, Horowitz JM (2010) Sp2 is a maternally inherited transcription factor required for embryonic development. *J Biol Chem* **285**: 4153-4164

Xie W, Wood AR, Lyssenko V, Weedon MN, Knowles JW, Alkayyali S, Assimes TL, Quertermous T, Abbasi F, Paananen J, Haring H, Hansen T, Pedersen O, Smith U, Laakso M, Dekker JM, Nolan JJ, Groop L, Ferrannini E, Adam KP, Gall WE, Frayling TM, Walker M (2013) Genetic variants associated with glycine metabolism and their role in insulin sensitivity and type 2 diabetes. *Diabetes* **62**: 2141-2150

Xu J, Zhang T, Wang T, You L, Zhao Y (2014) PIM kinases: an overview in tumors and recent advances in pancreatic cancer. *Future Oncol* **10**: 865-876

Yabe D, Seino Y (2011) Two incretin hormones GLP-1 and GIP: comparison of their actions in insulin secretion and beta cell preservation. *Prog Biophys Mol Biol* **107**: 248-256

Yamashita H, Takenoshita M, Sakurai M, Bruick RK, Henzel WJ, Shillinglaw W, Arnot D, Uyeda K (2001) A glucose-responsive transcription factor that regulates carbohydrate metabolism in the liver. *Proc Natl Acad Sci U S A* **98**: 9116-9121

Yamato E, Tashiro F, Miyazaki J (2013) Microarray analysis of novel candidate genes responsible for glucose-stimulated insulin secretion in mouse pancreatic beta cell line MIN6. *PLoS One* **8**: e61211

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

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**: R137

Zhao L, Guo M, Matsuoka TA, Hagman DK, Parazzoli SD, Poitout V, Stein R (2005) The islet beta cell-enriched MafA activator is a key regulator of insulin gene transcription. *J Biol Chem* **280**: 11887-11894

Zhao Y, Zhang Y, Zhou M, Wang S, Hua Z, Zhang J (2012) Loss of mPer2 increases plasma insulin levels by enhanced glucose-stimulated insulin secretion and impaired insulin clearance in mice. *FEBS Lett* **586**: 1306-1311

Zhao YC, Zhu J, Song GY, Li XS (2015) Relationship between thioredoxin-interacting protein (TXNIP) and islet beta-cell dysfunction in patients with impaired glucose tolerance and hypertriglyceridemia. *Int J Clin Exp Med* **8**: 4363-4368

Zhou Q, Law AC, Rajagopal J, Anderson WJ, Gray PA, Melton DA (2007) A multipotent progenitor domain guides pancreatic organogenesis. *Dev Cell* **13**: 103-114

Zhu XD, Pan G, Luetke K, Sadowski PD (1995) Homology requirements for ligation and strand exchange by the FLP recombinase. *J Biol Chem* **270**: 11646-11653

Résumé

La régulation de l'homéostasie du glucose dans l'organisme est la fonction principale des cellules beta sécrétrices d'insuline dans le pancréas endocrine. Le facteur de transcription à domaine « winged helix », RFX6, a récemment, été identifié comme un nouveau régulateur de la différenciation endocrine pancréatique en aval de Ngn3 chez le poisson zèbre, la souris et l'homme. De plus, diverses mutations de Rfx6 chez l'homme ont été identifiées et reliées au syndrome de Mitchell Riley notamment caractérisé par un diabète néonatal, une atrésie de l'intestin grêle et une malabsorption intestinale. Lors de mes travaux de thèse, une approche combinée de transcriptomique chez la souris et la recherche des sites de fixation de RFX6 dans une lignée cellulaire beta et dans les ilots pancréatiques a permis de démontrer son importance dans le maintien de l'identité et de la fonction de la cellule beta. Pour la première fois, l'identification des cibles directes de RFX6 *in vivo* a été réalisée et a permis l'identification de l'ensemble du répertoire des gènes régulés directement par RFX6 dans les cellules beta qui n'ont pas été révélés dans le système cellulaire. Cette étude aura également permis d'identifier Mlxipl comme principale cible directement régulée par Rfx6 à la fois chez la souris et l'homme. Les expériences réalisées ont ainsi permis de déterminer les gènes cibles directs de RFX6 et contribué à élucider le rôle de ce facteur de transcription dans la différenciation et la fonction des cellules beta sécrétrices d'insuline.

Summary

Glucose homeostasis regulation in the body is the main function of insulin secreting beta cells in the endocrine pancreas. The winged-helix transcription factor RFX6 has recently been identified as a new pancreatic endocrine differentiation regulator, downstream of Ngn3, in zebrafish, mouse and human. Moreover, several Rfx6 mutations in humans were discovered and linked to the Mitchell Riley syndrome, which is characterized by neonatal diabetes, intestinal atresia and malabsorption. My thesis consisted of using an approach combining transcriptomic analysis in mouse and the identification of RFX6 target genes in a beta cell line as well as in pancreatic islets. This work has demonstrated the crucial role of RFX6 in maintaining beta cell identity and function. For the first time, RFX6 target genes were identified *in vivo* as well as the whole repertoire of directly regulated RFX6 target genes in beta cells, which were previously unidentified in the beta cell line. These studies have also shown that Mlxipl is a main RFX6 regulated target gene in mice and human. Overall, this work has allowed the clear identification of RFX6 target genes, thus contributing in understanding the role of this crucial transcription factor in the differentiation and function of insulin secreting beta cells.