PDX1 Binds and Represses Hepatic Genes to Ensure Robust Pancreatic Commitment in Differentiating Human Embryonic Stem Cells

Adrian Kee Keong Teo,1,5 Norihiro Tsuneyoshi,1 Shawn Hoon,2 Ee Kim Tan,1 Lawrence W. Stanton,3 Christopher V.E. Wright,4 and N. Ray Dunn1,*
1Institute of Medical Biology, A*STAR (Agency for Science, Technology and Research), 8A Biomedical Grove, #06-06 Immunos, Singapore 138648, Singapore
2Molecular Engineering Lab, A*STAR, 61 Biopolis Drive, Proteos #03-13, Singapore 138673, Singapore
3Genome Institute of Singapore, A*STAR, 60 Biopolis Street, #02-01 Genome, Singapore 138672, Singapore
4Department of Cell and Developmental Biology and Vanderbilt University Program in Developmental Biology, Vanderbilt University Medical Center, 3144 MRBIII, 465 21st Avenue South, Nashville, TN 37232, USA
5Present address: Institute of Molecular and Cellular Biology, A*STAR, 61 Biopolis Drive, Proteos, Singapore 138673, Singapore
*Correspondence: ray.dunn@imb.a-star.edu.sg
http://dx.doi.org/10.1016/j.stemcr.2015.02.015
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

SUMMARY

Inactivation of the Pancreatic and Duodenal Homeobox 1 (PDX1) gene causes pancreatic agenesis, which places PDX1 high atop the regulatory network controlling development of this indispensable organ. However, little is known about the identity of PDX1 transcriptional targets. We simulated pancreatic development by differentiating human embryonic stem cells (hESCs) into early pancreatic progenitors and subjected this cell population to PDX1 chromatin immunoprecipitation sequencing (ChIP-seq). We identified more than 350 genes bound by PDX1, whose expression was upregulated on day 17 of differentiation. This group included known PDX1 targets and many genes not previously linked to pancreatic development. ChIP-seq also revealed PDX1 occupancy at hepatic genes. We hypothesized that simultaneous PDX1-driven activation of pancreatic and repression of hepatic programs underlie early divergence between pancreas and liver. In HepG2 cells and differentiating hESCs, we found that PDX1 binds and suppresses expression of endogenous liver genes. These findings rebrand PDX1 as a context-dependent transcriptional repressor and activator within the same cell type.

INTRODUCTION

The adult pancreas is comprised of two major functional compartments. The endocrine pancreas contains hormone-secreting cells, with the glucose-sensing, insulin-secreting β cells garnering tremendous clinical attention for their role in the pathology of diabetes. The exocrine pancreas contains digestive enzyme-secreting acinar cells and ductal epithelial cells that produce mucin and plumb the pancreas, providing transport channels for the secretion of acinar enzymes into the duodenum. All of these tissues originate embryonically from the definitive endoderm (DE), one of the three primary germ layers formed during gastrulation (reviewed extensively in Pan and Wright, 2011; Shih et al., 2013).

By embryonic day (E) 8.5 in the mouse embryo, the DE forms a sheet of simple squamous epithelium lining the ventral surface of the lordotic embryo. Anteriorly, a specialized pit termed the anterior intestinal portal (AIP) forms and its continued invagination results in a pocket—the foregut—whose descendants give rise to the esophagus, lungs, thyroid, liver, pancreas, and biliary system. Morphologically, the incipient pancreas is recognizable as dorsal and ventral evaginations on opposing sides of the posterior foregut around E9.5. At this stage, the pancreatic primordium is conspicuously labeled by the expression of the ParaHox gene Pancreatic and Duodenal Homeobox 1 (Pdx1) (Ahlgren et al., 1996; Guz et al., 1995; Jørgensen et al., 2007; Offield et al., 1996). Lineage tracing experiments show that these early Pdx1+ cells are multipotent, giving rise to all three components of the mature pancreas (Gannon et al., 2000; Gu et al., 2002). Loss-of-function studies further demonstrate that Pdx1 is not only an indelible marker of the early pancreatic lineage but also is required cell autonomously for early pancreas formation (Ahlgren et al., 1996; Jonsson et al., 1994; Offield et al., 1996). In Pdx1 null mutant embryos, patent dorsal and ventral buds form but quickly regress, resulting in pancreatic agenesis, severe postnatal hyperglycemia, and eventual death. PDX1 is similarly expressed in the developing human pancreas (Jennings et al., 2013), and in 1997, a case of rare human pancreatic agenesis was reported in an individual homozygous for a cytosine deletion in codon 63 of PDX1 (Pro63fsdelC) (Stoffers et al., 1997b). This result underscores why PDX1 activation is considered an obligatory mile marker for the successful in vitro production of therapeutically relevant β cells from pluripotent human stem cells.

Despite the central role of Pdx1/PDX1 in orchestrating pancreatic morphogenesis in mice and humans, vanishingly little is known about its direct transcriptional targets. Here, we chose to address this deficit by developing an in vitro human embryonic stem cell (hESC) differentiation protocol that specifically captures robust numbers of early multipotent, proliferative PDX1+...
pancreatic progenitor (ePP) cells. Based on extensive molecular marker analysis, ePP cells on day 17 of differentiation strongly resemble the early mammalian dorsal and ventral pancreatic buds. We therefore performed chromatin immunoprecipitation followed by massively parallel DNA sequencing (ChIP-seq) in an effort to elaborate the pancreatic gene regulatory network over which PDX1 presides. Our analyses identified more than 350 genes who are simultaneously bound by PDX1 (within ± 20 kb of the transcriptional start site [TSS]) and whose expression is upregulated on day 17 of differentiation. We also unexpectedly found that PDX1 binds classic liver marker genes such as ALB, AFP, and TTR, which label hepatoblasts, the precursors to hepatocytes, in the developing mammalian embryo. This result suggested that acquisition of pancreatic identity requires direct and simultaneous suppression of the liver transcriptional program by PDX1. We tested this hypothesis by transiently overexpressing PDX1 in HepG2 cells and by constitutively overexpressing PDX1 in hESCs and confirmed that PDX1 can indeed bind and repress a battery of hepatic genes. These results provide fundamental insight into how organ-specific transcriptional programs are established during embryonic development and recapitulate PDX1 as both a transcriptional activator and repressor.

RESULTS

Early Pancreatic Progenitor Cells from Differentiated hESCs

We previously reported the production of ePP cells from the differentiation of hESCs in adherent culture (Teo et al., 2012). These ePP cells were characterized by high levels of PDX1 expression on day 12, but quantitation by fluorescence-activated cell sorting (FACS) revealed that they numbered no more than 35% of the entire culture. We therefore explored other culture methodologies and platforms aimed at improving differentiation efficiency to ePP and discovered that PDX1+ cell numbers were increased substantially by initially plating hESC on fibronectin-coated transwell dishes and by extending retinoic acid (RA) treatment by 2 days and supplementing with FGF2, nicotinamide, and DAPT (FND) (see Figure 1A). On day 14, FND was replenished, and cultures were typically harvested on day 17 (Figure 1A). In this revised protocol, hESCs expectedly form a cobblestone-like lawn of DE cells by day 5 (Figure 1A). By day 10, distinct cell clusters emerge and shortly thereafter appear to undergo microlumen formation and fusion reminiscent of the tubulogenesis that occurs in vivo in the developing mouse pancreas (Figure 1A) (Kesavan et al., 2009; Villasenor et al., 2010). With continued differentiation, thickened ridges extend and intersect across the transwell in a honeycomb-like meshwork (Figure 1A).

To determine how faithfully ePP production in transwell culture followed the normal program of organogenesis, we used qRT-PCR to examine the expression of a panel of signature marker genes from days 0 to 17 of differentiation. These analyses document the transition from pluripotency to mesendoderm, which was characterized by the downregulation of pluripotency genes NANOG, OCT4, and SOX2 and the upregulation of BRACHYURY (T), GOOSECOID, MIXL1, and EOMES (Figure S1A). This event was followed shortly thereafter by upregulation of pan-DE (GATA4, CXCR4, SOX17, FOXA2, GATA6), ventral DE (HHEX), and gut tube (HNF1B) marker genes (Figures 1B and S1A). Between days 5 and 17, dorsal endoderm (ISL1, MNX1) and posterior foregut/pancreatic progenitor (SOX9, PDX1, HNF6, NKX6-2, HNF1A, SOX4) markers were increasingly expressed (Figures 1B and S1A). Western blot analyses on differentiated hESC from days 0 to 17 confirmed the transition from DE (SOX17, FOXA2) to pancreatic epithelial progenitors (SOX9, PDX1, E-CAD) (Figure S1B). Immunodetection analyses further showed abundant FOXA2-, HNF1B-, SOX9-, or E-CAD-positive cells on day 10 (Figure 1C). Two days later, significant numbers of SOX9;PDX1 double-positive cells were observed, consistent with the emergence of mitogenic multipotent ePP cells in vivo (Jennings et al., 2013; Seymour et al., 2007) (Figure S1C). Significant numbers of HNF1B-, SOX9-, PDX1-, or E-CAD-positive cells were maintained on day 17, particularly in the thickened ridges that crisscross the transwell (Figure 1C). Based on the expression profile of hESC-derived ePP cells (Figures 1B and S1A), we reasoned that these cells closely approximate the Pdx1+ dorsal and ventral pancreatic buds in E9.5 mouse embryos, at a stage prior to the robust activation of endocrine-commitment genes such as Ngn3 and Nkx6-1 (Ahlgren et al., 1996; Jennings et al., 2013; Jørgensen et al., 2007; Offield et al., 1996).

PDX1 Binds a Battery of Foregut/Midgut and Early Pancreatic Genes in hESC-Derived ePP Cells

PDX1 plays a preeminent, evolutionarily conserved role in orchestrating pancreatic morphogenesis, but surprisingly little is known about the identity of its transcriptional targets during embryonic development. We therefore combined high-affinity polyclonal PDX1 antibodies with chromatin immunoprecipitation and deep sequencing (ChIP-seq) in an effort to uncover those immediate downstream genes that govern the early growth and development of the human pancreatic anlagen. For these studies, we selected day 17 of differentiation—a time point that consistently yielded large numbers (≥65%) of PDX1+ ePP cells (Figure 1D). These analyses revealed 15,436 PDX1-bound regions that map to 6,212 genes (false discovery
Figure 1. Directed Differentiation of hESCs into Early Pancreatic Progenitors

(A) Schematic of 17-day pancreatic differentiation protocol. On day -2, HES3 cells are plated into fibronectin-coated transwell plates. Differentiation is initiated on day 0. Growth factors (activin A, BMP4, and FGF2) and small molecules (RA, Nic, and DAPT) were added at the indicated days (see the Experimental Procedures for additional details). The typical morphological changes that occur during differentiation are shown below the schematic. Scale bar represents 100 μm.

(legend continued on next page)
rate [FDR] < 0.1 with no distance cutoff; Table S1, part A). The PDX1/PBX1-complex homeodomain-binding motif was the most highly enriched among the sequence reads, followed by the FOXA1/FOXA2 forkhead/winged helix DNA-binding motif (Figure 2A). PBX1 binds to its half-site ATGATT, whereas PDX1/HOX binds to the half-site TTAAATGG, with an overlap at the middle TT (underlined), and these proteins heterodimerize to modulate gene transcription (Dutta et al., 2001; Knoepfler et al., 1996; Liu et al., 2001; Swift et al., 1998). These findings provide strong evidence that our ChIP-seq data are highly enriched for specific PDX1 binding events and further suggest that some PDX1-bound targets in ePP cells are co-regulated by PBX1 and FOXA proteins.

We next analyzed the binding profile of PDX1 and observed an increased binding frequency near the TSS of genes, suggesting that PDX1 plays a major role in transcriptional regulation (Figure 2B); 3,498 PDX1 binding sites are within ±20 kb of the TSS (FDR < 0.1), corresponding to 2,817 genes (Figure 2B; Table S1, part B). Among these genes were typically associated with gut endoderm, specifically the posterior foregut region where Pdx1 shows localized expression during early mouse pancreatic development. These genes include endoderm (CXCR4, LHX1, GATA4, GATA6), pancreas (MNX1, HNF1B, HNF1A, ONECUT1 (HNF6), PROX1, SOX9, PDX1, MEIS1, RFX6, PAX2, GLIS3, NR5A2, NKX2-2, SOX4, SOX5, KRT4, KRT18, KRT19, CPE, IAPP, SLC2A2, PLA2G1B), and numerous HOX (HOXA1, HOXA2, HOXA3, HOXA4, HOXA6, HOXB2, HOXC13, HOXD12) genes (Figures 2C and S2A; Table S1, part B). Genes typically associated with intestinal development and homeostasis (CDX2, LGR5, DPP4) were also identified (Figure S2A), which is consistent with the early expression of Pdx1 in the duodenum and its genetic

(B) Kinetics of endodermal (GATA4), gut tube (HNF1B), and pancreatic progenitor (SOX9, PDX1, HNF6) marker gene expression by qPCR over the course of 17 days of differentiation. Error bars represent the SD of three biological replicates.

(C) Immunostaining for FOXA2, HNF1B, E-CADHERIN, or PDX1 on days 10 or 17 of differentiation as indicated. Scale bar represents 100 μm.

(D) Quantitation of PDX1+ pancreatic progenitors by FACS on day 17.
requirement for normal rostral duodenum patterning (Ahlgren et al., 1996; Guz et al., 1995; Jennings et al., 2013; Offield et al., 1996). ChIP-qPCR analyses independently confirmed several of these binding events (Figures 2D and S2D). Expression analysis also revealed that HOXA4 and CDX2 levels continuously increased from day 10 of differentiation onward, paralleling PDX1, while DPP4 peaked on day 10/12 and then declined (Figures 1B, 2C, and S2E).

DPP4 peaked on day 10/12 and then declined (Figures 1B, 2C, and S2E). Genomic Regions Enrichment of Annotations Tool (GREAT) (http://great.stanford.edu/public/html/index.php) analysis refined the functional significance of PDX1-bound regions to GO terms such as “endocrine pancreas development” and “columnar/cuboidal epithelial cell differentiation” with the corresponding pathway designated as “regulation of gene expression in early pancreatic precursor cells” (Figure S2C). PDX1-bound targets were also generally associated with genes expressed in the foregut/midgut region when compared against the Mouse Genome Informatics (MGI) database (Figure S2C).

Together, these bioinformatics approaches provided further confidence that PDX1-bound targets in day 17 ePP cells are developmentally relevant to the formation of the human pancreas.

Candidate PDX1 Transcriptional Targets through Comparative Microarray Analysis

Microarray analyses performed in quadruplicate on days 0, 10, and 17 of differentiation revealed that these selected expression signatures (Figure 3A; Table S2, parts A–C). Consistent with our qPCR studies, numerous pancreas-related genes were upregulated by day 17, including TM4SF4, MAFB, CDH1 (E-CAD), SPINK1, PDX1, PROX1, CDH1 (E-CAD), SPINK1, ITM2B, PARM1, RXF6, PCSK2, HOXA2, and HNF1A, whereas several others such as PITX2, MEIS2, DACH1, ONECUT2, JAG1, FOXA2, MEIS1, ISL1, SOX9, HNF1B, KRT19, GATA4, GATA6, and MNX1 were expressed from day 10 onward (Figure 3B). Pluripotency genes NANOG, OCT4, and SOX2 were predictably downregulated by day 10, while mesendoderm (EOMES) and DE (SOX17) genes were downregulated on day 17 (Figure 3B). The slight increase in SOX2 on day 17 was consistent with the expression of Sox2 in the mouse anterior foregut (Que et al., 2007) (Figure 3B).

Next, we compared the list of PDX-bound genes to the day 17 microarray results to identify genes that are potentially directly regulated by PDX1. Among the 2,817 genes bound by PDX1 (+/−20 kb from the TSS), 357 were upregulated, whereas 102 were downregulated on day 17 as compared with day 0 (Figure 3C; parts D and E). Autoregulation of PDX1 is well documented, and PDX1 was thus as expected among these 357 genes (Table S2, part D) (Chakrabarti et al., 2002; Marshak et al., 2000). To home in on key PDX1 targets, we took advantage of two previously published data sets that attempted to identify Pdx1-regulated genes in cells or tissue isolated directly from the developing mouse embryo. In the first, Pdx-eGFP+ cells were isolated by cell sorting from transgenic Pdx1-eGFP E10.5 mouse embryos, and their gene expression profile determined with Affymetrix microarrays (Gu et al., 2004). From 56 enriched mouse genes thus identified, only 15 human homologs are bound by PDX1 (Figure 3D; Table S3, part A). These include the DE marker IGFBP5 and pancreatic regulator MEIS1, as well as a series of genes poorly characterized for their role in pancreas growth and development (NR2F1 (COUP-TF1), TNAIP8, TANCI, and WSBI (Figure 3D), all of which show increased expression on day 17 (Tamplin et al., 2008). Consistent with its expression in ePP cells (Pan and Wright, 2011; Shih et al., 2013), Onecut1 (Hnf6) was among the 56 upregulated genes in Pdx1eGFP+ sorted cells reported by Gu et al. (2004). Our results showed that ONECUT1 expression increased during ePP differentiation (Figure 1B) and that the ONECUT1 locus was bound by PDX1 (Figures 2C and 2D).

In a second study, comparative microarrays were performed on microdissected WT and Pdx1−/− E10.5 dorsal pancreatic buds (Svensson et al., 2007). (Beyond this stage, the unbranched Pdx1−/− pancreatic buds grow arrest and regress [Ahlgren et al., 1996; Offield et al., 1996].) From 73 genes downregulated in Pdx1 null mutant buds, 20 human homologs were bound by PDX1 (Figure 3E; Table S3, part B). Among these 20, 12 showed increased expression on day 17 and include known endoderm and pancreas markers (WFDC2, CDH1 (E-CAD), ANXA4, ITM2B, PARM1), as well as a diverse group of genes whose functions in pancreas morphogenesis are unknown (POSTN (PERIOSTIN), SORT1, SPINT2, SERPINB1, and DYNLT3) (Figure 3E) (Hoffman et al., 2008; Tamplin et al., 2008). Taken together, these two comparative studies underscore the developmental significance of PDX1-bound target genes revealed by our in vitro studies of differentiating hESCs.

PDX1 is also expressed in adult β cells, and heterozygous mutations in PDX1 are linked to autosomal dominant early onset type 2 diabetes (MODY4) (Ohlsson et al., 1993; Stoffers et al., 1997a). Moreover, in mice, Pdx1 is genetically required for the maintenance of mature β cell function (Ahlgren et al., 1998; Gannon et al., 2008). We therefore asked whether there are common PDX1 gene targets in day 17 ePP cells and in a recently published study of human
islets harvested from nondiabetic donors (Khoo et al., 2012). We found 1,096 PDX1-bound genes shared between the two tissue types (Figure S3A; Table S3, part C) that GREAT classifies into biological processes such as “pancreas development,” “type B pancreatic cell differentiation,” and “endoderm development” (Figure S3C). We next identified the mouse homologs of the 2,817 human genes bound by PDX1 (±20 kb from TSS) (Figure 3C) and overlapped these with a more limited ChIP-on-Chip study using the NIT-1 mouse insulinoma cell line (Figure S3B; Table S3, part D) (Keller et al., 2007). Consistent with a β cell origin of this cell line, among the 99 shared genes were known β cell regulators, including MEIS2, NKX2-2, PDX1, TLE1, and TLE3 (Table S3, part D). Taken together, these findings suggest that there is substantial overlap between PDX1 transcriptional targets during development and adult β cell homeostasis.

PDX1 Binds and Represses Hepatic Genes

In the developing mouse embryo, both explant culture experiments and comprehensive lineage tracing show that the liver and ventral pancreas arise from bipotent, and possibly multipotent, precursors in the foregut endoderm (Angelo et al., 2012; Deutsch et al., 2001; Miki et al., 2012; Tremblay and Zaret, 2005). Consistent with this tight lineage relationship, closer inspection of our microarray data revealed companion upregulation of early liver lineage genes, including APOA2, TBX3 FOXA1, TTR, AFP, APOB,
nests of AFP+ cells often abutting or near PDX1 labeled cell populations on day 17 of differentiation, with small 2012). Consistent with this, PDX1 and AFP labeled distinct AIP in the early-somite-stage mouse embryo (Miki et al., 1996). Recent high-resolution immunohistochemistry and whole-mount in situ hybridization further demonstrate that Pdx1 and Afp do not co-localize in the pancreatic buds, the caudal stomach, bile duct, and rostral duodenum (Jonsson et al., 1994; Jørgensen et al., 2007; Offield et al., 1996). Recent high-resolution immunohistochemistry and whole-mount in situ hybridization further demonstrate that Pdx1 and Afp do not co-localize in the AIP in the early-somite-stage mouse embryo (Miki et al., 2012). Consistent with this, PDX1 and AFP labeled distinct cell populations on day 17 of differentiation, with small nests of AFP+ cells often abutting or near PDX1 labeled ridges (Figure 4C). This result raises the possibility that a critical aspect of PDX1 function is to bind and repress hepatic genes in pancreatic progenitor (PP) cells, ensuring stable commitment to the pancreatic lineage.

To test this hypothesis, we first asked whether PDX1 overexpression could repress liver genes in the hepatocellular carcinoma cell line HepG2. We transiently overexpressed WT PDX1, or a mutant form of PDX1 (PDX1N196S) that cannot bind DNA (Horb et al., 2003) and observed suppression of endogenous AFP, ALB, TTR, APOA2, and FOXA1 gene expression by WT, but not PDX1N196S (Figure 4D). In support of direct repression by PDX1, we confirmed by ChIP-qPCR that, when overexpressed, WT PDX1 binds endogenous AFP, ALB, TTR, and FOXA1 regulatory sequences (Figure 4B; Table S1, part B). These results show that the DNA-binding activity of PDX1 is required for its ability to act as a strong transcriptional repressor when overexpressed in a terminally differentiated heterologous cell line.

We next asked whether PDX1 behaved similarly during hESC differentiation. On day 17, we observed that assorted liver marker genes were upregulated both by microarray and qPCR (Figures 4A and S4C). We therefore reasoned that constitutive overexpression of PDX1 in hESC may suppress or altogether eliminate the emergence of liver progenitors alongside ePP cells. Thus, we generated hESC lines that stably overexpressed either GFP or WT human PDX1. Importantly, overexpression of GFP or PDX1 neither impaired pluripotency nor the ability to sequentially activate endodermal or pancreatic marker genes (data not shown; see also Bernardo et al., 2009 and Lavon et al., 2006). Western analysis revealed that PDX1 almost entirely suppressed AFP protein expression (Figure 4E), corresponding to a nearly 90% reduction at the mRNA level by qPCR (Figure 4F). In addition, PDX1 overexpression strongly suppressed the expression of AFP, ALB, TTR, APOA2, and FOXA1 (Figure 4F). Taken together, these data suggest that PDX1 engages a robust commitment to the pancreatic lineage at least in part by binding the regulatory regions of liver genes and directly repressing their transcription.

**DISCUSSION**

Given its conspicuous expression in the early pancreatic anlagen, its ability to directly regulate the Insulin gene promoter and its dramatic loss-of-function phenotype, PDX1 is often designated as “master regulator” of pancreatic fate (Pan and Wright 2011; Shih et al., 2013). Consequently, numerous studies have attempted to use ectopic PDX1 to reprogram “permissive” tissues and assorted cell lines into the pancreatic lineage and, ultimately, functional beta cells for cell replacement therapy (Meivar-Levy and Ferber, 2010). For example, adenoviral approaches have overexpressed either WT PDX1 or an activated PDX1-VP16 fusion (C-terminal addition of the 80 amino acid transactivation activation domain of the Herpes simplex virus VP16 protein) in the adult mouse liver, an organ whose embryonic origin is in close proximity to the ventral pancreas (Angelo et al., 2012; Deutsch et al., 2001; Miki et al., 2012; Tremblay and Zaret, 2005). Upregulation of signature pancreatic genes and islet hormones was observed to varying degrees, and depending on the study, different criteria were taken as a sign of “function” in the resulting proto-beta cells, the most stringent being the ability to positively affect blood glucose levels (Ber et al., 2003; Ferber et al., 2000; Imai et al., 2005; Kaneto et al., 2005; Kojima et al., 2003). Curiously, these in vivo studies principally focused on the acquisition of a handful of markers of pancreatic identity, but not the degree of preservation or loss of hepatic fate.

PDX1 overexpression experiments provide some evidence that this factor can partially engage the pancreatic program while conversely suppressing endogenous liver genes. In studies with primary rodent hepatocytes or liver epithelial cells, Alb levels were variably downregulated by overexpressing PDX1 or PDX1-VP16 (Fodor et al., 2007; Jin et al., 2008; Meivar-Levy et al., 2007; Yamada et al., 2006). Li et al. (2005) further showed that HepG2 cells expressing TTR-XIiHbox8-VP16, encoding the *Xenopus* homolog of PDX1 (XIiHbox8) driven by the liver-specific TTR promoter, activated endocrine genes such as Insulin.
Figure 4. PDX1 Represses Liver Marker Genes

(A) Microarray gene expression heatmap showing increasing levels of liver lineage-related genes between days 0 and 17 of pancreatic differentiation.

(B) ChIP-qPCR provides independent confirmation that PDX1 binds in proximity to liver genes. Error bars indicate the SD of three technical replicates from a single ChIP-qPCR experiment. These data were independently confirmed in repeat pull-down qPCR experiments (data not shown). Normal goat IgG is the negative control.

(C) AFP and PDX1 immunostaining on day 17 of pancreatic differentiation. AFP and PDX1 label distinct cell populations that are often observed in close proximity. Scale bar represents 100 μm.

(D) Expression of hepatic genes (AFP, ALB, TTR, APOA2, FOXA1, FOXA2, FOXA3, and HHEX) by qPCR in HepG2 cells after transient over-expression of GFP, PDX1, or the PDX1(N196S) mutant that cannot bind DNA. Both WT PDX1 and PDX1(N196S) are robustly overexpressed after transfection. All error bars indicate SD of three biological replicates. For all genes, data are shown relative to the GFP-transfected controls. p values are indicated.

(E) Western blot analyses for PDX1 and AFP on days 0 and 17 of pancreatic differentiation in GFP- or human PDX1-overexpressing hESC clones.

(F) Expression of liver markers (AFP, ALB, TTR, APOA2, FOXA1, HHEX) in hESCs stably overexpressing either GFP or human PDX1 and differentiated for 17 days. Error bars indicate the SD of three biological replicates. p values were calculated when compared with GFP control.
and suppressed Alb, Transferin (Ttr), and α-1 anti-trypsin (Serpina1a) (Li et al., 2005). More akin to our PDX1 overexpression experiments in differentiating hESCs, Horb et al. (2003) reported that transgenic TTR-XIHubox8-VP16 Xenopus tadpoles showed some conversion of liver to pancreas, scored by reduction of Ttr and activation of Ins, Gcg and Amylase (Amy) (Horb et al., 2003). This in vivo transdifferentiation by XIHubox8-VP16 required the XIHubox8 DNA-binding homeodomain, but curiously WT XIHubox8 lacking the VP16 transactivation domain was entirely ineffective in ectopically activating pancreatic genes in the developing liver (Horb et al., 2003). In hESC-derived ePP cells, we demonstrated by ChIP-seq that WT PDX1 binds hepatic genes and formally established in HepG2 cells that such binding is necessary for the direct suppression of several cardinal liver markers (AFP, ALB, TTR, APOA2, and FOXA1). Given our ChIP-seq results, one prediction from work in the frog is that the XIHubox8-VP16 fusion protein should be recruited to and potently activate, rather than repress, endogenous liver genes. Such a result was not reported. We posit that the addition of the VP16 module so drastically alters WT XIHubox8 (Pdx1) that it is difficult to compare the results of Horb et al. (2003) with our own. In addition, the bulk of the studies employing XIHubox/Pdx1-VP16 fusions did not carefully study temporally the transdifferentiation of liver to pancreas by assessing the expression of Pdx1-bound hepatic markers (e.g., Afp, Ttr, Alb, or Foxa1) at multiple time points, and consequently, one cannot exclude the possibility, ever how unlikely in our opinion, that an upregulation of hepatic genes via XIHubox8 Pdx1 VP16 did occur, even transiently. Taking together these published studies and our own findings, we favor a model whereby Pdx1-directed “transdifferentiation” of liver to pancreas requires PDX1 to act as both a transcriptional activator and repressor.

Consistent with our in vitro data, two recent genetic studies in the mouse provide strong in vivo evidence that Pdx1 is a bona fide transcriptional repressor. By E11.5, the pancreatic buds in Pdx1 null mutant embryos arrest and begin to regress (Ahlgren et al., 1996; Offield et al., 1996). Recently, Seymour et al. (2012) performed high-resolution quantitative immunohistochemistry on Pdx1−/− embryos and observed significant numbers of ectopic Afp+ cells within the dorsal pancreatic bud. This suggestion of partial conversion to the hepatic cell fate is consistent with prior work demonstrating that hepatic competence is not restricted to the region of the ventral foregut where the liver normally forms (Bossard and Zaret, 1998, 2000; Gualdi et al., 1996) and with our data that PDX1 represses AFP in vitro (Figures 4D–4F). Seymour et al. (2012) also showed that Pdx1 deficiency caused varying degrees of Sox9 downregulation and that focal loss of Sox9 in the developing pancreas led to elevated expression of hepatic markers. Our results show that in hESC-derived ePP cells PDX1 binds SOX9 between exons 2 and 3 (Table S1, part A), which suggests that SOX9 is positively regulated by PDX1, as expected for these principal regulators of the pancreatic program. Phylogenetic sequence conservation in this region of mouse and human SOX9 genes (data not shown) fits the idea that this Pdx1-Sox9 regulatory relationship is central to the pro-pancreatic gene regulatory network. Taken together, these findings indicate that early-stage Pdx1+ progenitor cells are not stably determined (“metastable”), with Pdx1 positively regulating Sox9 and actively repressing liver potential during a substantial period of early pancreas organogenesis.

In a second and very recent study, Gao et al. (2014) inactivated Pdx1 in the adult β cell using Cre-Lox methods with concurrent indelible YFP labeling of the derived Pdx1−/− cells (Gao et al., 2014). Expectedly, these mice became rapidly hyperglycemic—a result consistent with prior work (Ahlgren et al., 1998; Gannon et al., 2008)—but unexpectedly lineage-labeled Pdx1−/− cells contained glucagon and expressed MafB, a transcription factor that in the adult mouse is restricted to the islet α-cells. These authors used ChIP from mouse insulinoma cell lines to detect PDX1 binding within 1.5 kb of the MafB TSS. Taken together, these findings strongly suggest that Pdx1 directly represses MafB transcription in adult β cells. Interestingly, MafB is required for the production of both α and β cells during pancreas development, but its expression is extinguished in β cells soon after birth (Hang and Stein, 2011). Consistent with this in vivo expression kinetic, we observed MAFB levels increasing from days 0 to 17 of hESC differentiation (Figure 3B). However, MAFB transcriptional regulation is apparently independent of PDX1 at these stages, as PDX1 binding was observed a great distance from the MAFB TSS (≥ 150 kb) (Table S1, part A). In addition, it is important to highlight recent data showing that, in contrast to mice, MAFB persists in a subset of human adult β cells (~9%) (Dai et al., 2012). This finding suggests that in mice Pdx1 adopts its role as a transcription repressor late in β cell ontogeny and that in humans PDX1+MABF− and PDX1+MABF− represent distinct β cell subtypes.

Our findings raise an important outstanding question: what is the mechanism underlying PDX1 transcriptional repression? We speculate that one answer lies in the top-ranking motifs enriched in our ChIP-seq data—PBX1 and FOXA1/A2. Nearly 20 years ago, Pdx1 was shown to bind the HOX-cofactor Pbx1 (pre-B cell leukemia factor 1), a member of the TALE (three-amino-acid loop extension) family of atypical homeodomain-containing proteins (Peers et al., 1995). Pbx1 alternative splicing yields two isoforms differing at their C termini, the longer Pbx1a and shorter Pbx1b. Pdx1:Pbx1b complexes transcriptionally activate
target genes, while Pdx1:Px1a forms a repressor complex through the recruitment of co-repressor proteins such as NCoR-SMRT (or HDAC) (Asahara et al., 1999; Saleh et al., 2000). PBX1A and PBX1B are both expressed during human ePP differentiation (Figure S1A; A.K.K.T. and N.R.D., unpublished data), raising the possibility of differential recruitment within the same cell to activate or repress appropriate gene targets to direct lineage choice and stabilization. Similarly, FOXA transcription factors can also recruit HDAC via the co-repressor Groucho-related protein 3 (Grg3; formally Tle3), which is highly expressed in the pancreas during embryonic development, to silence genes central to hepatic differentiation (Lam et al., 2013; Santisteban et al., 2010). Our data also show that FOXA1/A2 binding sites were significantly enriched in sequence reads from day 17 PDX1 ChIP-seq (Figure 2A), and FOXA2:PDX1 co-binding was observed frequently in nearly 2,000 loci in mouse islets (Hoffman et al., 2010). Canonical FOXA1/A2 motifs exist close to PDX1 binding in both AFP and TTR, also suggesting context-dependent functions of FOXA proteins. Finally, it is important to note that approximately 100 genes are bound by PDX1 on day 17, but their expression is significantly downregulated by microarray (comparing day 17 to day 0) (Figure 3C; Table S2, part E). Obvious candidates for direct repression among these ~100 include FGFB, TWIST2, ETV4, and ZIC3, whose mouse orthologs are typically expressed in mesoderm or mesodermally derived tissues during embryonic development (http://www.informatics.jax.org/genes.shtml) (Table S2, part E). Methods to address the issue of direct repression or activation of PDX1 target loci include the development of PDX1-deficient hESC lines or lines carrying inducible knockdown tools for context and time-dependent inactivation.

EXPERIMENTAL PROCEDURES

Cell Culture

HES3 (ESO3; http://www.nationalstemcellbank.org) were cultured on γ-irradiated human Ortec feeders in hESC medium supplemented with 50-ng/ml FGF2 or in mTeSR 1 medium (STEMCELL Technologies) (Teo et al., 2012). Colonies were passaged every 7 days by treating with 1-ng/ml collagenase IV followed by fine scoring with a pipet tip. HepG2 cells (ATCC) were cultured in DMEM high glucose (Invitrogen), 10% fetal bovine serum (FBS), and 2-mM L-glutamine.

Transfection

Human PDX1 cDNA was cloned into pTP6. The N196S point mutation was introduced using the Directed Mutagenesis kit (Agilent). Transfection of HepG2 and HES3 cells was performed with Lipofectamine 2000 (Invitrogen); 4 μg of pTP6-hPDX1 and pTP6-hPDX1N196S were transfected into HepG2 cells for 48 hr followed by selection in puromycin (2 μg/ml) for 6 days. Similarly, 4 μg of pTP6-hPDX1 was transfected into HES3 cells and 1 μg/ml of puromycin was included in the hESC medium for 1 week to select for hPDX1-overexpressing HES3 lines.

Pancreatic Differentiation

Feeder cells were depleted using a 70-μm cell strainer (BD Falcon) prior to directed differentiation (Teo et al., 2012). hESCs were plated onto fibronectin-coated transwell plates (Corning) and cultured in hESC medium for 2 days. They were induced to differentiate in RPMI-1640, 2% B-27, 2-mM L-glutamine, 55-μM β-mercaptoethanol, and 1 × MEM NEAA medium supplemented with 50-ng/ml activin A and 50-ng/ml BMP4 for 3 days followed by 2 additional days in 50 ng/ml activin A (schematized in Figure 1A). Cultures were then exposed to 50-ng/ml FGF2, 3-μM all-trans-RA, and 10-mM nicotinamide (Nic) for 5 days; 50-ng/ml FGF2, 3-μM RA, 10-mM Nic, and 20-μM DAPT for 4 days; and 50-ng/ml FGF2, 10-mM Nic, and 20-μM DAPT for 3 days.

qPCR, Western Blot, Immunostaining, FACS, ChIP-Seq, and Microarray Hybridization and Data Analyses

Methods for qPCR, western blot, immunostaining, FACS, ChIP-seq, and microarray analyses were previously described (Teo et al., 2012; Tsuneyoshi et al., 2012). More detailed methods for ChIP-seq, microarray hybridization, and data analysis are provided in the Supplemental Experimental Procedures. Primers and antibodies are listed in Table S4.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2015.02.015.

AUTHOR CONTRIBUTIONS

A.K.K.T. and N.R.D. designed the study and analyzed the data. A.K.K.T. performed hESC differentiation, qPCR, ChIP-seq, ChIP-qPCR, western blots, FACS, and comparative microarray studies. S.H. and L.W.S. provided bioinformatics support and access to deep sequencing. N.T. performed immunofluorescence, provided additional bioinformatics support, and designed all figures. E.K.T. performed subcloning, construct mutagenesis, qPCR, and transfections. C.V.E.W. provided key reagents, including the acid-purified goat anti-Pdx1 antibody used for all ChIP experiments and contributed to data analysis and interpretation. N.R.D. wrote the paper.

ACKNOWLEDGMENTS

The authors thank Sarah Chan and Jonathan Scolnick for technical assistance with deep sequencing and Yusuf Ali, Hiram Chipperfield, Rory Johnson, and Pauline Tay for their early work on the pancreatic differentiation protocol and PDX1 chromatin immunoprecipitation. This work is supported by A*STAR, Singapore.
REFERENCES

Ahlgren, U., Jonsson, J., and Edlund, H. (1996). The morphogenesis of the pancreatic mesenchyme is uncoupled from that of the pancreatic epithelium in IPF1/PDX1-deficient mice. Development 122, 1409–1416.

Ahlgren, U., Jonsson, J., Jonsson, L., Simu, K., and Edlund, H. (1998). beta-cell-specific inactivation of the mouse Ip1/Pdx1 gene results in loss of the beta-cell phenotype and maturity onset diabetes. Genes Dev. 12, 1763–1768.

Angelo, J.R., Guerrero-Zayas, M.I., and Tremblay, K.D. (2012). A gene therapy approach. Diabetologia 55, 121–130.

Gannon, M., Herrera, P.L., and Wright, C.V. (2000). Mosaic Cre-mediated recombination in pancreas using the pdx-1 enhancer/promoter. Genesis 26, 143–144.

Gannon, M., Ables, E.T., Crawford, L., Lowe, D., Offield, M.F., Magnuson, M.A., and Wright, C.V. (2008). pdx-1 function is specifically required in embryonic beta cells to generate appropriate numbers of endocrine cell types and maintain glucose homeostasis. Dev. Biol. 314, 406–417.

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, E., Aronow, B., and Melton, D.A. (2004). Global expression analysis of gene regulatory pathways during endocrine pancreatic development. Development 131, 165–179.

Gualdi, R., Bossard, P., Zheng, M., Hamada, Y., Coleman, J.R., and Zaret, K.S. (1996). Hepatic specification of the gut endoderm in vitro: cell signaling and transcriptional control. Genes Dev. 10, 1670–1682.

Guz, Y., Montminy, M.R., Stein, R., Leonard, J., Gamer, L.W., Wright, C.V., and Teitelman, G. (1995). Expression of murine STF-1, a putative insulin gene transcription factor, in beta cells of pancreas, duodenal epithelium and pancreatic exocrine and endocrine progenitors during ontogeny. Development 121, 11–18.

Hang, Y., and Stein, R. (2011). MafA and MafB activity in pancreatic beta cells. Trends Endocrinol. Metab. 22, 364–373.

Hoffman, B.G., Zavaglia, B., Witzsche, J., Ruiz de Alcaraz, T., Beach, M., Hoodless, P.A., Jones, S.J., Marra, M.A., and Helgason, C.D. (2008). Identification of transcripts with enriched expression in the developing and adult pancreas. Genome Biol. 9, R99.

Hoffman, B.G., Robertson, G., Zavaglia, B., Beach, M., Cullum, R., Lee, S., Soukhatcheva, G., Li, I., Wederell, E.D., Thiessen, N., et al. (2010). Locus co-occupancy, nucleosome positioning, and H3K4me1 regulate the functionality of FOXA2-, HNF4A-, and PDX1-bound loci in islets and liver. Genome Res. 20, 1037–1051.

Horb, M.E., Shen, C.N., Tosh, D., and Slack, J.M. (2003). Experimental conversion of liver to pancreas. Curr. Biol. 13, 105–115.

Imai, J., Katagiri, H., Yamada, T., Ishigaki, Y., Ogihara, T., Uno, K., Hasegawa, Y., Gao, J., Ishihara, H., Sasano, H., et al. (2005). Constitutively active PDX1 induced efficient insulin production in adult murine liver. Biochem. Biophys. Res. Commun. 326, 402–409.

Jennings, R.E., Berry, A.A., Kirkwood-Wilson, R., Roberts, N.A., Hearn, T., Salisbury, R.J., Blaylock, J., Piper Hanley, K., and Hanley, N.A. (2013). Development of the human pancreas from foregut to endocrine commitment. Diabetes 62, 3514–3522.

Jin, C.X., Li, W.L., Xu, F., Geng, Z.H., He, Z.Y., Su, J., Tao, X.R., Ding, X.Y., Wang, X., and Hu, Y.P. (2008). Conversion of immortal liver progenitor cells into pancreatic endocrine progenitor cells by persistent expression of Pdx-1. J. Cell. Biochem. 104, 224–236.
Jönsson, J., Carlsson, L., Edlund, T., and Edlund, H. (1994). Insulin-promoter-factor 1 is required for pancreas development in mice. Nature 371, 606–609.

Jørgensen, M.C., Ahnhelt-Rønne, J., Hald, J., Madsen, O.D., Serup, P., and Hecksher-Sørensen, J. (2007). An illustrated review of early pancreas development in the mouse. Endocr. Rev. 28, 685–705.

Kaneto, H., Nakatani, Y., Miyatsubu, T., Matsuo, T.A., Matsuhisa, M., Hori, M., and Yamasaki, Y. (2005). PDX-1/VP16 fusion protein, together with NeuroD or Ngn3, markedly induces insulin gene transcription and ameliorates glucose tolerance. Diabetes 54, 1009–1022.

Keller, D.M., McWeeny, S., Arsenlis, A., Drouin, J., Wright, C.V., Wang, H., Wollheim, C.B., White, P., Kaestner, K.H., and Goodman, R.H. (2007). Characterization of pancreatic transcription factor Pdx-1 binding sites using promoter microarray and serial analysis of chromatin occupancy. J. Biol. Chem. 282, 32084–32092.

Kesavan, G., Sand, F.W., Greiner, T.U., Johansson, J.K., Kobberup, S., Wu, X., Brakebusch, C., and Semb, H. (2009). Cdc42-mediated tubulogenesis controls cell specification. Cell 139, 791–801.

Khoo, C., Yang, J., Weinrott, S.A., Kaestner, K.H., Naji, A., Schug, J., and Stoffers, D.A. (2012). Research resource: the pdx1 cistrome of pancreatic islets. Mol. Endocrinol. 26, 521–533.

Knoepfler, P.S., Lu, Q., and Kamps, M.P. (1996). Pbx-1 Hox heterodimers bind DNA on inseparable half-sites that permit intrinsic DNA binding specificity of the Hox partner at nucleotides 3’ to a TAAT motif. Nucleic Acids Res. 24, 2288–2294.

Kojima, H., Fujiyama, M., Matsumura, K., Younan, P., Imaeda, H., Maeda, M., and Chan, L. (2006). NeuroD-betacellulin gene therapy regulated genes in pancreatic progenitor cells. BMC Dev. Biol. 6, 636–648.

Koizumi, T., Su, L., Chien, Y., and Johnson, L. (1999). The pancreatic islet factor STF-1 binds cooperatively with Pbx to a regulatory element in the somatostatin promoter: importance of the FPWMK motif and of the homeodomain. Mol. Cell. Biol. 15, 7091–7097.

Krejci, J., Okubo, T., Goldenring, J.R., Nam, K.T., Kurotani, R., Morrissey, E.E., Tananaka, O., Pevny, L.H., and Hogan, B.L. (2007). Multiple dose-dependent roles for Sox2 in the patterning and differentiation of anterior foregut endoderm. Development 134, 2521–2531.

Li, W.C., Horb, M.E., Tosh, D., and Slack, J.M. (2005). In vitro transdifferentiation of hepatoma cells into functional pancreatic cells. Mech. Dev. 122, 835–847.

Liu, Y., MacDonald, R.J., and Swift, G.H. (2001). DNA binding and transcriptional activation by a PDX1.PBX1b.MEIS2b trimer and cooperation with a pancreas-specific basic helix-loop-helix complex. J. Biol. Chem. 276, 17985–17993.

Marshak, S., Benshushan, E., Shoshkes, M., Havin, L., Cerasi, E., and Melloul, D. (2000). Functional conservation of regulatory elements in the pdx-1 gene: PDX-1 and hepatocyte nuclear factor 3beta transcription factors mediate beta-cell-specific expression. Mol. Cell. Biol. 20, 7583–7590.

Meivar-Levy, I., and Ferber, S. (2010). Adult cell fate reprogramming: converting liver to pancreas. Methods Mol. Biol. 636, 251–283.

Meivar-Levy, I., Sapr, T., Gefen-Halevi, S., Aviv, V., Barshack, I., Onaca, N., Mor, E., and Ferber, S. (2007). Pancreatic and duodenal homeobox gene 1 induces hepatic dedifferentiation by suppressing the expression of CCAAT/enhancer-binding protein beta. Hepatology 46, 898–905.
Swift, G.H., Liu, Y., Rose, S.D., Bischof, L.J., Steelman, S., Buchberg, A.M., Wright, C.V., and MacDonald, R.J. (1998). An endocrine-exocrine switch in the activity of the pancreatic homeodomain protein PDX1 through formation of a trimeric complex with PBX1b and MRG1 (MEIS2). Mol. Cell. Biol. 18, 5109–5120.

Tamplin, O.J., Kinzel, D., Cox, B.J., Bell, C.E., Rossant, J., and Lickert, H. (2008). Microarray analysis of Foxa2 mutant mouse embryos reveals novel gene expression and inductive roles for the gastrula organizer and its derivatives. BMC Genomics 9, S11.

Teo, A.K.K., Ali, Y., Wong, K.Y., Chipperfield, H., Sadasivam, A., Poobalan, Y., Tan, E.K., Wang, S.T., Abraham, S., Tsuneyoshi, N., et al. (2012). Activin and BMP4 synergistically promote formation of definitive endoderm in human embryonic stem cells. Stem Cells 30, 631–642.

Tremblay, K.D., and Zaret, K.S. (2005). Distinct populations of endoderm cells converge to generate the embryonic liver bud and ventral foregut tissues. Dev. Biol. 280, 87–99.

Tsuneyoshi, N., Tan, E.K., Sadasivam, A., Poobalan, Y., Sumi, T., Nakatsuji, N., Suemori, H., and Dunn, N.R. (2012). The SMAD2/3 corepressor SNON maintains pluripotency through selective repression of mesendodermal genes in human ES cells. Genes Dev. 26, 2471–2476.

Villasenor, A., Chong, D.C., Henkemeyer, M., and Cleaver, O. (2010). Epithelial dynamics of pancreatic branching morphogenesis. Development 137, 4295–4305.

Yamada, S., Yamamoto, Y., Nagasawa, M., Hara, A., Kodera, T., and Kojima, I. (2006). In vitro transdifferentiation of mature hepatocytes into insulin-producing cells. Endocr. J. 53, 789–795.
PDX1 Binds and Represses Hepatic Genes
to Ensure Robust Pancreatic Commitment
in Differentiating Human Embryonic Stem Cells

Adrian Kee Keong Teo, Norihiro Tsuneyoshi, Shawn Hoon, Ee Kim Tan, Lawrence W. Stanton, Christopher V.E. Wright, and N. Ray Dunn
Supplementary Materials

Supplemental Experimental Procedures

Chromatin Immunoprecipitation-sequencing (ChIP-seq)

ChIP-seq was carried out as previously described (Teo et al. 2011), without extra cross-linkers (XL) (Brown et al. 2011). Two 6-well transwell plates of day 17 differentiated hESC were sonicated for 2 min x 90% pulse x 5 times using an Ultrasonic 3000 Ultrasonic homogenizer (Biologics, Inc). ChIP-QPCR validation was carried out as described previously (Teo et al. 2011). Results were expressed as fold enrichment over control region (not shown, PDX1 control region) and comparisons carried out between PDX1 antibody pull-down and IgG control. Primers and antibodies are listed in Supplementary Table 4.

ChIP-seq libraries were prepared using the Solexa ChIP-seq DNA Sample Prep Kit (IP102-1001). After amplification, DNA fragments between 180 and 400 bp were gel extracted and sent for sequencing. Libraries were sequenced using the Illumina Genome Analyzer II as single-end 35 bp reads. Raw data have been uploaded to GEO with accession number GSE58686.

Peak annotation

Sequence reads were mapped to the human reference genome (hg19) using the bowtie program version 0.12.7 with settings (-m 1 --chunkmbs 256 --best -p 8 -S hg19) (Langmead et al. 2009). Aligned reads were analyzed by the MACS peak finder program with default settings (Zhang et al. 2008). 15,436 peaks (FDR 10% cutoff) were identified that correspond to putative PDX1-bound sites. Only reads that uniquely mapped to the human genome were used. In the initial gene annotation, PDX1-bound regions were associated with the nearest Ensembl transcript (no distance cutoff). In subsequent comparisons with microarray or published data, bound regions were associated to EntrezGene IDs within 20 kb of the TSS of genes.

Motif and Gene Ontology (GO) analyses

De novo motif analysis was performed using the HOMER software (findMotifsGenome.pl) with settings -size 200 -len 8,10,12 -S 50 (Heinz et al. 2010). Peaks were annotated using the annotatePeaks.pl (with setting -size 500) script from the HOMER package to assign peaks to the nearest TSS. Using a cutoff distance of 20 kb, 2,817 genes were located near a PDX1-bound peak. The same script was used to perform Gene Ontology enrichment analysis. The Genomic Regions Enrichment of Annotations Tool (GREAT) was also used to annotate peaks (McLean et al. 2010). All 15,436 peak locations (FDR 10% cutoff) were used in this analysis with default settings.

Analyses of overlap in gene sets

To overlap PDX1-regulated genes identified in published mouse studies (Gu et al. 2004; Keller et al. 2007; Svensson et al. 2007; Khoo et al. 2012), we used ENSEMBL biomart to map mouse IDs to their human orthologs (Flicek et al. 2008).
Microarray hybridization and data analysis

Four biological replicates of total RNA extracted from day 0 HES3 cells, day 10 and day 17 differentiated HES3 cells were subjected to reverse transcription and microarray hybridization. RNA samples were amplified and biotin-labeled according to manufacturer’s recommendations (Illumina TotalPrep-96 RNA Amplification Kit). Amplified cRNA was hybridized to the Human HT-12 v4 Beadchip (Illumina) for 16 h. The hybridized beadchip was washed, Cy3-stained the next day and scanned by Illumina BeadArray Reader at Scan Factor 1.5, PMT 504 after drying. The expression signal of the scanned results was summarized to gene level from multiple probes if any for each sample. The data were then imported into Partek® Genomics Suite™ for analysis. Changes in global gene expression between day 0 and day 10, and between day 0 and day 17 were compared, and a list of genes with significant confidence values (FDR < 0.05) was generated. All hybridizations have been uploaded to the Gene Expression Omnibus (GEO) data repository (NCBI, http://www.ncbi.nlm.nih.gov/geo/). The accession number for this study is GSE58686.

Data visualization

For data visualisation, heat maps of gene expression were created by importing relevant subsets of Robust Multichip Average (RMA) processed microarray gene expression data into Partek® Genomics Suite™. For each gene (row), the heat map colours depict gene expression in units of standard deviation from the mean across all samples (columns).

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure 1. Data in support of Figure 1. (A) Expression of pluripotency (NANOG, OCT4 and SOX2), mesendodermal (T/BRACHYURY, GSC, MIXL1 and EOMES), DE (CXCR4, SOX17, FOXA2 and GATA6), ventral endoderm (HHEX), dorsal endoderm (ISL1, MNX1) and pancreatic progenitor (PBX1, NKX6-2, HNF1A, SOX4) marker genes by QPCR during pancreatic differentiation over 17 days. Error bars indicate standard deviation of three biological replicates. (B) Western blot analyses for DE (SOX17, FOXA2) and pancreatic progenitor (SOX9, PDX1, E-CADHERIN) markers during pancreatic differentiation over 17 days. (C) Double SOX9 and PDX1 immunofluorescence on day 12 of differentiation. SOX9;PDX1+ cells increase in number with extended differentiation (see also Teo et al., 2012).

Supplementary Figure 2. Data in Support of Figure 2. (A) Representative views of PDX1 binding within the PBX1, MEIS1, KRT4, LGR5, CDX2 and DPP4 genomic loci. (B) GO analysis of PDX1-bound genes on day 17 of pancreatic differentiation using PARTEK software. The GO term is indicated along the y-axis and the P value for the significance of enrichment for the top 20 GO terms along the x-axis. (C) GO biological process, pathway commons and MGI expression analyses were carried out on PDX1-bound genes on day 17 of pancreatic differentiation using the GREAT software tool. P and Q values are as indicated in the tables. (D) ChIP-QPCR confirms the presence of PDX1 on regulatory regions of selected intestinal genes. Error bars indicate the standard deviation of three technical replicates from a single ChIP-QPCR experiment. These data were independently confirmed in repeat pull-down QPCR experiments (data not shown). Normal goat IgG as negative control. (E) Expression of HOX/duodenal (HOXA4) and intestinal (DPP4, CDX2) marker genes during pancreatic differentiation by QPCR. Error bars indicate standard deviation of three biological replicates.
Supplementary Figure 3. Comparisons of PDX1 binding on day 17 of hESC differentiation and previously characterized PDX1-bound target genes in human islets and NIT-1.  (A) Previously published human islet ChIP-Seq yielded 15,000 PDX1 peaks (Khoo et al., 2012), which correspond to 3,687 genes, 1,096 of which overlap with PDX1 peak data from day 17 of pancreatic differentiation. (B) Common PDX1-target genes between day 17 PDX1 ChIP-Seq (2,817 within ± 20 kb from the TSS) and a limited, previously published ChIP-on-Chip study using the insulin-secreting, NIT-1 mouse insulinoma cell line (Keller et al., 2007). (C) GO biological process and MGI expression analyses were carried out on the 1,096 common genes in (A) using the GREAT software tool. P and Q values are as indicated in the tables.

Supplementary Figure 4. Data in support of Figure 4.  (A) Representative views of PDX1 binding within FOXA1, ALB, HHEX, TTR and APOA2 genomic loci. (B) PDX1 binds endogenous liver genes (FOXA1, AFP, TTR and ALB) in HepG2 cells. Cells were transfected using Lipofectamine™ 3000 (Invitrogen) with 10 µg pTP6-hPDX1 and harvested 48 hours later for ChIP. Relative occupancy values are shown as the apparent immunoprecipitation efficiency (percentage) (ratio = immunoprecipitated DNA/inputDNA). Normal goat IgG as negative control. Error bars indicate the standard deviation of three technical replicates from a single ChIP-QPCR experiment. These data were independently confirmed in repeat pull-down QPCR experiments (data not shown). (C) Expression of hepatic markers (FOXA1, FOXA3, HNF4A, AFP, ALB, TTR and APOA2) over 17 days of differentiation. Error bars indicate standard deviation of three biological replicates.

SUPPLEMENTARY TABLE LEGENDS

Supplementary Table 1 (Supports Figures 2 and 3). PDX1 ChIP-seq. (A) Genomic regions bound by PDX1. (B) PDX1-bound targets within ± 20 kb from the TSS of genes.

Supplementary Table 2 (Supports Figure 3). (A,B) Genes up- and downregulated on day 10 (A) and 17 (B) of pancreatic differentiation as compared to day 0. (C) Genes up- and downregulated on day 17 as compared to day 10. (D) PDX1-bound genes (± 20 kb from the TSS of genes) that are upregulated on day 17 as compared to day 0 (> 2 fold; FDR < 0.05). (E) PDX1-bound genes (± 20 kb from the TSS of genes) that are downregulated on day 17 as compared to day 0 (> 2 fold; FDR < 0.05). (F) PDX1-bound genes (± 20 kb from the TSS of genes) that are upregulated on day 17 as compared to day 10 (> 2 fold; FDR < 0.05). (G) PDX1-bound genes (± 20 kb from the TSS of genes) that are downregulated on day 17 as compared to day 10 (> 2 fold; FDR < 0.05).

Supplementary Table 3. (A) Candidate PDX1-target genes common between day 17 PDX1 ChIP-seq and Gu et al. (2004) dataset. (B) Candidate PDX1-target genes common between day 17 PDX1 ChIP-seq and Svensson et al. (2007) dataset. (C) PDX1-bound targets common between day 17 PDX1 ChIP-seq and Khoo et al. (2012) dataset. (D) Candidate PDX1-target genes common between day 17 PDX1 ChIP-seq and Keller et al. (2007) dataset.

Supplementary Table 4. Primer sequences and antibodies.
References for Supplementary Materials

Brown S, Teo A, Pauklin S, Hannan N, Cho CH, Lim B, Vardy L, Dunn NR, Trotter M, Pedersen R et al. 2011. Activin/Nodal signaling controls divergent transcriptional networks in human embryonic stem cells and in endoderm progenitors. *Stem Cells* **29**(8): 1176-1185.

Flicek P, Aken BL, Beal K, Ballester B, Caccamo M, Chen Y, Clarke L, Coates G, Cunningham F, Cutts T et al. 2008. Ensembl 2008. *Nucleic Acids Res* **36**(Database issue): D707-714.

Gu G, Wells JM, Dombkowski D, Preffer F, Aronow B, Melton DA. 2004. Global expression analysis of gene regulatory pathways during endocrine pancreatic development. *Development* **131**(1): 165-179.

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular cell* **38**(4): 576-589.

Keller DM, McWeeney S, Arsenlis A, Drouin J, Wright CV, Wang H, Wollheim CB, White P, Kaestner KH, Goodman RH. 2007. Characterization of pancreatic transcription factor Pdx-1 binding sites using promoter microarray and serial analysis of chromatin occupancy. *J Biol Chem* **282**(44): 32084-32092.

Khoo C, Yang J, Weinrott SA, Kaestner KH, Naji A, Schug J, Stoffers DA. 2012. Research resource: the pdx1 cistrome of pancreatic islets. *Molecular endocrinology* **26**(3): 521-533.

McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. 2010. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* **28**(5): 495-501.

Svensson P, Williams C, Lundeberg J, Ryden P, Bergqvist I, Edlund H. 2007. Gene array identification of Ipf1/Pdx1-/- regulated genes in pancreatic progenitor cells. *BMC Dev Biol* **7**: 129.

Teo AKK, Ali Y, Wong KY, Chipperfield H, Sadasivam A, Poobalan Y, Tan EK, Wang ST, Abraham S, Tsuneyoshi N et al. 2012. Activin and BMP4 synergistically promote formation of definitive endoderm in human embryonic stem cells. *Stem Cells* **30**: 631-642.

Teo AK, Arnold SJ, Trotter MW, Brown S, Ang LT, Chng Z, Robertson EJ, Dunn NR, Vallier L. 2011. Pluripotency factors regulate definitive endoderm specification through eomesodermin. *Genes Dev* **25**(3): 238-250.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**(9): R137.
Suppl. Figure 2. Teo et al.

A

164,525,000  164,575,000  164,625,000  164,675,000  164,725,000  164,775,000  164,825,000

Chr 1

PBX1

66,680,000  66,680,000  66,700,000  66,720,000  66,740,000  66,760,000  66,780,000  66,800,000

Chr 2

MEIS1

53,240,000  53,230,000  53,220,000  53,210,000  53,200,000  53,200,000  53,190,000  53,180,000

Chr 12

KRT14

71,840,000  71,860,000  71,880,000  71,900,000  71,920,000  71,940,000  71,960,000  71,980,000

Chr 12

LGR5

28,575,000  28,565,000  28,555,000  28,545,000  28,535,000  28,525,000  28,515,000

Chr 13

CDX2

162,940,000  162,920,000  162,900,000  162,880,000  162,860,000  162,840,000  162,820,000  162,800,000

Chr 2

DPP4


PDX1 binding region

Exon

B

GO term category - Biological Process

developmental process

multicellular organismal development

anatomical structure development

anatomical structure morphogenesis

organ development

nervous system development

anatomical structure formation involved in morphogenesis

central nervous system development

multicellular organismal process

brain development

signaling

embryo development

cellular developmental process

organ morphogenesis

vasculature development

blood vessel development

cell morphogenesis involved in differentiation

C

GO Biological Process

Term Name | Binomial Raw P-Value | Binomial FDR Q-Value
---|---|---
Endocrine pancreas development | 1.1000e-21 | 4.3854e-19
Columnar / cuboidal epithelial cell differentiation | 9.7300e-13 | 1.3403e-10
Cardiac septum development | 2.8900e-9 | 1.9737e-7
Cardiac septum morphogenesis | 2.5900e-7 | 1.2565e-5

Pathway Commons

Term Name | Binomial Raw P-Value | Binomial FDR Q-Value
---|---|---
Regulation of gene expression in early pancreatic precursor cells | 1.6627e-8 | 2.3278e-6

MGI Expression: Detected

Term Name | Binomial Raw P-Value | Binomial FDR Q-Value
---|---|---
Midgut | 9.7027e-25 | 4.5169e-22
Foregut-midgut junction | 1.7766e-14 | 1.0166e-12
Septum transversum/hepatic component | 5.1017e-14 | 2.7832e-12
Foregut-midgut junction endoderm | 5.4569e-14 | 2.9311e-12
Islets of Langerhans | 2.0426e-9 | 4.4436e-8

D

Intestine related genes

\[ \text{fold change} \]

E

HOXA4

DPP4

CDX2

Relative expression level
Suppl. Figure 3. Teo et al.

A

2017 genes PDX1 ChIP-seq
(± 20kb from TSS)
(this study)

3687 genes Human Islet
PDX1 ChIP-seq
(± 20kb from TSS)
(Khoo et al., 2012)

1721
1006
2591

B

2017 genes PDX1 ChIP-seq
(± 20kb from TSS)
(this study)

Total 579 genes
NIT-1 cell line
Pdx1 ChIP-on-chip

2718
89
480

C

GO Biological Process

| Term Name                                      | Binomial Raw P-Value | Binomial FDR Q-Value |
|------------------------------------------------|----------------------|----------------------|
| Pancreas development                          | 2.41705e-10          | 2.11758e-6           |
| Type II pancreatic cell differentiation        | 5.82472e-6           | 3.92541e-3           |
| Endoderm development                          | 1.76506e-4           | 2.76629e-2           |
| Positive regulation of cellular catabolic process | 2.41651e-4          | 3.25660e-2           |

Mouse Phenotype

| Term Name                                      | Binomial Raw P-Value | Binomial FDR Q-Value |
|------------------------------------------------|----------------------|----------------------|
| Abnormal pancreas secretion                    | 1.44501e-6           | 8.10049e-4           |
| Abnormal pancreatic beta cell physiology      | 1.69250e-6           | 8.81672e-4           |
| Abnormal endocrine pancreas physiology        | 1.83077e-6           | 8.38591e-4           |
| Abnormal insulin secretion                     | 1.86913e-6           | 8.14728e-4           |
| Abnormal pancreatic islet size                | 8.16173e-6           | 2.12584e-3           |
| Small pancreatic islets                       | 1.18836e-5           | 3.01366e-3           |
| Abnormal pancreatic beta cell morphology      | 4.32069e-5           | 9.26788e-3           |
| Decreased pancreatic beta cell number         | 1.78973e-4           | 2.37318e-2           |
