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

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

Josselin Soyer[1,*], Lydie Flasse[2,*], Wolfgang Raffelsberger[1], Anthony Beucher[1], Christophe Orvain[1], Bernard Peers[2], Philippe Ravassard[3], Julien Vermot[1], Marianne L. Voz[2,†], Georg Mellitzer[1,†] and Gérard Gradwohl[1,†]

SUMMARY

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

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

INTRODUCTION

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

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

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

*These authors contributed equally to this work
†Authors for correspondence (mellit@igbmc.fr; gradwohl@igbmc.fr; mvoz@ulg.ac.be)

Accepted 12 November 2009

(Mellitzer et al., 2006) or *Nkx2.2* (Watada et al., 2003) might be direct targets of Ngn3, it has not yet been proven that Ngn3 binds the promoter of these genes in vivo. Our knowledge of the Ngn3-regulated program is thus still incomplete, and we do not know much about how this transcription factor integrates the generic program of endocrine differentiation with the programs that specify the different islet cell types. Therefore, to determine the gene expression profile of islet progenitor cells and thus identify potential novel downstream effectors of Ngn3 function, we performed Affymetrix microarray analysis on purified Ngn3-positive progenitor cells. Here we show that this strategy led to the identification of Rfx6, a novel Ngn3-dependent winged helix transcription factor. We report that in both mouse and zebrafish, *Rfx6* is found in islet progenitor cells and maintained in developing and mature islet cells. Furthermore functional studies in zebrafish revealed that *rfx6* is an important regulator of endocrine cell differentiation. These findings suggest that the study of other genes identified in this study might reveal the full genetic program implementing Ngn3 endocrinogenic function. This information might in turn be relevant to promote and scrutinize the differentiation of hESCs to the beta lineage.

MATERIALS AND METHODS

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

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

In situ hybridization and immunohistochemistry on mouse tissues

Tissues were fixed in paraformaldehyde, embedded and frozen using standard methods. Detailed protocols for in situ hybridization immunofluorescence and immunohistochemistry on frozen sections are available on request. Mouse cRNA probes used included: *Ngn3* and *NeuroD* (Gradwohl et al., 2000), *Insm1*/*IA1* (Mellitzer et al., 2006), *Rfx3* (transcribed from a 2.7 kb cDNA fragment, image clone 4483833, IRAV35-E2) and *Rfx6* (transcribed from a 0.9 kb cDNA fragment, cloned from E13.5 pancreatic RNA with oligo 5′ CGGAATTCGCCACGTGGAGACATCCTAT and 3′GGACTAGTAATCTGGGTTTGCAAGTTGG). The following primary antibodies were used: guinea pig or rabbit anti-Pdx1 at 1:1000 (provided by C. Wright, Vanderbilt University, Nashville, TN, USA), guinea pig anti-Ngn3 at 1:1000 (provided by M. Sander, California University, Irvine, CA, USA), anti-insulin at 1:1000 (Linco), anti-glucagon at 1:2000 (Linco), anti-polypeptide pancreatic (PP) at 1:1000 (Linco), rabbit anti-somatostatin at 1:200 (Dako), rabbit anti-Rfx6 at 1:1000, rat anti-Rfx6 at 1:200, mouse anti-Ki-67 at 1:100 (Novocastra). Secondary antibodies used were: anti-rabbit Alexa 488 at 1:1000 (Molecular Probes), Cy3 anti-rabbit, anti-rat and anti-guinea pig at 1:1000 (Jackson Immunoresearch), biotin-coupled anti-rabbit at 1:200 (Vector Laboratories). For rat anti-Rfx6, signal amplification was performed using biotin anti-rat coupled antibody at 1:200 (Vector Laboratories) and streptavidin-Alexa 488 conjugate at 1:500 (Moleculer Probes). Nuclei were stained with DAPI and slides mounted in Aqua-Poly/Mount (Polysciences).

Cloning of the *rfx6* ortholog from zebrafish

rfx6 partial cDNA was cloned by two rounds of PCR performed on cDNAs of embryos at 24 hours post-fertilization (hpf). The primers used for amplification were O146 (TGCCCTTTTTGACCAGATTGTAGTG) and O139 (GAACGACTGGAGCTGCTGATGGAT) for the first PCR, followed by a nested PCR with O146 and O147 (GCTACGCTTTCTCTGGAC-ATCACCT), giving rise to a 972 bp fragment in the coding region. This fragment was cloned into a pCRII-TOPO vector (Invitrogen) and used as template for preparing labelled antisense RNA probes.

Morpholino sequences and injections

The *rfx6* morpholinos (MOs) were designed by Gene Tools and are complementary to either the exon 2 splice donor site (MO1: GTCCTCAAGCCTAATGAAACAAAAC) or the exon 2 splice acceptor site (MO2: AATAAAAACGCCTCTTACCTTTCCG). A standard control MO, having the sequence 5′-CCTCTTACCTCAGTTACAATTTATA-3′, has also been designed by Gene Tools in a way that it should have no target and no significant biological activity. The MOs were dissolved at a concentration of 3 μg/μl in 1× Danieau buffer containing 0.5% rhodamine dextran and microinjected at the 1- to 2-cell stage at a dose of 3 ng. Injected embryos were then grown in the presence of 0.003% 1-phenyl-2-thiourea until the desired stage, fixed overnight in 4% paraformaldehyde and stored in 100% methanol before analysis.

Riboprobes, wholemount in situ hybridizations (WISH) and imaging

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

Cilia imaging

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

RESULTS

Differential expression of transcription factors in the islet lineage

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

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

Rfx6 is a novel islet-specific winged helix transcription factor

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

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

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

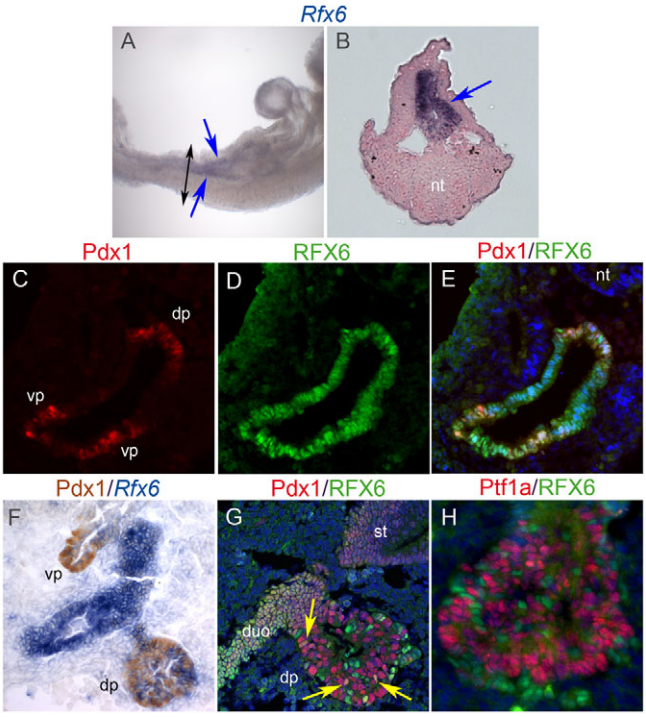


Fig. 1. Expression of *Rfx6* in the gut endoderm and in the developing pancreatic buds in the mouse embryo.
(**A**) Wholemount in situ hybridization (ISH) on an E9.0 mouse embryo (lateral view) with an *Rfx6* antisense probe (purple). (**B**) Transversal section of embryo shown in A at the level of the black arrows. (**C-E**,**G**,**H**) Double immunofluorescence for Rfx6 (green) and Pdx1 (red) on cryosections of E9.5 (C-E), and E10.5 (G) embryos and Rfx6 and Ptf1a (red) at E10.5 (H). (**F**) ISH for *Rfx6* followed by immunohistochemistry for Pdx1. Nuclei are stained with DAPI (blue) in E, G and H. *Rfx6* is found in the gut endoderm (blue arrows in A,B) and in the developing pancreatic regions in Pdx1-positive progenitor cells at E9.5 (C-E). At E10.5, Rfx6 is excluded from Pdx1- (F,G) and Ptf1a- (H) positive pancreatic endodermal progenitor cells and restricted to differentiating endocrine cells. All sections are sagittal, B is transverse. In E and G arrows point to Pdx1/Rfx6 double-positive cells. dp, dorsal pancreas; duo, duodenum; nt, neural tube; st, stomach; vp, ventral pancreas.

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

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

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

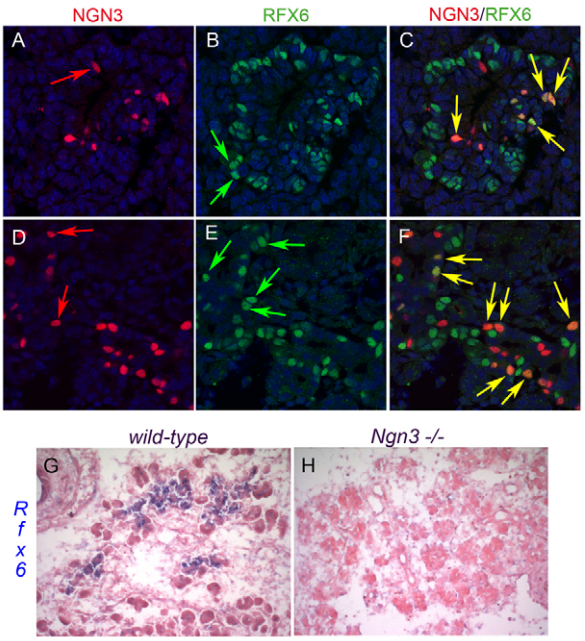


Fig. 2. *Ngn3*-dependent expression of *Rfx6* in the mouse islet lineage. (**A-F**) Double immunofluorescence showing partial overlapping expression of Ngn3 (red) and Rfx6 (green) at E10.5 (A-C) and E15.5 (D-F). Yellow, green and red arrows point to double-labelled, single Rfx6-positive and single Ngn3-positive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

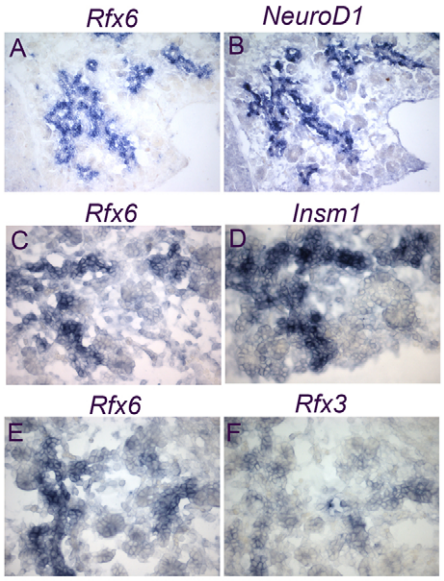


Fig. 3. Overlapping expression of *Rfx6* and other islet transcription factors in the pancreas of mouse embryos. (**A-F**) In situ hybridization (blue) experiments on adjacent cryosections showing overlapping expression of *Rfx6* with *NeuroD*, *Insm1* and *Rfx3*. All experiments were performed on E15.5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

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

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

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

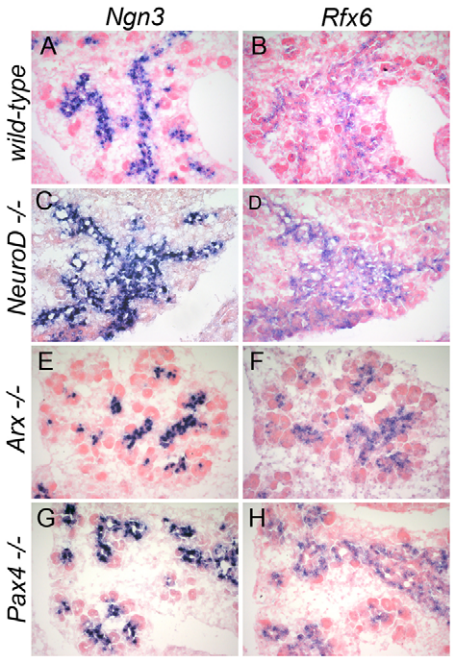


Fig. 4. *Rfx6* expression is unaffected in the pancreas of *Pax4-Arx-* and *NeuroD*-deficient mouse embryos. (**A-H**) In situ hybridization (blue) for *Rfx6* and *Ngn3* on adjacent pancreas cryosections of wild-type (A,B), *NeuroD*- (C,D), Arx- (E,F) and *Pax4*- (G,H) deficient embryos at E14.5.

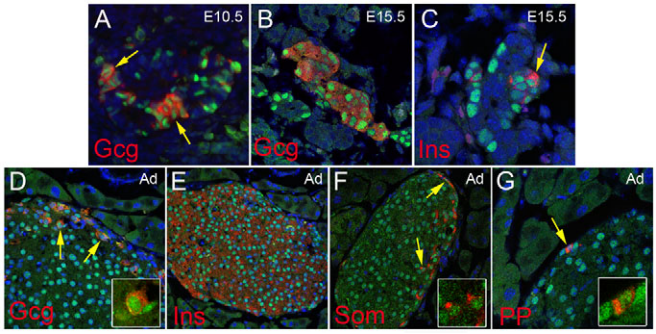


Fig. 5. Rfx6 expression in mature islet cells in the mouse. (**A-G**) Double immunofluorescence revealing nuclear expression of Rfx6 (green) in hormone-expressing cells (red) in developing alpha (A,B) and beta (C) cells, as well as in mature alpha (D), beta (E), delta (F) and PP (G) cells in adult islets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

of endocrine cells (Mavropoulos et al., 2005) and for which we could confirm an absence of colocalization with the hormones at 24 hpf (Fig. 6F). Two other pancreatic markers, *isl1* and *neurod*, were used in the double in situ hybridization experiments. *Isl1* is known to be expressed in postmitotic endocrine cells in mouse embryos (Ahlgren et al., 1997; Thor et al., 1991). The same seems to be true in zebrafish as, at 24 hpf, all mature hormone-expressing cells also express *isl1* (Fig. 6E) and there is essentially no colocalization between *isl1* and the endocrine marker *sox4b* (Fig. 6C). As for *NeuroD*, it is expressed in the murine *Ngn3* pancreatic precursors as well as in the mature hormone-expressing cells (Huang et al., 2000; Itkin-Ansari et al., 2005; Naya et al., 1997). In the same way, in zebrafish, *neurod* is expressed in the pancreatic progenitor cells as demonstrated by its colocalization with the *sox4b* factor (Mavropoulos et al., 2005), and in the mature hormone-expressing cells (Fig. 6D). It is important to note that the pancreatic endocrine precursors, labelled by *sox4b*, are localised in the ventral part of the pancreatic endoderm, whereas the more differentiated cells are located more dorsally (Fig. 6C and diagram).

The location of *rfx6* transcripts was then compared with that of these three pancreatic endocrine markers, *sox4b*, *neurod* and *isl1* (Fig. 6G-N). At stage 18S (18 hpf), *rfx6* showed a perfect colocalization with *sox4b* (Fig. 6G). Progressively, *sox4b* expression became restricted to the ventral part of the *rfx6* expression domain (Fig. 6H,I), the dorsal part corresponding to *isl1*-expressing cells (Fig. 6M,N). Finally, a total colocalization between *rfx6* and *neurod* was observed at all stages analysed (ie. 18, 24 and 30 hpf) (Fig. 6J-L). As expected based on *isl1* and *neurod* colocalization, we detected *rfx6* transcripts in all endocrine cell types, ie. the *insulin*-, *glucagon*-, *ghrelin*- and *somatostatin*-expressing cells (see Fig. S4 in the supplementary material). Taken together, these data show that the pancreatic expression pattern of *rfx6* is similar in zebrafish and in mice, as *rfx6* is expressed in the pancreatic endocrine progenitor cells as well as in the mature endocrine cells. By contrast, divergence exists for its expression in the gut.

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

To assess the role of zebrafish Rfx6 in pancreatic endocrine development, we abrogated Rfx6 protein expression in the zebrafish embryo by injecting two distinct antisense MOs. The first MO targets the exon 2 splice donor site and the second targets the exon 2 splice acceptor site, leading to *rfx6* mRNA splicing disruption as shown by RT-PCR (see Fig. S5 in the supplementary material). As expected based on the restricted expression of *rfx6* in the pancreas, the injection of the MOs did not disturb the general morphology of the embryos (data not shown). By contrast, pancreas development was strongly perturbed as injection of either one or the other MO led to the same phenotype: an almost complete depletion of *glucagon*- and *ghrelin*-expressing cells together with a drastic reduction in the number of *somatostatin*-expressing cells (Fig. 7A-L). Conversely, the number of *insulin*-expressing cells was only slightly decreased (Fig. 7, right), although they were no longer clustered in a compact islet (Fig. 7K,L) To understand by which mechanism Rfx6 affects endocrine cell differentiation, we analysed the expression of *sox4b*, *neurod* and *isl1* genes in the morphants. The number of pancreatic endocrine progenitor cells labelled by *sox4b* was drastically increased at 24 hpf upon *rfx6* knock-down (Fig. 7P-R), whereas the number of more differentiated endocrine cells, labelled by *isl1*, was significantly reduced (Fig. 7M-O). The lateral view of the developing pancreas at 30 hpf, stained for both *sox4b* and *isl1*, highlights the drastic reduction of the *isl1* dorsal domain together with the ventral expansion of *sox4b* (Fig. 7 Y-ZII). This increase does not seem to be due to an enhanced proliferation of the progenitors as we could not detect any change in the proportion of *pcna* (proliferating cell nuclear antigen)/*sox4b* double-positive cells (data not shown). On the contrary, the number of *neurod*-expressing cells was not affected in the morphants (Fig. 7S-U), indicating that the total number of endocrine progenitor and differentiated cells was not significantly perturbed. This strongly suggests that the increase of progenitor cells occurs at the expense of more differentiated cells

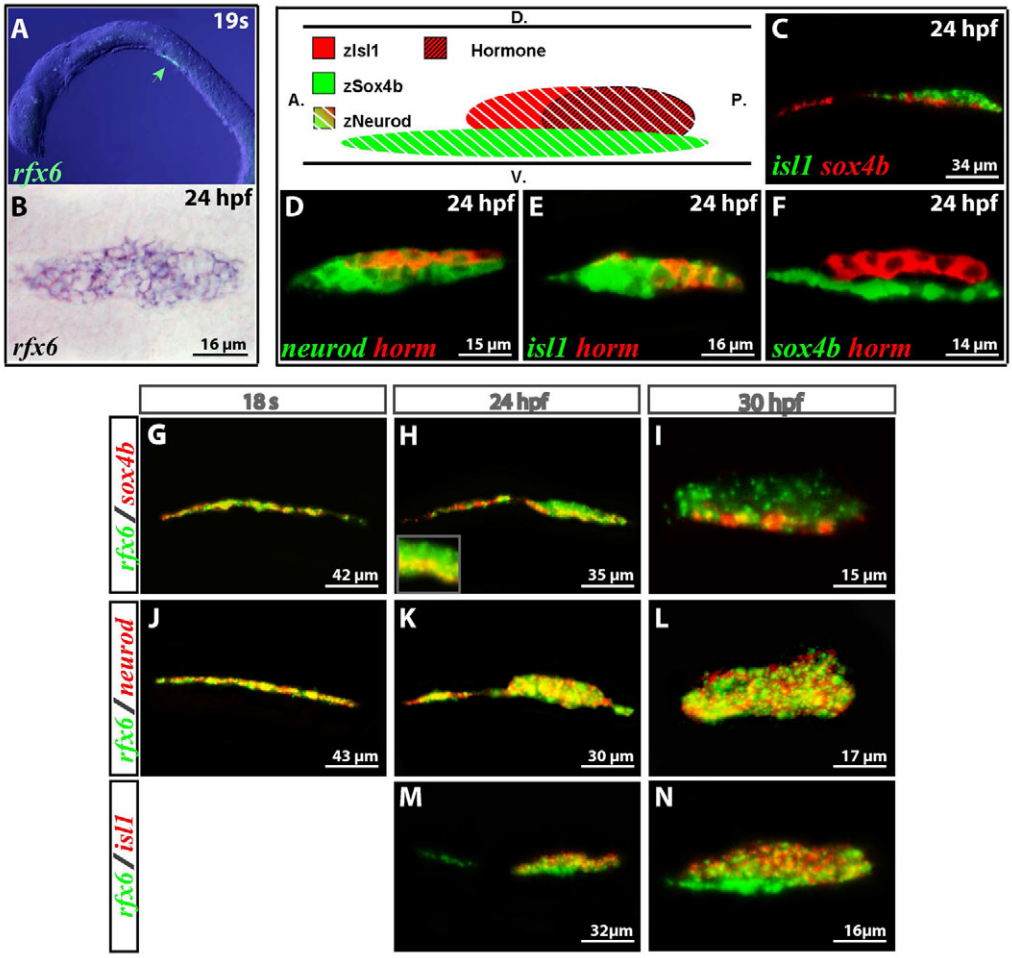


Fig. 6. Zebrafish *rfx6* is expressed in pancreatic endocrine progenitors as well as in mature endocrine cells. (**A**) Lateral view of fluorescent wholemount in situ hybridization (WISH) performed on stage 19S embryos with an *rfx6* antisense probe (arrowhead). (**B**) Ventral view of WISH with an *rfx6*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

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

DISCUSSION

To identify novel downstream effectors of the proendocrine function of the transcription factor Ngn3, we determined the gene expression profile of isolated Ngn3-positive progenitors. Our data are complementary to, and further extend, similar published studies (Gu et al., 2004; White et al., 2008) owing to the combination of an original mouse model with unaltered levels of *Ngn3* transcripts and the use of a very representative Affymetrix array. This is illustrated by the identification of the uncharacterized winged helix transcription factor Rfx6 which has not been found in the above mentioned studies, but was reported recently to be expressed in Ngn3 progenitor cells based on RT-PCR experiments (Miyatsuka et al., 2009). In the present study, we characterize *Rfx6* expression in the mouse and its ortholog in zebrafish, and report its crucial role in the progression of islet cell differentiation in the latter.

Rfx transcription factors in the mouse pancreas

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

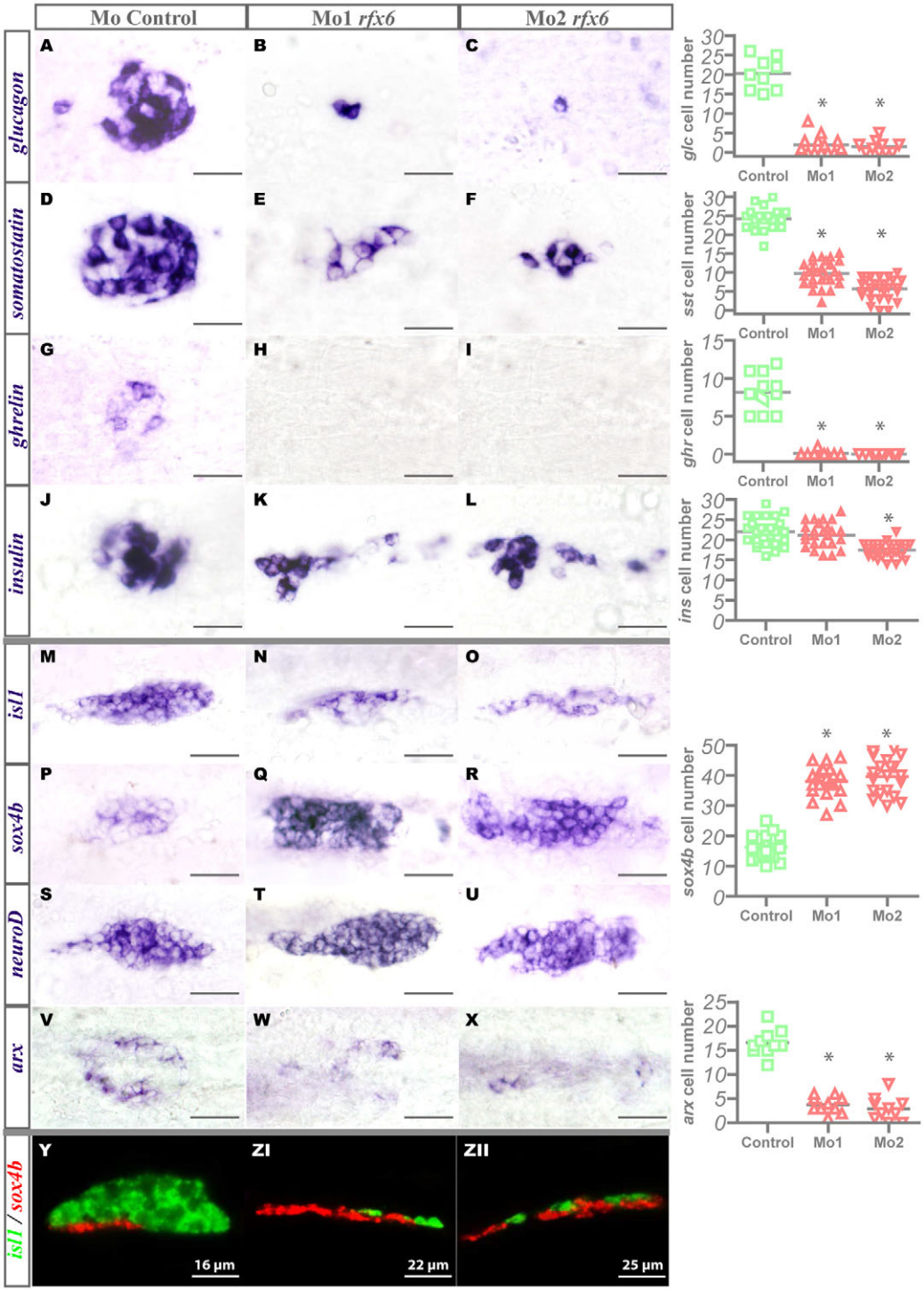


Fig. 7. Impaired endocrine cell differentiation, accumulation of islet progenitor cells and failure of *insulin*-expressing cells clustering in *rfx6* morphants in the zebrafish. (**A-X**) Ventral views of the pancreas area from embryos analysed by WISH, anterior to the left. (A-L) Hormone expression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

function of Rfx proteins arose from studies in invertebrates. daf-19, the unique Rfx gene in *C. elegans*, has been shown to be a crucial regulator of ciliogenesis (Swoboda et al., 2000). Similarly, Rfx in *Drosophila* is necessary for ciliated sensory neuron differentiation (Dubruille et al., 2002). In invertebrates, Rfx transcription factors are thought to control the transcription of proteins involved in cilia assembly in a process called intraflagellar transport (IFT). Cilia are organelles found almost ubiquitously in vertebrate cells and are involved in numerous developmental process and human genetic disorders (ciliopathies) (Gerdes et al., 2009). Recently, *rfx2* has been shown to control ciliogenesis in the zebrafish pronephros (Liu et al., 2007). In the mouse, the expression of *Rfx3* was reported in the endocrine pancreas and it has been shown that *Rfx3* loss-of-function resulted in impaired islet cell composition and glucose tolerance (Ait-Lounis et al., 2007). This phenotype was associated with abnormal formation of primary cilia on islet cells. However, because the pancreatic deletion of *Kif3a*, a gene involved in IFT, did not result in endocrine failure (Cano et al., 2006), it is not clear whether the islet phenotype in *Rfx3*-deficient mice results from defects in cilia formation. In this study, we report for the first time the pancreatic

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

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

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

Role of *rfx6* in zebrafish endocrine cell differentiation

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

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

Acknowledgements

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

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

Competing interests statement

The authors declare no competing financial interests.

Supplementary material

Supplementary material for this article is available at
http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.041673/-/DC1

References

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

Ahlgren, U., Pfaff, S. L., Jessell, T. M., Edlund, T. and Edlund, H. (1997). Independent requirement for ISL1 in formation of pancreatic mesenchyme and islet cells. *Nature* **385**, 257-260.

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

Argenton, F., Zecchin, E. and Bortolussi, M. (1999). Early appearance of pancreatic hormone-expressing cells in the zebrafish embryo. *Mech. Dev.* **87**, 217-221.

Biemar, F., Argenton, F., Schmidtke, R., Epperlein, S., Peers, B. and Driever, W. (2001). Pancreas development in zebrafish: early dispersed appearance of endocrine hormone expressing cells and their convergence to form the definitive islet. *Dev. Biol.* **230**, 189-203.

Cano, D. A., Sekine, S. and Hebrok, M. (2006). Primary cilia deletion in pancreatic epithelial cells results in cyst formation and pancreatitis. *Gastroenterology* **131**, 1856-1869.

Claiborn, K. C. and Stoffers, D. A. (2008). Toward a cell-based cure for diabetes: advances in production and transplant of beta cells. *Mt. Sinai J. Med.* **75**, 362-371.

Collombat, P., Mansouri, A., Hecksher-Sorensen, J., Serup, P., Krull, J., Gradwohl, G. and Gruss, P. (2003). Opposing actions of Arx and Pax4 in endocrine pancreas development. *Genes Dev.* **17**, 2591-2603.

D'Amour, K. A., Bang, A. G., Eliazer, S., Kelly, O. G., Agulnick, A. D., Smart, N. G., Moorman, M. A., Kroon, E., Carpenter, M. K. and Baetge, E. E. (2006). Production of pancreatic hormone-expressing endocrine cells from human embryonic stem cells. *Nat. Biotechnol.* **24**, 1392-1401.

Devos, N., Deflorian, G., Biemar, F., Bortolussi, M., Martial, J. A., Peers, B. and Argenton, F. (2002). Differential expression of two somatostatin genes during zebrafish embryonic development. *Mech. Dev.* **115**, 133-137.

Dubruille, R., Laurencon, A., Vandaele, C., Shishido, E., Coulon-Bublex, M., Swoboda, P., Couble, P., Kernan, M. and Durand, B. (2002). Drosophila regulatory factor X is necessary for ciliated sensory neuron differentiation. *Development* **129**, 5487-5498.

Ekker, S. C. (2000). Morphants: a new systematic vertebrate functional genomics approach. *Yeast* **17**, 302-306.

Gerdes, J. M., Davis, E. E. and Katsanis, N. (2009). The vertebrate primary cilium in development, homeostasis, and disease. *Cell* **137**, 32-45.

Gradwohl, G., Dierich, A., LeMeur, M. and Guillemot, F. (2000). neurogenin3 is required for the development of the four endocrine cell lineages of the pancreas. *Proc. Natl. Acad. Sci. USA* **97**, 1607-1611.

Gu, G., Dubauskaite, J. and Melton, D. A. (2002). Direct evidence for the pancreatic lineage: NGN3+ cells are islet progenitors and are distinct from duct progenitors. *Development* **129**, 2447-2457.

Gu, G., Wells, J. M., Dombkowski, D., Preffer, F., Aronow, B. and Melton, D. A. (2004). Global expression analysis of gene regulatory pathways during endocrine pancreatic development. *Development* **131**, 165-179.

Heller, R. S., Jenny, M., Collombat, P., Mansouri, A., Tomasetto, C., Madsen, O. D., Mellitzer, G., Gradwohl, G. and Serup, P. (2005). Genetic determinants of pancreatic epsilon-cell development. *Dev. Biol.* **286**, 217-224.

Huang, H. P., Liu, M., El Hodiri, H. M., Chu, K., Jamrich, M. and Tsai, M. J. (2000). Regulation of the pancreatic islet-specific gene BETA2 (neuroD) by neurogenin 3. *Mol. Cell. Biol.* **20**, 3292-3307.

Itkin-Ansari, P., Marcora, E., Geron, I., Tyrberg, B., Demeterco, C., Hao, E., Padilla, C., Ratineau, C., Leiter, A., Lee, J. E. et al. (2005). NeuroD1 in the endocrine pancreas: localization and dual function as an activator and repressor. *Dev. Dyn.* **233**, 946-953.

Jenny, M., Uhl, C., Roche, C., Duluc, I., Guillermin, V., Guillemot, F., Jensen, J., Kedinger, M. and Gradwohl, G. (2002). Neurogenin3 is differentially required for endocrine cell fate specification in the intestinal and gastric epithelium. *EMBO J.* **21**, 6338-6347.

Johansson, K. A., Dursun, U., Jordan, N., Gu, G., Beermann, F., Gradwohl, G. and Grapin-Botton, A. (2007). Temporal control of neurogenin3 activity in pancreas progenitors reveals competence windows for the generation of different endocrine cell types. *Dev. Cell* **12**, 457-465.

Korzh, V., Edlund, T. and Thor, S. (1993). Zebrafish primary neurons initiate expression of the LIM homeodomain protein Isl-1 at the end of gastrulation. *Development* **118**, 417-425.

Korzh, V., Sleptsova, I., Liao, J., He, J. and Gong, Z. (1998). Expression of zebrafish bHLH genes ngn1 and nrd defines distinct stages of neural differentiation. *Dev. Dyn.* **213**, 92-104.

Kroon, E., Martinson, L. A., Kadoya, K., Bang, A. G., Kelly, O. G., Eliazer, S., Young, H., Richardson, M., Smart, N. G., Cunningham, J. et al. (2008). Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells in vivo. *Nat. Biotechnol.* **26**, 443-452.

Lee, C. S., Perreault, N., Brestelli, J. E. and Kaestner, K. H. (2002). Neurogenin 3 is essential for the proper specification of gastric enteroendocrine cells and the maintenance of gastric epithelial cell identity. *Genes Dev.* **16**, 1488-1497.

Liu, Y., Pathak, N., Kramer-Zucker, A. and Drummond, I. A. (2007). Notch signaling controls the differentiation of transporting epithelia and multiciliated cells in the zebrafish pronephros. *Development* **134**, 1111-1122.

Madsen, O. D. and Serup, P. (2006). Towards cell therapy for diabetes. *Nat. Biotechnol.* **24**, 1481-1483.

Mavropoulos, A., Devos, N., Biemar, F., Zecchin, E., Argenton, F., Edlund, H., Motte, P., Martial, J. A. and Peers, B. (2005). sox4b is a key player of pancreatic alpha cell differentiation in zebrafish. *Dev. Biol.* **285**, 211-223.

Mellitzer, G., Martin, M., Sidhoum-Jenny, M., Orvain, C., Barths, J., Seymour, P. A., Sander, M. and Gradwohl, G. (2004). Pancreatic islet progenitor cells in neurogenin 3-yellow fluorescent protein knock-add-on mice. *Mol. Endocrinol.* **18**, 2765-2776.

Mellitzer, G., Bonne, S., Luco, R. F., Van de Casteele, M., Lenne-Samuel, N., Collombat, P., Mansouri, A., Lee, J., Lan, M., Pipeleers, D. et al. (2006). IA1 is NGN3-dependent and essential for differentiation of the endocrine pancreas. *EMBO J.* **25**, 1344-1352.

Milewski, W. M., Duguay, S. J., Chan, S. J. and Steiner, D. F. (1998). Conservation of PDX-1 structure, function, and expression in zebrafish. *Endocrinology* **139**, 1440-1449.

Miyatsuka, T., Li, Z. and German, M. S. (2009). Chronology of islet differentiation revealed by temporal cell labeling. *Diabetes* **58**, 1863-1868.

Murtaugh, L. C. (2007). Pancreas and beta-cell development: from the actual to the possible. *Development* **134**, 427-438.

Naya, F. J., Huang, H. P., Qiu, Y., Mutoh, H., DeMayo, F. J., Leiter, A. B. and Tsai, M. J. (1997). Diabetes, defective pancreatic morphogenesis, and abnormal enteroendocrine differentiation in BETA2/neuroD-deficient mice. *Genes Dev.* **11**, 2323-2334.

Pauls, S., Zecchin, E., Tiso, N., Bortolussi, M. and Argenton, F. (2007). Function and regulation of zebrafish nkx2.2a during development of pancreatic islet and ducts. *Dev. Biol.* **304**, 875-890.

Ploner, A., Calza, S., Gusnanto, A. and Pawitan, Y. (2006). Multidimensional local false discovery rate for microarray studies. *Bioinformatics* **22**, 556-565.

Schonhoff, S. E., Giel-Moloney, M. and Leiter, A. B. (2004). Neurogenin 3-expressing progenitor cells in the gastrointestinal tract differentiate into both endocrine and non-endocrine cell types. *Dev. Biol.* **270**, 443-454.

Schwitzgebel, V. M., Scheel, D. W., Conners, J. R., Kalamaras, J., Lee, J. E., Anderson, D. J., Sussel, L., Johnson, J. D. and German, M. S. (2000). Expression of neurogenin3 reveals an islet cell precursor population in the pancreas. *Development* **127**, 3533-3542.

Shapiro, A. M., Lakey, J. R., Ryan, E. A., Korbutt, G. S., Toth, E., Warnock, G. L., Kneteman, N. M. and Rajotte, R. V. (2000). Islet transplantation in seven patients with type 1 diabetes mellitus using a glucocorticoid-free immunosuppressive regimen. *New Engl. J. Med.* **343**, 230-238.

Smith, S. B., Gasa, R., Watada, H., Wang, J., Griffen, S. C. and German, M. S. (2003). Neurogenin3 and hepatic nuclear factor 1 cooperate in activating pancreatic expression of Pax4. *J. Biol. Chem.* **278**, 38254-38259.

Song, J., Kim, H. J., Gong, Z., Liu, N. A. and Lin, S. (2007). Vhnf1 acts downstream of Bmp, Fgf, and RA signals to regulate endocrine beta cell development in zebrafish. *Dev. Biol.* **303**, 561-575.

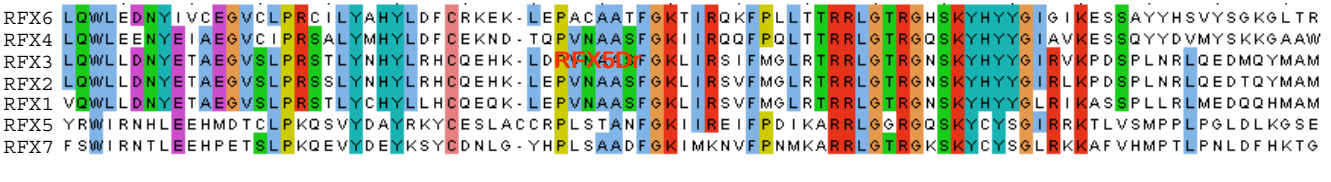
Sosa-Pineda, B., Chowdhury, K., Torres, M., Oliver, G. and Gruss, P. (1997). The Pax4 gene is essential for differentiation of insulin-producing beta cells in the mammalian pancreas. *Nature* **386**, 399-402.

Swoboda, P., Adler, H. T. and Thomas, J. H. (2000). The RFX-type transcription factor DAF-19 regulates sensory neuron cilium formation in C. elegans. *Mol. Cell* **5**, 411-421.

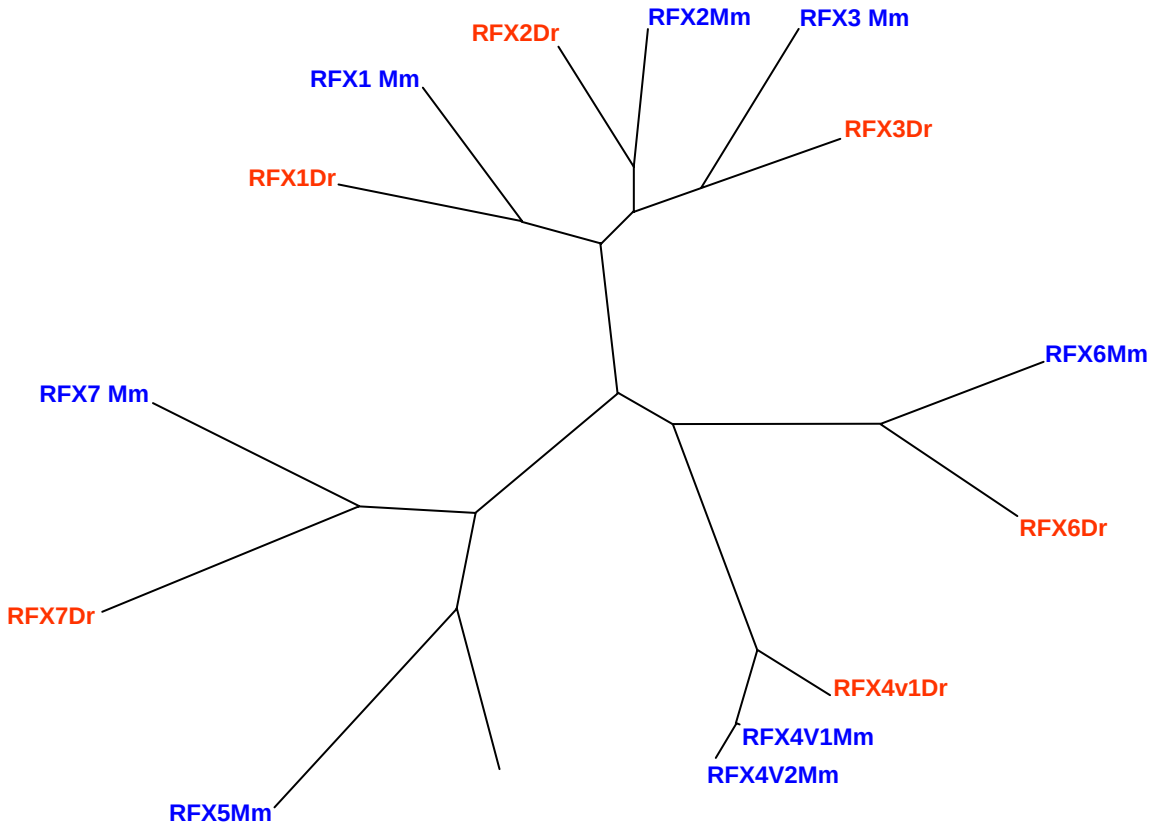
Thor, S., Ericson, J., Brannstrom, T. and Edlund, T. (1991). The homeodomain LIM protein Isl-1 is expressed in subsets of neurons and endocrine cells in the adult rat. *Neuron* **7**, 881-889.

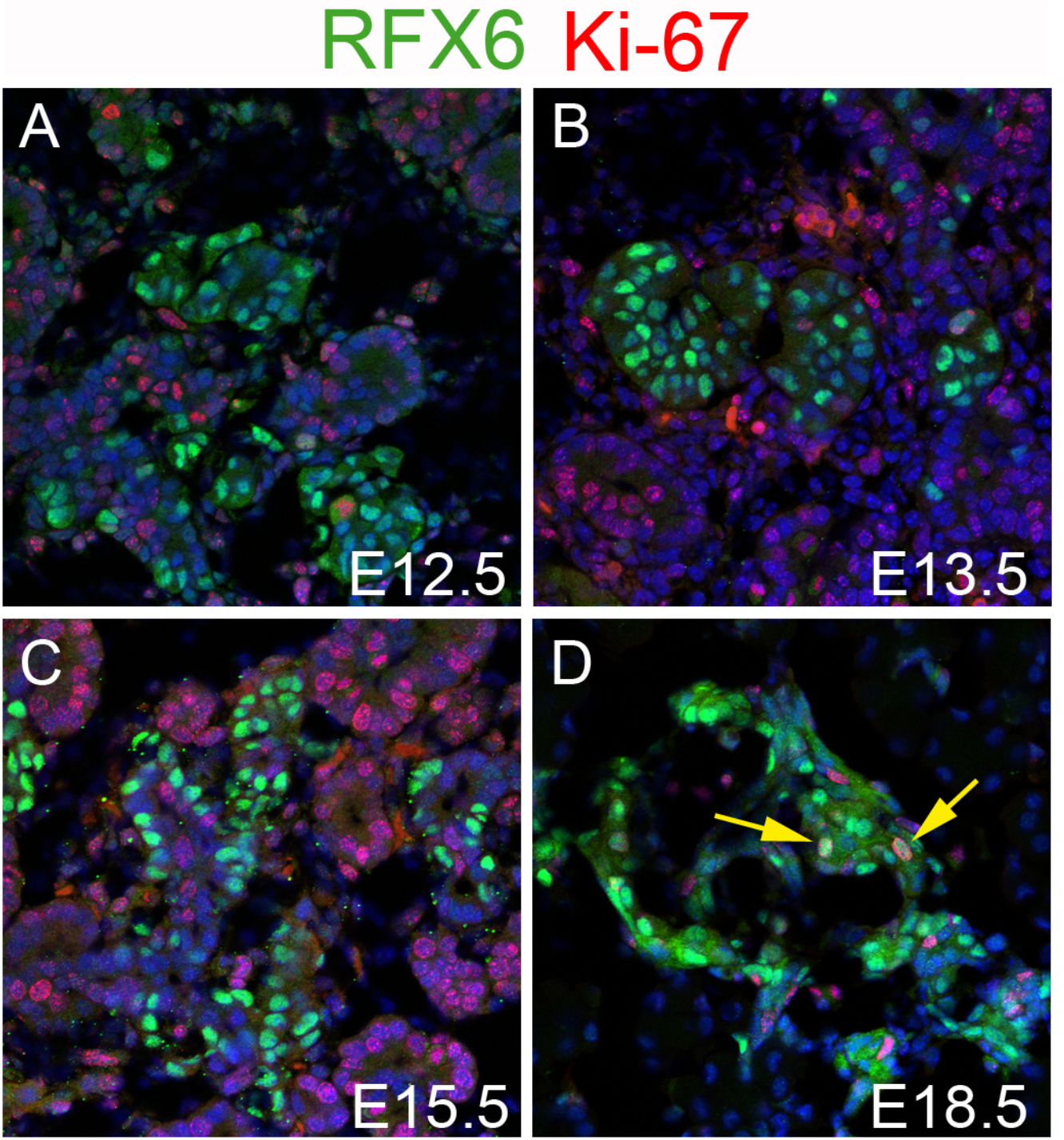
Watada, H., Scheel, D. W., Leung, J. and German, M. S. (2003). Distinct gene expression programs function in progenitor and mature islet cells. *J. Biol. Chem.* **278**, 17130-17140.
White, P., May, C. L., Lamounier, R. N., Brestelli, J. E. and Kaestner, K. H. (2008). Defining pancreatic endocrine precursors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

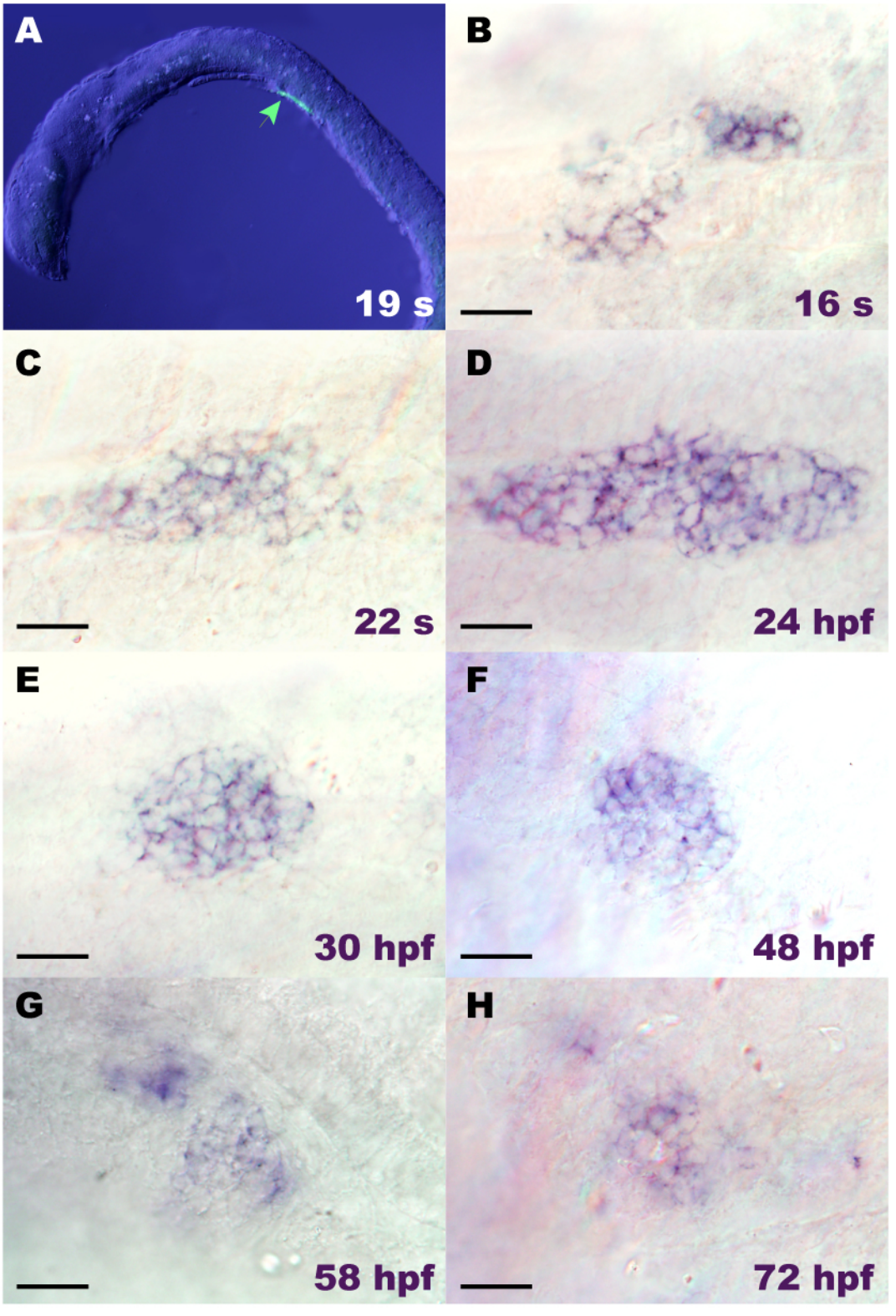
A

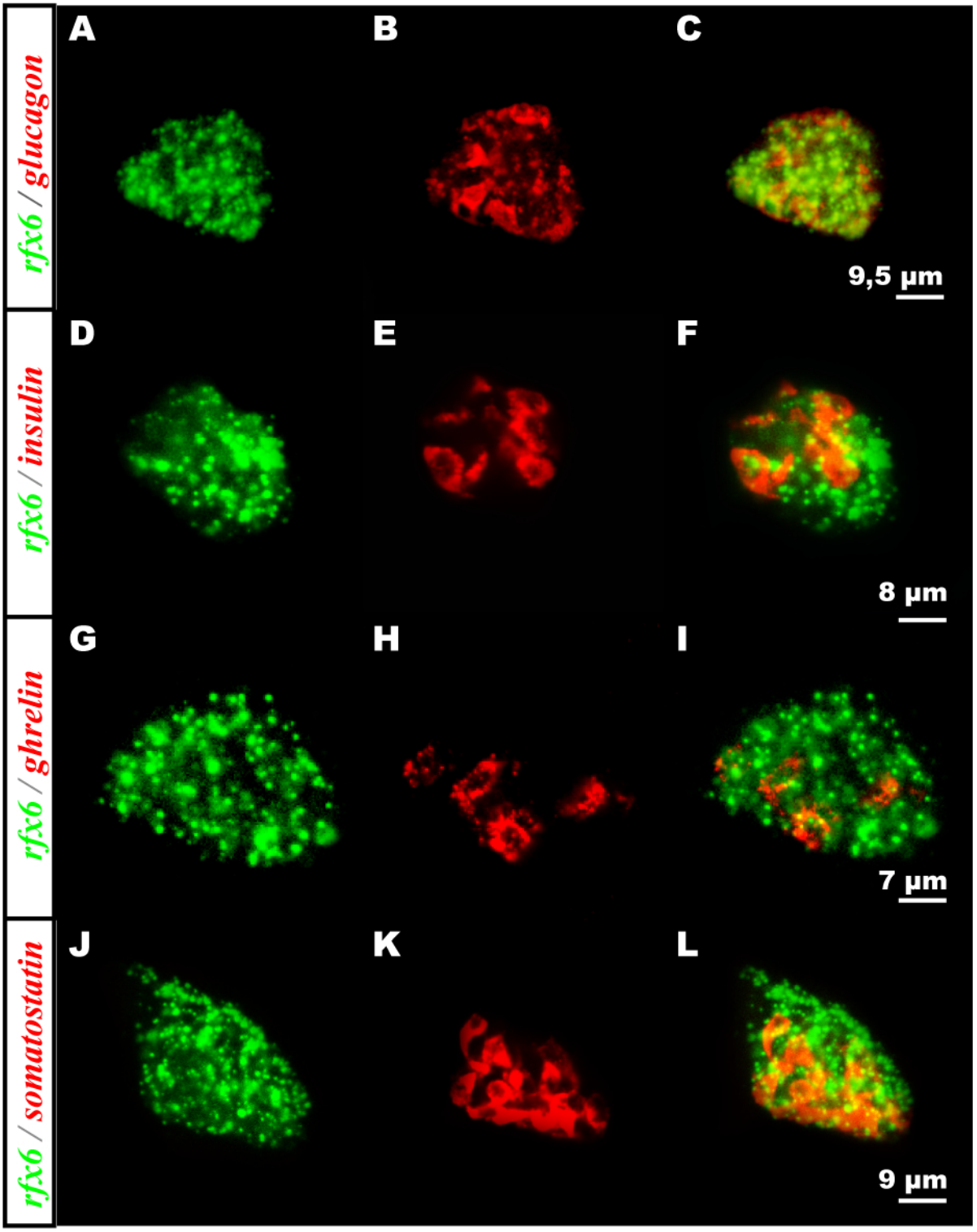

RFX6
RFX4
RFX3
RFX2
RFX1
RFX5
RFX7

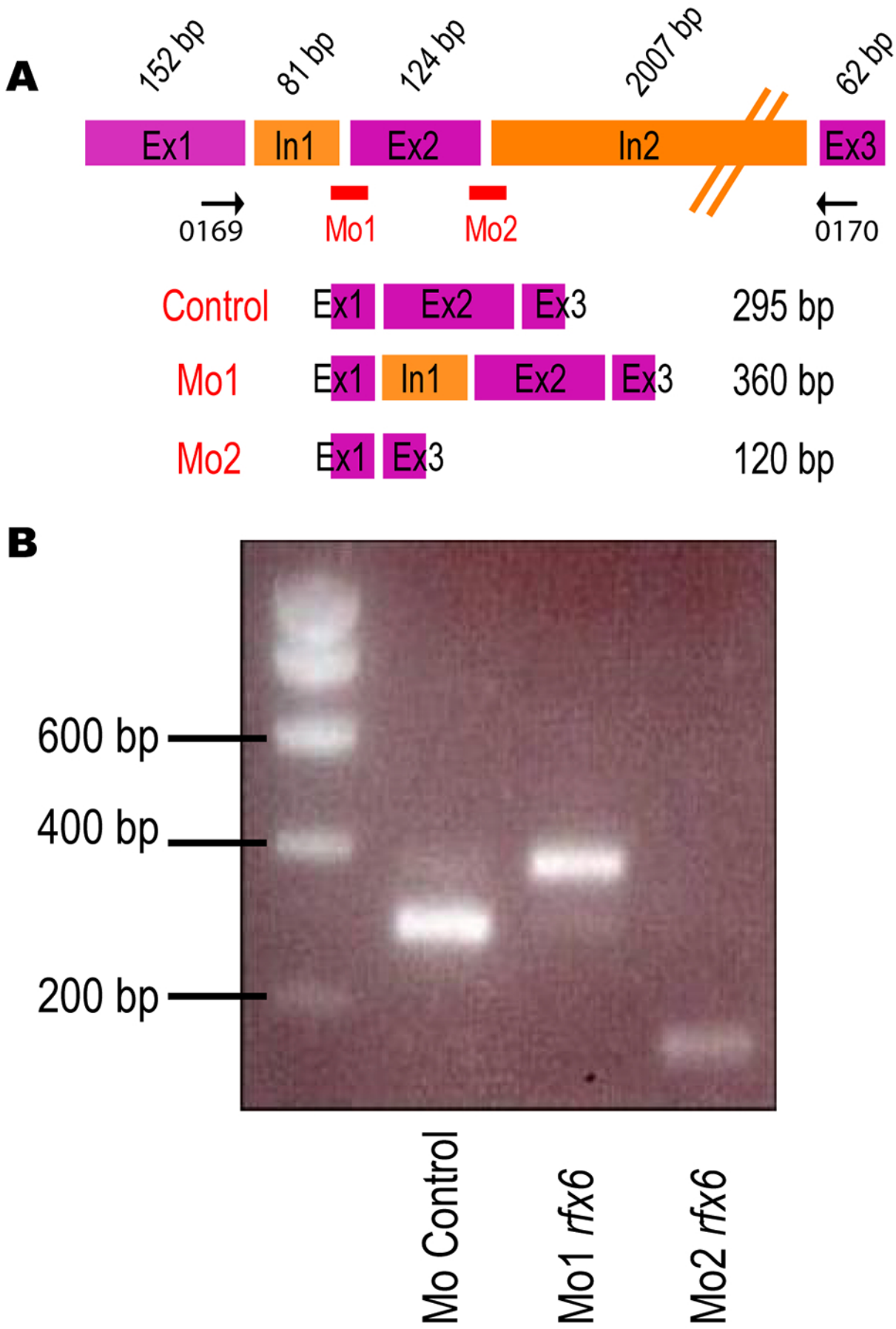
B


RFX2Dr
RFX2Mm
RFX3 Mm
RFX1 Mm
RFX3Dr
RFX1Dr
RFX6Mm
RFX7 Mm
RFX6Dr
RFX7Dr
RFX4v1Dr
RFX4V1Mm
RFX4V2Mm
RFX5Mm


RFX6 Ki-67
A
E12.5
B
E13.5
C
E15.5
D
E18.5


A
19 s
B
16 s
C
22 s
D
24 hpf
E
30 hpf
F
48 hpf
G
58 hpf
H
72 hpf


A
B
C
9,5 µm
rfx6 / glucagon
D
E
F
8 µm
rfx6 / insulin
G
H
I
7 µm
rfx6 / ghrelin
J
K
L
9 µm
rfx6 / somatostatin


A
152 bp
81 bp
124 bp
2007 bp
62 bp
Ex1
In1
Ex2
In2
Ex3
0169
Mo1
Mo2
0170
Control
Ex1
Ex2
Ex3
295 bp
Mo1
Ex1
In1
Ex2
Ex3
360 bp
Mo2
Ex1
Ex3
120 bp
B
600 bp
400 bp
200 bp
Mo Control
Mo1 rfx6
Mo2 rfx6

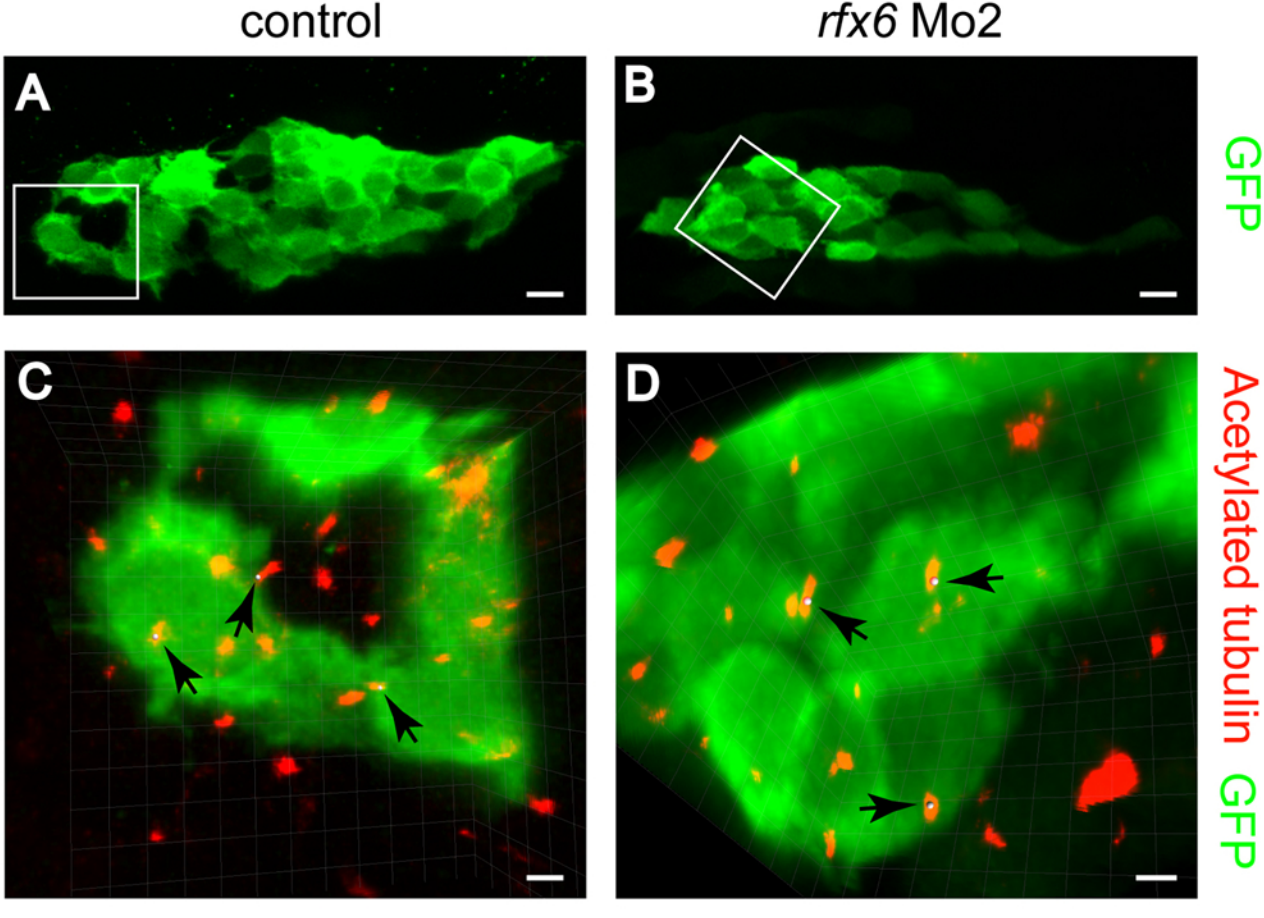

control
rfx6 Mo2
A
B
C
D
GFP
Acetylated tubulin
GFP

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

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

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