**ABSTRACT**

Insulinoma associated 1 (*Insrn1*) plays an important role in regulating the development of cells in the central and peripheral nervous systems, olfactory epithelium and endocrine pancreas. To better define the role of *Insrn1* in pancreatic endocrine cell development, we generated mice with an *Insrn1*cre reporter allele and used them to study *Insrn1*-expressing and null populations. Endocrine progenitor cells lacking *Insrn1* were less differentiated and exhibited broad defects in hormone production, cell proliferation and cell migration. Embryos lacking *Insrn1* contained greater amounts of a non-coding Neurog3 mRNA splice variant and had fewer Neurog3*Insrn1* co-expressing progenitor cells, suggesting that *Insrn1* positively regulates Neurog3. Moreover, endocrine progenitor cells that express either high or low levels of Pdx1, and thus may be biased towards the formation of specific cell lineages, exhibited cell type-specific differences in the genes regulated by *Insrn1*. Analysis of the function of *Ripplly3*, an *Insrn1*-regulated gene enriched in the Pdx1-high cell population, revealed that it negatively regulates the proliferation of early endocrine cells. Taken together, these findings indicate that in developing pancreatic endocrine cells *Insrn1* promotes the transition from a ductal progenitor to a committed endocrine cell by repressing a progenitor cell program and activating genes essential for RNA splicing, cell migration, controlled cellular proliferation, vasculogenesis, extracellular matrix and hormone secretion.

**KEY WORDS:** Pancreas development, Endocrine progenitor cells, Gene expression, Transcription factors, Mouse

**INTRODUCTION**

The genetic program responsible for the generation of pancreatic endocrine cells from endocrine progenitors remains incompletely understood (Oliver-Krasinski and Stoffers, 2008; Rieck et al., 2012; Stanger and Hebrok, 2013). Beginning at around embryonic day (E) 9-9.5 in the mouse, and peaking around E15.5, pancreatic epithelial cells express Neurog3, which encodes a master regulatory factor that activates a cascade of secondary transcription factor genes, including *Insrn1*, *Neurod1*, *Pax4*, *Arx* and *Rfx6*. Together, these and other transcription factors orchestrate the formation of the five endocrine cell types found in adult islets (Huang et al., 2000; Schwitzgebel et al., 2000; Gu et al., 2002; Collombat et al., 2003; Smith et al., 2003; Mellitzer et al., 2006; Soyer et al., 2010). Neurog3 triggers pre-endocrine cells to delaminate from a pancreatic ductal epithelium by an epithelial-to-mesenchymal transition (EMT), lose basal-apical cell polarity and epithelial adhesions, then migrate and aggregate to form islets of Langerhans (Cole et al., 2009; Gouzi et al., 2011). Moreover, *Neurog3* expression coincides both with a markedly lower rate of cell proliferation and an increase in the expression of cyclin-dependent kinase inhibitors such as Cdkn1c (p57Kip2) and Cdkn1a (p21Cip1) (Georgia et al., 2006; Miyatsuka et al., 2011).

Multiple transcription factors regulate pancreatic endocrine cell development, and they have interacting and sometimes opposing functions. For instance, Arx drives the formation of glucagon-producing α-cells. In its absence, there is a preponderance of insulin-producing β-cells and somatostatin-producing δ-cells. Similarly, *Pax4* opposes the effect of Arx and is essential for the formation of β-cells, since mice lacking this factor are characterized by an expansion in α-cells (Sosa-Pineda et al., 1997; Collombat et al., 2003). Moreover, nascent β-cells express higher amounts of Pdx1, a transcription factor crucial for the early specification of pancreatic epithelium, compared with other pre-endocrine cells (Ohlsson et al., 1993; Ahlgren et al., 1998; Fujitani et al., 2006; Nishimura et al., 2006; Gannon et al., 2008). Other transcription factors important for β-cell specification and development, such as Nkx2.2, Neurod1, Nkx6.1, Mafb and Mafa, also function in an interrelated manner (Sosa-Pineda et al., 1997; Susel et al., 1998; Nishimura et al., 2006; Nelson et al., 2007; Schaffer et al., 2013).

The expression of *Insrn1*, which encodes a zinc-finger transcription factor, is activated by the binding of Neurog3 to sites within its promoter (Mellitzer et al., 2006). Endocrine cell differentiation is markedly perturbed in mice lacking *Insrn1* (Gierl et al., 2006). In the absence of this factor, there is a reduction in the number of insulin-expressing cells, with many cells lacking any hormone expression. In addition to being expressed in developing endocrine cells throughout the gut, *Insrn1* is also expressed in the developing central nervous system, where it contributes to the formation and expansion of intermediate (basal) neural progenitors from early apical progenitor cells (Farkas et al., 2008), in the peripheral neural system and in the olfactory epithelium, where it is involved in regulating the differentiation of neurogenic progenitor cells (Wildner et al., 2008; Rosenbaum et al., 2011).

The acquisition of robust quantitative global gene transcription datasets, which are necessary for understanding the gene regulatory

---

*This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

Received 10 October 2013; Accepted 26 May 2014
network that dictates the formation and function of endocrine cells, requires the combined use of fluorescent reporter alleles, fluorescence-activated cell sorting (FACS) and next-generation sequencing technology. To this end, we have derived mice containing an Insm1GFPCre reporter allele that enabled us to isolate highly purified populations of Insm1-expressing and -deficient pre-endocrine cells, and to characterize these cell populations by RNA-Seq. In addition, we performed multi-channel FACS to isolate and differentially characterize Insm1-positive cell populations that express either high or low levels of Pdx1. In doing so, we identified many genes that are likely to contribute to both the formation and function of pancreatic endocrine cell types. In addition, the function of Ripply3 and the alternative RNA processing of Neureg3 mRNA were examined. Together, these studies provide multiple new insights into the gene regulatory network controlling pancreatic endocrine cell formation and function.

RESULTS

Generation of Insm1GFPCre reporter mice

A two-step strategy utilizing both gene targeting and recombinase-mediated cassette exchange (RMCE) was used to derive mice that express a green fluorescent protein-Cre fusion protein (GFPCre) under control of the endogenous Insm1 gene locus (Fig. 1A; supplementary material Fig. S1A-F). Insertion of GFPCre sequences into the Insm1 gene locus disrupted Insm1 protein expression, as confirmed by western blot analysis of homozygous null embryos (supplementary material Fig. S1F). Mice heterozygous for this allele (hereafter termed Insm1+/−) appeared normal, whereas animals that were homozygous for Insm1GFPCre (hereafter termed Insm1−/−) died before E15.5 and E18.5 due to defects in catecholamine biosynthesis and secretion (Wildner et al., 2008). Green fluorescence was observed in pancreatic pre-endocrine cells (Fig. 1C) and GFPCre was readily detectable by immunofluorescence in islet cells after birth [at postnatal day (P) 1, P7 and P60; data not shown] and in the developing central nervous system between E9.5 and E18.5 (Fig. 1B). GFPCre expression was also detected in the peripheral nervous system and gut endocrine cells (data not shown).

Co-staining with anti-GFP and anti-Insm1 antibodies at E15.5-18.5 in Insm1−/− pancreata showed that the majority of Insm1-positive cells co-expressed both proteins (supplementary material Fig. S2). Functionality of the Cre portion of the GFPCre fusion was also confirmed (data not shown).

Insm1 knockout mice have altered pancreatic hormone cell differentiation, replication, size and migration

To investigate the role of Insm1 in pancreas development we quantified the percentage of different pancreatic hormone-positive cells among GFPCre-expressing endocrine cells in Insm1 heterozygous and knockout animals at E18.5 (supplementary material Fig. S3). Consistent with the results of Gierl et al. (2006), 54% of endocrine cells expressed insulin in heterozygous animals, whereas only 8% of GFPCre-expressing cells were insulin positive (a 6-fold decrease) in Insm1−/− embryos. There were also less pronounced but significant decreases in cells expressing glucagon (from 24% to 11%), somatostatin (from 11% to 7%) and ghrelin (from 8% to 5%) in the null embryos. Also, the number of pancreatic polypeptide-positive cells increased from 7% to 12% in the knockout animals, as is also consistent with the findings of Gierl et al. (2006).

Since it has been suggested that Insm1 inhibits the progression of endocrine cells through the cell cycle (Zhang et al., 2009), we quantified the proliferation rates of GFPCre-positive endocrine cells by immunostaining for Ki67, a marker of cell proliferation (Fig. 2A). At E18.5 we observed up to a 7-fold decrease (from 11.6% to 1.5%) in the number of Ki67/GFPCre double-positive cells in heterozygous versus knockout pancreata (Fig. 2B). Concurrently, the size of the Insm1−/− endocrine progenitor cells is increased (Fig. 2C,D) and their shape becomes irregular. No differences in the number of apoptotic cells were observed between heterozygous and homozygous null animals by TUNEL assay (data not shown). To assess whether the proliferation defect is visible at an earlier stage, we also quantified proliferation rates at E15.5 (supplementary material Fig. S4). Although the majority of endocrine cells are postmitotic at E15.5, we were still able to detect a slight decrease in proliferation from 1.4% in Insm1+/− cells to 0.9% in Insm1−/− cells.

In addition to a reduced proliferation rate of Insm1−/− cells, we also observed that the GFPCre-positive endocrine cells remained closely associated with pancreatic ductal cells. We quantified the apparent cell migration defect by calculating the percentage of GFPCre-expressing cells that were in contact with ductal cells (DBA-positive cells) and found an increase (from 24% to 55%) in the knockout animals compared with heterozygous animals (Fig. 2E,F). This finding suggests that Insm1, in addition to its other roles, is necessary for the migration of the differentiating pre-endocrine cell away from pancreatic ductal regions.

Identification of Insm1-dependent genes in developing endocrine cells

To quantitatively assess the gene expression changes brought about by the absence of Insm1, we performed RNA-Seq using FACS-purified progenitor cells isolated at E15.5. On average, we obtained 13,456 (±2886) cells per Insm1+/− embryo and 8505 (±1602) cells per Insm1−/− embryo at E15.5, consistent with a proliferation defect. RNA-Seq resulted in the detection of 13,865 individual mRNAs with an RPKM >0.1. Pairwise comparison of the transcriptional profiles from Insm1+/− and Insm1−/− mice revealed that the expression levels for the majority of genes were highly correlated, but with a subset of the genes being differentially expressed (Fig. 3B; supplementary material Table S1). After counting only those transcripts altered in

![Fig. 1. The Insm1GFPCre allele. (A) Schematic of the Insm1GFPCre allele. Insm1 coding sequences were replaced with those encoding GFPCre using combined gene targeting/recombinase-mediated cassette exchange (RMCE) as described in supplementary material Fig. S1. The triangles represent heterozygous loxP sites and the circle a remnant FLP recognition target (FRT) site. (B) Green fluorescence in a whole mouse embryo at E11.5 broadly marks the neural system. (C) Green fluorescence in a pancreas at E15.5 marks pre-endocrine cells. Fluorescence images were overlaid with images taken with white light.](image-url)
of the more highly expressed transcripts in the knockout cells were grouped into broader categories, including cytoskeleton remodeling, transcriptional regulation and components of developmental, differentiation and morphogenesis pathways (Fig. 3C; supplementary material Fig. S5).

Genes most dependent on Insm1 for expression include those encoding pancreatic hormones, especially insulin and glucagon, as well as components of the hormone secretory machinery such as Cgth, Sec3, Snap25, Nnat and Vat11. Consistent with our immunohistochemical findings, Ppy was also more highly expressed in the Insm1−/− cells, as was Carpt, which encodes a neuroepitope. In the absence of Insm1 there was also a reduction in the expression of several transmembrane channels and transporters (Kctd12b, Slc2a2, Slc30a8, Caca1a), gap junction proteins (Gjd2), a calcium-sensing protein (Hpcal), several extracellular matrix proteins (Cryba2, Muc4, Spock2), and other less well characterized transmembrane proteins (Tmem215, Lrp11, Amigo2). Insm1−/− cells showed an increase in the expression of receptors previously associated with different progenitor cell types (Sspan8, Prom1, Ilgam, Cds9a, Lgr4) and neural progenitor cells (Nrp2, Pmp22, Fat4, Gfra1, Laml1). Interestingly, among the nucleic acid-binding proteins in Insm1−/− cells, there was a marked increase in the expression of genes encoding RNA-processing proteins such as Elav4, Cel5b and Rodl1, all of which have been shown to be involved in alternative RNA splicing (Chapple et al., 2007; Spellman et al., 2007; Ince-Dunn et al., 2012). Genes for DNA-binding proteins, such as Fve, Sl18, Rippy3, Lmo3 and Jazf1, were upregulated to a lesser extent. Conversely, Insm1−/− cells exhibited a marked increase in the expression of transcription factors that play important roles in the maintenance of progenitor cells (Bhlhe22, Rest, Nr5a2, Hes1, Acscl1, Foxp2, Sox6). Markers of EMT (Zeb2, Mmp14, Cldn11, Cldb, Vim, Cdh11, Fn) and components of Notch signaling pathway (Hes1, Hes5, Hey1, Dll1, Notch2, Notch1) were also increased in the Insm1−/− progenitor cells. Similarly, we observed an upregulation in genes involved in other signaling pathways controlling cell progenitor states, such as TGF/BMP (Smad3, Bmpr1b, Tgfβ2, Bmp1), FG (Fgfr1, Fgfr4) and IGF (Igfbp1, Igfbp3). Several genes involved in cell proliferation (Cdkn1c, Bcl2, Merk, Anxa2, Ccdn1) were increased in Insm1−/− cells, whereas others (Cdkn1b, Manf and Akt3) were more highly expressed in the Insm1+/− cells. Insm1−/− cells also showed greater expression of genes involved in cytoskeleton reorganization and cell migration (Gng12, Pak3, Filip1), vascularization-promoting growth factors (Pgf, Vegfa) and a β-cell-specific MPK8 scaffold protein (Mapk8ip2).

We performed chromatin immunoprecipitation (ChiP) analysis of the Insm1-regulated genes Rest, Cldcn1c, Cldcn1b and Ripply3, all of which contain putative Insm1 binding sites within their promoters based on the JASPAR database (Mathelier et al., 2014) (supplementary material Table S2). All genes showed enrichment that was similar or greater than that of the Insm1 promoter itself, which was previously shown to be downregulated by Insm1 (Breslin et al., 2002) (supplementary material Fig. S6).
junctions for a number of pro-endocrine genes. Interestingly, Neurog3 mRNA was found to be alternatively spliced in endocrine progenitors at E15.5, with two splice variants present – long coding and a short non-coding, with the shorter variant having about twice as many sequence tags in Insm1−/− cells than in Insm1+/− cells. To validate the presence of alternatively spliced Neurog3 transcripts, we designed primers to simultaneously amplify both splicing variants or just the long or short variant (Fig. 4A), and performed semi-quantitative RT-PCR. This experiment, as well as RT-qPCR, confirmed the expression of both Neurog3 mRNA isoforms (Fig. 4B,C; supplementary material Fig. S7), with the ratio of long to short variant expression being 2-fold higher in Insm1+/− than in Insm1−/− cells. Immunostaining of E15.5 embryos also showed a decrease in the number of GFP/Neurog3 co-expressing cells (19% in Insm1+/− compared with 12% in Insm1−/− tissues; Fig. 4D,E), as might be expected if there was a reduction in Neurog3 protein levels.

Validation of RNA-Seq results by RT-qPCR and further characterization of novel genes

To further validate and quantify the differential expression of selected genes, we performed RT-qPCR using 96-well TaqMan array cards that were designed to detect genes showing higher expression in Insm1+/− versus Insm1−/− cells as well as some pancreatic endocrine genes as a reference (supplementary material Table S3). Cells were isolated at E15.5 and E18.5, with the latter time point serving to better detect changes occurring as the endocrine progenitor cells became more differentiated. This analysis confirmed the upregulation or downregulation of all these genes at E15.5 (supplementary material Fig. S8A) and allowed us to distinguish between those required for later stages of endocrine differentiation and maturation (upregulated in Insm1+/− versus Insm1−/− samples at E18.5, e.g. Mnx1, Mafa, Ripply3 and Fev) and those active at earlier stages of endocrine specification (downregulated in Insm1+/− versus Insm1−/− samples at E18.5, e.g. Neurod1, Runx1t1, Pax6 and Pax4). The expression of some of these genes in pancreatic endocrine cells was confirmed by examining in situ RNA hybridization images in the EUREXPRESS database obtained from E14.5 mouse embryos (Diez-Roux et al., 2011) (supplementary material Fig. S8B). In addition, we performed staining with antibodies against surface proteins, such as cadherin Celsr3, adhesion molecule Amigo2, and low-density lipoprotein receptor-related protein Lrp11 (supplementary material Fig. S9), and found preferential staining in each case for pre-endocrine cells.
**Temporal expression profiling of endocrine populations expressing high and low levels of Pdx1**

Beginning at ~E14.5 during pancreas development, Pdx1 is expressed at higher levels in β-cells than in non-β-cells (Ohlsson et al., 1993). Thus, to identify *Insm1*-regulated genes that might be specifically important for β-cell development and function we used a *Pdx1* ^CrePP^ allele (Potter et al., 2011) to isolate populations of *Insm1*-positive endocrine cells that express either high (HI) or low (LO) levels of *Pdx1* at time points from E15.5 to E18.5. Confocal imaging of direct fluorescence confirmed that *Insm1* ^Cre/+;Pdx1CFP/+^ and *Insm1* ^Cre/+;Pdx1CFP/^ mice expressed both GFP and CFP in endocrine cells, with different levels of CFP expression marking endocrine subpopulations (Fig. 5A), which we were able to separate by FACS (Fig. 5B). This allowed us to assess changes in the relative proportions of different cell types during pancreas development over a 4 day interval (Fig. 5C). *Insm1* ^+Pdx1^ double-positive endocrine cells comprised ~20-30% of total sorted cells, with the greatest increase in cell numbers occurring between E16.5 and E17.5. As expected, the proportion of *Pdx1*-HI cells in the double-positive cell populations increased between E15.5 and E18.5, mirroring the increase in β-cells during pancreas development.

To explore the differences between the *Insm1*/Pdx1-HI and *Insm1*/Pdx1-LO pre-endocrine cell populations as a function of time, we performed both RT-qPCR and RNA-Seq analysis. Pairwise comparison of RNA-Seq results between these cell populations at E15.5 and E18.5 revealed an increase in the expression of a number of genes important for β-cell specification (supplementary material Tables S4 and S5). Gene ontology analysis showed that most of the *Insm1*/Pdx1-HI enriched genes are involved in hormone secretion as well as ion transport and glucose metabolism (supplementary material Fig. S10A,B). Ninety-three genes were commonly upregulated in *Insm1*/Pdx1-HI versus *Insm1*/Pdx1-LO cells at E15.5 and E18.5. Correlation analysis showed that many of these genes have similar differential expression levels and many are known to be important for β-cell function (supplementary material Fig. S10C).

RT-qPCR analysis of pre-endocrine cell populations from E15.5 to E18.5 allowed hierarchical clustering based on relative levels of expression of the genes in the *Insm1*/Pdx1-HI versus *Insm1*/Pdx1-LO cells (supplementary material Fig. S11). Genes expressed at higher levels in the *Insm1*/Pdx1-LO cells included the α-cell transcription factor Arx, the hormones Glc and Ghrl, as well as many transcription factors known to be involved in early endocrine specification, such as Neurod1, Neurog3, Isl1 and Pax4. These results confirmed our prediction that the *Insm1*/Pdx1-LO population consists mostly of non-β-cells and non-committed pre-endocrine cells. Interestingly, somatostatin and Ppy were expressed at a higher level in the *Insm1*/Pdx1-HI population, especially at E15.5. Other genes that were more highly expressed in the *Insm1*/Pdx1-HI cells included known β-cell genes and transcription factors such as Ins2, Glut2, Mafa, Mnx1 and Nkx6.1. This population also contained the transcriptional regulators Mafg and Ripply3 and the signaling molecule Igfbp7.

---

**Fig. 4. Neurog3 is alternatively spliced in endocrine progenitor cells.** (A) Neurog3 gene and mRNA splicing isoforms. Primer pair N1+N3 is designed to amplify both isoforms, N1+N2 the long coding isoform, and N1+N4 the short non-coding isoform. (B) Semi-quantitative PCR analysis for Neurog3 mRNA isoforms. The proportion of the short non-coding isoform is higher in *Insm1* ^+/−^ than in *Insm1* ^−/−^ cells. (C) Quantification of relative expression of the Neurog3 long versus short isoforms by real-time PCR. (D) Immunofluorescence labeling of pancreatic tissues from *Insm1* ^+/−^ and *Insm1* ^−/−^ mice at E16.5 with anti-GFP antibodies (green) and anti-Neurog3 antibodies (red). (E) The percentage of Neurog3-positive cells among GFP-positive cells is decreased in *Insm1* ^+/−^ mice. Error bars indicate s.e.m. (n≥6); *P*-values were determined by Student’s t-test. Scale bar: 50 μm.
**DISCUSSION**

**Role of _Insm1_ in pancreatic pre-endocrine cell differentiation**

In agreement with previously published data (Gierl et al., 2006), mice without _Insm1_ have a reduced number of endocrine cells and hormone expression in these cells is either lacking or imbalanced in terms of specific endocrine cell types. Given that _Insm1_ has long been known to contain a SNAG repressor domain (Chiang and Ayyanathan, 2013), our finding that approximately twice as many _Insm1<sup>−/−</sup>_ than in _Insm1<sup>+/−</sup>_ mice were increased in expression is consistent with the major role of _Insm1_ being as a repressor. In the absence of _Insm1_, pancreatic pre-endocrine cells display many features of earlier stage progenitor cells, including an increase in the expression of transcription factors and cell surface receptors associated with the progenitor state, EMT markers, and components of morphogenic signaling pathways (e.g. Notch, Tgf/BMP and Wnt). Although EMT markers increased in _Insm1_ null cells, they failed to migrate away from the ducts. This might be due to incomplete EMT, since some of the genes upregulated in _Insm1<sup>−/−</sup>_ cells are involved in actin remodeling and cell migration. Moreover, signaling molecules and extracellular matrix proteins that are expressed upon _Insm1_-dependent endocrine differentiation might be necessary for the migration and aggregation of newly differentiated endocrine cells into islets. _Insm1_ null cells also have decreased rates of proliferation that starts at E15.5 and is well pronounced at E18.5. This decrease was not accompanied by an increase in apoptosis, and the knockout cells appeared larger in size, a characteristic that often accompanies cell cycle arrest at G1/G2 phase (Bjorklund et al., 2006). Consistent with cell cycle arrest without apoptosis, _Insm1_ null cells express more of the cell cycle inhibitor Cdkn1c and the pro-survival gene Bcl2. Interestingly, _Insm1_-expressing cells make more of the cell cycle inhibitor Cdkn1b (p27<sup>Kip1</sup>), which is involved in controlling the proliferation of newly differentiated β-cells (Georgia and Bhushan, 2006). This suggests that _Insm1_ promotes a switch in the mode of cell cycle regulation from being Cdkn1c based to being Cdkn1b regulated, with both genes being directly regulated by _Insm1_. A similar developmental switch in cell cycle regulators has been observed during pituitary development, in which non-cycling precursors express Cdkn1c and more differentiated cells express Cdkn1b (Bilodeau et al., 2009). Although the differentiation of endocrine cells in the presence of _Insm1_ is associated with an increase in proliferation, cell proliferation itself is under the control of multiple genes, including cell cycle inhibitors such as Cdkn1b and _Ripply3_. The complex regulation of endocrine progenitor cell proliferation is not surprising given that overexpression of _Insm1_ in the developing neural cortex has been shown to decrease the proliferation of less differentiated apical progenitor cells but to increase the proliferation of more differentiated intermediate basal progenitor cells (Farkas et al., 2008).
This study also brings to light several genes that might play important but previously unrecognized roles in the formation of islet cells. Indeed, although a key function of *Insm1* might be to repress genes (e.g. *Rest*) that maintain pancreatic ductal epithelial cells in a progenitor-like state, which would be expected to promote endocrine cell differentiation, many genes were also found to be upregulated in *Insm1*-expressing cells. These include genes necessary for hormone secretion, membrane transporters and channels, extracellular matrix proteins, RNA-processing proteins and cell migration proteins. For example, *Insm1*-expressing progenitor cells are characterized by higher levels of transcription factors that are found in differentiating β-cells (*Mafa*, *Mafb*, *Mnx1*, *Vdr*, *Hmgn3*, *Fev*, *Lmx1b*), as well as novel genes that have not been studied in the pancreas (*Lmo3*, *Ripply3*, *Jazf1*, *Nhlh1*). Among the *Myt1* family, *St18*, *Myt1* and *Myt1l1* were all upregulated, suggesting that all three of these genes are expressed in differentiating endocrine cells, contrary to a previous report (Gu et al., 2004). Considering that many of the *Insm1*-regulated transcription factors have also been shown to play key roles during neural development, and that neural and islet cells share many phenotypic properties, these genes probably also contribute to endocrine development. Interestingly, one such gene, the transcription factor *Jazf1*, is associated with the development of type 2 diabetes (Zeggini et al., 2008).

**Identification of novel endocrine surface markers**

This study revealed a number of surface proteins that mark early endocrine cells. For example, mRNAs for *Amigo2*, *Celsr3* and *Lrp11* were strongly upregulated in *Insm1*-positive cells, and immunostaining for the encoded proteins confirmed their localization along the membrane of endocrine cells. *Amigo2* showed the strongest endocrine-specific staining pattern and may function in cell adhesion and migration (Rabenau et al., 2004). *Celsr3*, the cadherin associated with planar polarity, displayed localized membrane staining and was recently shown to promote endocrine cell differentiation (Tissir and Goffinet, 2006; Cortijo et al., 2012). Another early endocrine cell marker validated by our study is *Lrp11*, as was previously suggested by Hald et al. (2012). Interestingly, although *Lrp11* is present in all endocrine cells, it seems to be most highly expressed in δ-cells.

**Neurog3 is alternatively spliced in endocrine progenitor cells**

RNA-Seq revealed that ~30% of the differentially expressed genes had two or more differentially expressed RNA variants (data not shown). This suggests that the *Insm1*-dependent differentiation program might also function at a post-transcriptional level by promoting endocrine lineage-specific alternative splicing. Although the mechanisms involved in this regulation are unclear, this could occur through the activation of genes encoding RNA-binding proteins.
proteins, such as Elav4, Celf3 and Rodl. All of these transcripts were strongly upregulated in Insm1-positive cells and so could be important in regulating splicing changes during endocrine differentiation. In particular, we found that Neurog3 mRNA is alternatively spliced in endocrine progenitor cells, with a long coding and short non-coding isoform present. The latter is formed through splicing to the alternative splice acceptor site located in the second exon of Neurog3 (supplementary material Fig. S7). Interestingly, the ratio of the Neurog3 coding to non-coding isoform was higher in Insm1-positive cells, suggesting that alternative splicing of the Neurog3 pre-mRNA in an Insm1-dependent manner provides a positive-feedback mechanism for maintaining Neurog3 expression during development.

Transcriptional profiling of subpopulations of pre-endocrine cells

Combinatorial FACS using two or more fluorescent reporter alleles is a powerful means of isolating specific progenitor cell subpopulations. For this reason, we also made use of a recently derived Pdx1CFP fluorescent reporter allele to explore the differences in gene expression between developing endocrine progenitor cells that express either high or low levels of Pdx1. It is well known that Pdx1 marks all endocrine cells (Herrera, 2000; Gu et al., 2002) and that developing β-cells, as well as a small fraction of δ-cells, express high levels of Pdx1 (Leonard et al., 1993; Ohlsson et al., 1993; Gu et al., 1995). As expected, Insm1/Pdx1-LH cells express more β-cells-specific genes, such as Ins2, Slc2a2a and Mafa. However, we were surprised to find that this cell population (at E15.5) also contains many somatostatin and Ppy transcripts. This suggests that the Insm1/Pdx1-LH cell population might contain a common β/δ-progenitor, as well as newly formed δ-cells, the existence of which has been predicted by Pax4 and Arx knockout studies (Collombat et al., 2003, 2005). Our data are also consistent with previous lineage-tracing studies that have shown that a Ppy-Cre transgene marks both β-cells and Ppy-producing cells (PP-cells) (Herrera et al., 1994; Herrera, 2000), a finding that suggests that PP-cells and a common β/δ-cell might share a common progenitor, or that Ppy is expressed in the common β/δ-cell progenitor. Analysis of differential expression between Insm1/Pdx1-LH and Insm1/Pdx1-LO endocrine subpopulations allowed us to define Insm1-dependent genes that are enriched in Pdx1-LH populations and therefore potentially important for specific functions of β-cells and/or δ-cells.

Ripply3 as a downstream modulator of endocrine cell proliferation

In this study we discovered that Ripply3 is also expressed in endocrine progenitor cell populations and that Insm1 binding to the Ripply3 promoter. Ripply3 belongs to a family of related proteins, the other two members of which, Ripply1 and Ripply2, are necessary for the proper formation of somite boundaries and function by antagonizing T-box transcription factor activity during development (Kawamura et al., 2005; Hitachi et al., 2009; Takahashi et al., 2010). Ripply3 has also been shown to regulate pharyngeal apparatus development through suppression of Tbx1 activity (Okubo et al., 2011). The importance of other Ripply gene family members prompted us to explore the role of Ripply3 in pancreas development. Although our analysis did not reveal any defects in endocrine cell differentiation in Ripply3 null pancreatic tissues, we observed that both β-cell and α-cell areas and proliferation were increased.

The increase in proliferation in Ripply3 null tissues suggests that, similar to Cdkn1b, Ripply3 might control the proliferation of newly committed β-cells in order to allow them to properly differentiate.
Mouse lines
The Insm1GFPCre allele was made using a sequential gene targeting and RMCE strategy using methods and vectors that have been described previously (Chen et al., 2011). Insm1GFPCre and Pdx1GFPCre mice (Potter et al., 2011) were maintained on a CD1 outbred background (90% or higher) to extend the survival of the embryos, as previously described (Gierl et al., 2006). Approximately 10% of the Insm1 homozygous null pups survived to E18.5, consistent with the results of Gierl et al. (Gierl et al., 2006), with all other embryos dying between E15.5 and E18.5. Ripply3fl/fl mice were of mixed C57BL/6J×129Sv background. All animal procedures and husbandry protocols were approved by the Vanderbilt University Institutional Animal Care and Use Committee.

Embryo isolation and imaging, tissue dissociation and FACS
Timed matings were carried out by designating noon on the day vaginal plugs were identified as E0.5. Embryos at different ages were obtained by dissection and imaged by direct fluorescence visualization using a Leica MZ 16 FA stereo equipped with a QImaging RETIGA 4000R camera. For FACS, embryos were isolated at E15.5-18.5 and initially genotyped by direct fluorescence. Pancreata were dissected in cold PBS and dissociated with Accumax solution (Sigma-Aldrich) containing 5 μg/ml DNase I (Sigma-Aldrich) at 37°C for 1 h with periodic resuspension. Quenching solution [Li15 medium, 10 μM Hepes (pH 7.0), 1 mg/ml bovine serum albumin (Life Technologies), 5 μg/ml DNase I] was added to stop the reaction. The cell suspension was filtered through a 35 μm filter into FACS tubes, washed with quenching solution and resuspended in 0.5 ml quenching solution containing 7AAD viability dye (1:1000; Life Technologies). Fluorescent cells were sorted directly into Trizol LS reagent (Life Technologies) containing 40 μg/ml mussel glycogen (Sigma-Aldrich) using a FACSAria I cell sorter (BD Biosciences).

RNA isolation, library construction and RNA-Seq
Three independent RNA isolates from each genotype were used for cDNA amplification and sequencing. Cells from multiple embryos were pooled to obtain a sufficient amount of RNA for analysis. The methods used to isolate RNA, validate integrity and purity, amplify cDNA and make libraries have been described previously (Choi et al., 2012). Single-end sequencing (110 bp) was performed on an Illumina HiSeq2000 genome analyzer. Read alignment to the mouse genome (mm9) was performed using RNA-Seq Unified Mapper (RUM) (Grant et al., 2011). Expression level was quantified as reads per kilobase of gene model per million mapped reads (RPKM). Genome alignment of sequencing data yielded total mapped reads from each gene row. For analysis of alternative splicing variants, 1 ng amplified CDNA (used for RNA-Seq) was amplified using either real-time PCR with Power SYBR Green PCR Master Mix (Life Technologies) or semi-quantitative PCR using Phusion Taq (New England Biolabs).

Immunochemistry
Immunofluorescence staining of frozen sections was essentially as previously described (Burlison et al., 2008). Ductal cells were stained with biotinylated Dolichos biflorus agglutinin (DBA; Vector Laboratories) followed by streptavidin-Cy5 conjugate (Vector Laboratories). Coverslips were mounted with ProLong Gold antifade reagent with DAPI (Life Technologies). Fluorescence images were acquired using an Axioskop2 microscope (Zeiss) with 20× and 40× objectives and QImaging RETIGA EXi camera. Confocal microscopy images were acquired using an LSM 510 Meta microscope and imaging software (Zeiss). Images taken with different color channels were merged using Photoshop (Adobe Systems).

Morphometric analysis
The fractions of GFP-positive or β-galactosidase-positive endocrine cells co-expressing other proteins were calculated by manual cell counting on step sections (50 μm apart) of pancreatic tissues from at least six different animals of the same age and genotype. At least 500 cells were counted for each phenotype in each experiment. The percentage of GFP-positive cells adjacent to ducts was calculated based on the number of these cells in contact with DBA-positive cells. The cell area of GFP-positive endocrine cells was measured using ImageJ (NIH) software to outline E-cadherin-positive cell borders using a total of 1000 cells from different animals of each phenotype. The insulin-positive and glucagon-positive areas of Ripply3 heterozygous and mutant animals were measured using every fifth section from three different animals of the same phenotype.

ChiP
Pancreata were dissected from E15.5 wild-type embryos, cross-linked with 1% formaldehyde in PBS for 8 min, quenched with 125 mM glycine and washed with PBS. Tissues were lyzed in 100 μl lysis buffer [50 mM Tris-HCl (pH 8.0), 10 mM EDTA, 1% SDS, 1× protease inhibitor mix (Sigma), 1 mM PMSF], sonicated (Bioruptor, Diagenode) and lysates diluted to 1 ml with RIPA buffer [10 mM Tris-HCl (pH 7.5), 140 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1% Triton X-100, 0.1% (w/v) SDS, 1× protease inhibitor mix, PMSF]. 20 μg sonicated chromatin was incubated with 8 μg antibody coupled to IgG magnetic beads (Cell Signaling Technology) overnight at 4°C. The beads were washed five times with RIPA buffer and once with TE buffer [10 mM Tris (pH 8.0), 1 mM EDTA]. After washing, bound DNA was eluted at 65°C in elution buffer [20 mM Tris-HCl (pH 7.5), 5 mM EDTA, 50 mM NaCl, 1% SDS, 50 μg/ml proteinase K] for 2 h. After cross-linking reversal, the immunoprecipitated DNA was purified by phenol-chloroform extraction and ethanol precipitation and resuspended in 30 μl TE buffer. Enrichment at target promoters was determined by real-time PCR with Power SYBR Green PCR Master Mix. Relative fold enrichment at different sites was calculated by the 2−ΔΔCT method from five different ChIP experiments.

Acknowledgements
We thank Guoqiang Gu, Jennifer Stancill and Hannah Worcel for critically reviewing the manuscript; David K. Flaherty and Brittany Mallock for providing FACS expertise and technical assistance; Jennifer Skelton for help with deriving the mice; and Travis Clark for assistance with the next-generation sequencing.

Competing interests
The authors declare no competing financial interests.

Author contributions
A.B.O., Q.L., J.S., C.J.S. and M.A.M. were involved in the conception and design of the experiments and co-wrote the manuscript. A.B.O. performed most of the experiments and analyzed the data; Q.L. designed the gene targeting vector; R.G. performed embryo dissections and stereoscopy; S.B.H. performed the gene
targeting and RMCE procedures; E.M. and C.J.S. performed initial RNA-Seq data analysis; T.O. and S.T. provided tissues from Ripply3 knockout mice. All authors were provided with the opportunity to comment on the manuscript.

Funding
These studies were supported by National Institutes of Health (NIH) grants [DK72473 and DK85892 to M.A.M. Use of shared resources at Vanderbilt University School of Medicine was supported by NIH grants [CA68485, DK58404 and DK20593]. Deposited in PMC for immediate release.

Supplementary Material
Supplementary material available online at http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.104810/-/DC1

References


Vance, K. W., Carreira, S., Brosch, G. and Goding, C. R. (2005). Tbx2 is overexpressed and plays an important role in maintaining proliferation and suppression of senescence in melanomas. Cancer Res. 65, 2260-2266.

