Global expression analysis of gene regulatory pathways during endocrine pancreatic development

Guoqiang Gu[1,*,‡], James M. Wells[1,†,‡], David Dombkowski[2], Fred Preffer[2], Bruce Aronow[†] and Douglas A. Melton[1,§]

[1]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA
[2]Department of Pathology, Massachusetts General Hospital, Charlestown, Massachusetts 02129, USA
*Present address: Program of Developmental Biology, Department of Cell and Developmental Biology, Vanderbilt Medical Center, Nashville, TN 37232, USA
†Present address: Division of Developmental Biology, Cincinnati Children's Research Foundation, Cincinnati, OH 45229-3039, USA
‡These two authors contributed equally to this work and are listed alphabetically
§Author for correspondence (e-mail: dmelton@biohp.harvard.edu)

Accepted 30 September 2003

Development 131, 165-179
Published by The Company of Biologists 2004
doi:10.1242/dev.00921

Summary

To define genetic pathways that regulate development of the endocrine pancreas, we generated transcriptional profiles of enriched cells isolated from four biologically significant stages of endocrine pancreas development: endoderm before pancreas specification, early pancreatic progenitor cells, endocrine progenitor cells and adult islets of Langerhans. These analyses implicate new signaling pathways in endocrine pancreas development, and identified sets of known and novel genes that are temporally regulated, as well as genes that spatially define developing endocrine cells from their neighbors. The differential expression of several genes from each time point was verified by RT-PCR and in situ hybridization. Moreover, we present preliminary functional evidence suggesting that one transcription factor encoding gene (*Myt1*), which was identified in our screen, is expressed in endocrine progenitors and may regulate α, β and δ cell development. In addition to identifying new genes that regulate endocrine cell fate, this global gene expression analysis has uncovered informative biological trends that occur during endocrine differentiation.

Supplementary data available online

Key words: Myt1, endoderm, Pancreas, Endocrine, Islets, Microarray

Introduction

The vertebrate pancreas has two functions: producing digestive enzymes (exocrine), and regulating glucose homeostasis (endocrine). These separate functions are reflected in the complex architecture of the pancreas. The acini and ducts form the exocrine pancreas that produces and transports digestive enzymes into the duodenum. The endocrine islets contain four types of cells that secrete hormones to regulate glucose metabolism and other physiological processes (Slack, 1995). Thus, the developing pancreas presents a challenge for developmental biologists because of the complex morphogenetic processes underlying development of this organ. In addition, Type I or insulin dependent diabetes mellitus results from the autoimmune-mediated destruction of insulin-secreting β cells in islets, emphasizing the importance of understanding pancreas and β cell development (Mathis et al., 2001; Tisch and McDevitt, 1996).

The pancreas derives from the endoderm germ layer (Pictet et al., 1972; Slack, 1995), which in mouse is a cup of cells enveloping the mesoderm and ectoderm at embryonic day 7.5 (E7.5). At this time, the endoderm receives signals from adjacent mesoderm and ectoderm and becomes competent to respond to subsequent permissive signals that establish organ domains along the anterior-posterior axis (Wells and Melton, 1999). By E8.5, the endoderm begins to form a primitive gut tube, and the region destined to become the pancreas receives signals from the notochord and dorsal aorta, leading to the expression of essential pancreatic transcription factor genes such as pancreatic-duodenal homeobox 1 [*Pdx1*, also known as insulin-promoter factor 1 (Ipf1)] (Hebrok et al., 1998; Lammert et al., 2001). At E9.0, *Pdx1* expression marks both the dorsal and ventral domains of the developing pancreas, and defines where pancreatic buds will appear around E10 (Guz et al., 1995). As pancreatic buds expand and branch, signals from adjacent mesenchyme direct cells toward an endocrine or exocrine fate (Guz et al., 1995; Miralles et al., 1998a; Miralles et al., 1998b). Cells that have adopted an endocrine cell fate express the bHLH transcription factor neurogenin 3 (NGN3) (Gu et al., 2002).

Functional studies have identified several signaling pathways and transcription factors important for pancreatic development. Initial pancreatic specification of endoderm is mediated by the FGF, hedgehog, Notch and TGFβ/activin signaling pathways (Kim and Hebrok, 2001). These signals result in the expression of genes for several transcription factors in the developing pancreas including, *HNF1α* (*Tcf1α*), *HNF1β* (*Tcf1*β), *HNF4α* (*Tcf4α*), *Pdx1, NeuroD1, Ngn3, Pax4, Pax6* and others (Edlund, 1998). Mutations in some of these genes are associated with maturity onset

diabetes of the young (MODY 1, 3, 4, 5 and 6), and genetic analyses in mice have begun to elucidate how these transcription factors function during discrete stages of pancreas development (Stride and Hattersley, 2002). For example, loss of PDX1 results in defects of both early pancreatic specification and budding (Jonsson et al., 1994; Offield et al., 1996), whereas loss of NGN3 results in specific absence of endocrine cell development (Gradwohl et al., 2000). Moreover, cell lineage analysis supports the idea that PDX1 functions to establish the three basic lineages of the pancreas (ducts, acini, islets), whereas NGN3 functions specifically to establish the endocrine lineages (Gannon et al., 2000; Gu et al., 2002; Herrera, 2000; Herrera et al., 1998; Schwitzgebel et al., 2000).

Analyses of individual genes have begun to define some critical stages in the development of the endocrine pancreas, yet the complex interactions of extracellular signals and the responding genetic networks that control endocrine cell growth and differentiation are largely unstudied. For example, it is not known how *Pdx1* is induced and restricted to a defined region of the developing gut, nor is it known how *Ngn3* expression is temporally controlled resulting in the genesis of endocrine progenitor cells. Recently, 3,400 genes expressed in the pancreas were used to generate an endocrine pancreas microarray (PancChip), which is available through the β Cell Biology Consortium (Scearce et al., 2002). The PancChip will probably be a valuable diagnostic tool for the genetic analysis of pancreatic cell samples. However, the focus of the Endocrine Pancreas Consortium was not to provide a complete and quantitative analysis of the genes that are expressed during the formation of the endocrine pancreas. A transcriptional profile of pancreatic and endocrine progenitors would provide fundamental information about the processes regulating normal development of the endocrine pancreas. Moreover, regulatory factors identified in this screen might be used to promote regeneration of endocrine cells in vivo, or used to direct the differentiation of embryonic stem cells or adult stem/progenitor cells toward the β cell lineage in vitro.

We describe the fundamental gene expression profiles of several tissue or cell samples that define distinct stages during pancreatic and endocrine islet development. We used high-density microarrays from Affymetrix to systematically analyze the genes that are expressed at four key stages of pancreatic and endocrine development: E7.5 unspecified endoderm, E10.5 pancreatic cells that express *Pdx1*, E13.5 endocrine progenitor cells that express *Ngn3*, and mature islets of Langerhans. This genetic analysis is uniquely designed in several ways. First, we used a combination of dissection and cell-sorting using an eGFP reporter that was under the control of promoters of specific pancreatic genes to isolate highly purified cells from these well-defined stages of pancreatic development. Second, we compared both the temporal and spatial expression profile at each stage to more fully define these cell types. Third, we validated our profiles by demonstrating the cell-specific expression of several genes from each time point by RT-PCR and in situ hybridization (ISH). Finally, we demonstrated that one gene we identified, myelin transcription factor 1 (*Myt1*), might be a novel regulator of α, β and δ cell development in the pancreas.

Materials and methods

Transgene construction and transgenic mice generation

To generate the *Pdx1-eGFP* construct, a 1.5 kb DNA fragment containing the coding region of enhanced green fluorescence protein (*eGFP*) and a SV40 polyadenylation signal was amplified by PCR, digested by *Xba*I (via a site introduced in the 3′ end primer) and ligated to *Nco*I (blunt ended)-*Xba*I-digested plasmid *Pdx1-hsp68*-lacZ construct (a kind gift from C. E. V. Wright). The plasmid used as PCR template was pGreenlatern-1 (Clontech, Palo Alto, CA). The primers used are: forward: 5′-agcaagggcgaggaactgttc-3′ and reverse: 5′-catgatctagacatgataagatacattgatg-3′. The insert in the final construct (p#48) was released by *Sal*I digestion and used for transgenic animal production. The hybrid B6CBAF1 mouse strain was used to generate transgenic animals. Five transgenic lines were generated and the eGFP expression pattern was compared with that of PDX1 protein to ensure that eGFP expression mimics that of PDX1 protein. One line, P#48.9, whose expression recapitulates that of *Pdx1*, was used to obtain Pdx1-eGFP$^+$ cells.

An *Xba*I-*Sph*I (partial digestion) fragment that contains sufficient *Ngn3* enhancers (Gu et al., 2002) replaced the *Pdx1* enhancer region in *p#48* to generate the *Ngn3-eGFP* construct (p#63, Fig. 1). Insert was released by *Sal*I digestion to generate three transgenic lines. After verifying that eGFP expression mimics that of NGN3 by double ISH (Gu et al., 2002), one line P#63.1 was used to obtain embryos for cell sorting.

To generate function analyses constructs of *Myt1*, full length *Myt1a* or *Nzf2b* cDNA (a kind gift from L. D. Hudson) was inserted into the *Xho*I-*Sma*I site of the pCIG expression vector [under the control of CMV-beta actin promoter (Grapin-Botton et al., 2001)] to give plasmid p#116 and p#132. The dominant negative construct (*dnMyt1*) is generated using a similar approach, except the transcriptional activation sequence was deleted using a PCR approach. Specifically, two primers (p273: 5′-gacaattgaaggagcttctcacctgtcc-3′ and p252: 5′-ccatgtgtgcacctcagcatc-3′) were used to amplify a DNA fragment that had both the 5′ and 3′ ends of *Myt1a* cDNA and a vector sequence in the middle. This fragment was digested by *Eco*RI and *Mun*I and self-ligated. The resulting plasmid was a partial *Myt1* cDNA that had its transcription activation domain removed (*dnMyt1*). For transgenic animal production, *dnMyt1* was put under the control of the *Ngn3* promoter.

Tissue isolation and cell sorting

To obtain purified endoderm, mesoderm and ectoderm tissue, E7.5 embryos (90) were isolated from timed pregnant female ICR mice (Taconic, Germantown, New York) and the endoderm was manually dissected from the mesoderm and ectoderm with a polished tungsten needle (Wells and Melton, 2000). Isolated germ layers were combined into two pools. Each pool of isolated endoderm and mesoderm and ectoderm contained approximately 0.2-0.4 μg total RNA, which was used for cRNA probe generation (Baugh et al., 2001).

To isolate Pdx1-eGFP$^+$ cells, ICR or CD-1 mice were crossed with P#48 males, and the eGFP-expressing E10.5 embryos were identified under a fluorescence microscope. The pancreatic rudiments and the stomach and duodenum (Std) anlagen were separated by dissection. These tissues were trypsinized to single cells and sorted into eGFP$^+$ and eGFP$^-$ populations by FACS. From 350 eGFP$^+$ embryos, 1.3 and 1.8 million eGFP$^+$ cells were collected from the pancreatic or Std region, respectively. Meanwhile, five million eGFP$^-$ cells were also collected from both dissected samples. From these cells, 6, 8 and 14 μg total RNA was isolated and used for cRNA probe generation (each of these RNA were maintained in several small pools respectively). Ngn3-eGFP-expressing cells were isolated by a similar approach except that only the pancreatic rudiment was isolated, and the stage used was E13.5 (from 1300 eGFP$^+$ pancreata, 1.3 million Ngn3-eGFP$^+$ cells were collected and 5 μg total RNA was made and maintained as two pools).

Mouse islets were isolated by perfusing the pancreas with a collagenase solution (2 mg/ml), filtering the digested pancreas though a 300 μM wire mesh, and centrifugation on a histopaque 1077 cushion (Warnock et al., 1990). Islets were hand-picked to minimize contamination with exocrine tissue. For our analysis, pancreata from five adult animals were used to obtain 30 μg total RNA.

cRNA probe generation and hybridization to Affymetrix microarray chips

Total RNA samples were used to generate cRNA probes by two rounds of transcription (Baugh et al., 2001). Basically, a poly(dT) primer (with its 5′ end carrying T7 promoter sequence) was used to synthesize cDNA from total RNA. The cDNA were used to amplify cRNA using T7 polymerase. The cRNA product from this first round amplification was used to generate more cDNA by random priming, with the 3′ end carrying a T7 promoter sequence. This cDNA was used to transcribe biotinylated cRNA, which was used to hybridize to the Mu11K, Mu74Av1 or MU74Av2 microarrays produced by Affymetrix, following the manufacturer's protocol.

Data normalization and analysis

Two programs were used to analyze the data generated from the microarray hybridization.

First, using MicroArray Suite 5.0 (Affymetrix) image files were examined for uniform image quality without significant scratches or smudged fluorescence patterns. The images were processed into intensity data that was scaled per chip to a target intensity of 1500. Chip reports were examined for evidence of high quality and uniform RNA, RNA labeling, hybridization and scanning using approaches similar to those described at (http://cardiogenomics.med.harvard.edu/groups/proj1/pages/Method_QC.html). In brief, control oligonucleotide signal corresponding to spiked and constitutive RNAs were strong, uniform, sensitive and properly interpreted by the Affymetrix software. Background values were uniformly less than 100 and the scaling factor SF that is used to normalize the signal across the entire chip to 1500 signal units was within a twofold range for all chips. GeneSpring 5.0.1 (Silicon Genetics, Inc., Redwood City, CA) was used to analyze the resulting data values obtained from MicroArray Suite 5.0. The values used for filtering and clustering were 'Signal', 'Signal Confidence', 'Absolute Call' (Absent/Present). Data were normalized as follows: the 50th percentile of all measurements was used as a positive control for each array. Each measurement for each gene was divided by this synthetic positive control, assuming that this was at least 10. The bottom tenth percentile signal level was used as a test for correct background subtraction. The measurement for each gene in each sample was divided by the corresponding value in untreated samples, assuming that the value was at least 0.01. Throughout our analysis, only the genes that display more than threefold change between samples were listed and studied (P=0.01 in at least one statistical test).

Chick embryo electroporation

Chick embryo electroporations followed the reported protocol (Grapin-Botton et al., 2001). Briefly, electroporation was performed on embryos between the 18- and 25-somite stage (i.e., stage 13-15 HH). Eggs were windowed and DNA (2 μg/μl DNA in 1×PBS, 1 mM $MgCl_2$, 3 mg/ml carboxymethylcellulose, 50 μg/ml Nile Blue Sulfate) was injected in the blastocoel. A negative electrode was inserted below the embryo, and a positive electrode was held by a micromanipulator above the embryo and three square pulses of 17 volts for 50 mseconds each were applied (BTX T-820). After electroporation, eggs were incubated at 38°C for 48-60 hours, then collected and fixed in 4% paraformaldehyde/PBS, and sectioned for immunohistochemistry or in situ RNA analysis.

Immunohistochemistry

Electroporated embryos were sectioned and analyzed for hormone expression. Transgenic mouse embryos with the *Ngn3* promoter driving *dnMyt1* expression were analyzed by insulin and glucagon expression. The pancreata from five independently derived F_1 transgenic E14.5 embryos were fixed, completely sectioned (6 μm sections), immunostained with anti-insulin or glucagon antibodies, and the insulin$^+$ and/or glucagon$^+$ cells were counted on alternate paraffin sections. As a control, four littermate pancreata were counted in a similar fashion. Primary antibodies used were guinea pig anti-insulin (Dako, Carpinteria, CA), guinea pig anti-glucagon (Linco, St. Charles, MI) and rabbit anti-glucagon (Chemicon, Temecula, CA). Secondary antibodies used were peroxidase-conjugated donkey anti-guinea pig, FITC-conjugated donkey anti-guinea pig, and Cy3-conjugated donkey anti-rabbit (Jackson Immunoresearch, West Grove, PA). In order to obtain a significant number of insulin$^+$ or glucagon$^+$ cells, at least half of sections from each pancreas was counted.

RT-PCR and ISH

RT-PCR followed standard protocols. The primers used in our analyses were: ApoAIV forward: 5′-aaggtgaagggcaacacggaag-3′, reverse: 5′-cctcaagctggtacaagaagtgc-3′.

HPRT forward: 5′-gctggtgaaaggacctctc-3′, reverse: 5′-cacaggactagaacacctgc-3′ (Johansson and Wiles, 1995). Dkk1 forward: 5′-ggagatattccagcgctgtta-3′, reverse: 5′-ggtaagtgccacactgaggat-3′.

Prss12 forward: 5′-agagagaggccacagaaaacag-3′, reverse: 5′-ttgactccacatccatacccccc-3′.

Eya2 forward: 5′-ttactcccattacccacgggtc3′, reverse: 5′-gaagcctaaacaacgggcaaag-3′. Osteopontin forward: 5-gaagctttacagcctgcacccaga-3′; reverse: 5′-gcttttggttacaacggtgtttgc-3′; T7/osteopontin reverse: 5′-gtaatacgactcactatagggc aacagactaagctaagagccc-3′. Nkx2.2 forward: 5′-ccatgtcgctgaccaacacaaaga-3′; reverse; 5′-cgctcaccaagtccactgctgg-3′; T7/Nkx2.2 reverse: 5′-gtaatacgactcactatagggcggtgtgctgtcgggtactg-3′. Tm4sf3 (AF010499) forward: 5′-cagttccgctgtagcaatggctg-3′; reverse: 5′-cacacacactctaccactgagc-3′. T7/Tm4sf3 reverse: 5′-gtaatacgactcactatagggcagcacaaactacaaagaccca-3′. Spintz1 (AA57115): forward; 5′-gctgcaggcacacggatctctgc-3′; reverse: 5′-cagtgaatacctgtgaagatatc-3′. T7/Spintz1 reverse: 5′-gtaatacgactcactatagggcctcagtgagatacttcaataac-3′. Myt1 forward: 5′-gtctccggtggaagctcatggaca-3′; reverse: 5′-cttatggtgccctagtgtgtcatc-3′; T7/Myt1 reverse: 5′-gtaatacgactcactatagggccattaacataagagggtaa-3′. Rbp forward: 5′-ggctacatcataggtccctttcg-3′; reverse: 5′-tactgcctctctaggcacagctc-3′; T7/Rbp reverse: 5′-gtaatacgactcactatagggctgtctctgggctcaggc-3′. Galphao forward: 5′-gcatgcacgagtctctcatgctc-3′; reverse: 5′-ctagacagactagcctgacatg-3′; T7/Galphao reverse: 5′-gtaatacgactcactatagggcgaggcgccaggcccag-3′. Foxa3 forward: 5′-ataaccatggctattcagcaggct-3′; reverse: 5′-cacaggtcaatcaagattgccaac-3′; T7/Foxa3 reverse: 5′-gtaatacgactcactatagggccatccaacatcacgaccatc-3′; actin control forward: 5′-atgccaacac agtgctgtctggtgg-3′; reverse: 5′-gcgaccatcctcctcttaggagtg-3′

Sectioned in situ analysis was performed as described previously (Grapin-Botton et al., 2001). Paraffin sections (6 μm) were collected on glass slides (Superfrost Plus), dewaxed, treated with 1 μg/ml proteinase K for 7 minutes, and postfixed in 4% paraformaldehyde. Hybridization mix contained 1 μm/ml of probe, and hybridization was done overnight at 70°C. Sections were washed in maleic acid buffer and blocked with 20% lamb serum/2% Blocking Reagent (Boehringer Mannheim, Indianapolis, IN) and incubated overnight with anti-digoxigenin-alkaline phosphatase antibody (Boehringer Mannheim), 1:1000. Slides were washed again and developed with NBT and BCIP.

Whole-mount ISH was performed as described previously (Wilkinson and Nieto, 1993). Briefly, E7.5 embryos were fixed, dehydrated in methanol, rehydrated, treated with 6% hydrogen peroxide, proteinase K treated for 1.5 minutes, and postfixed in 4% paraformaldehyde. Embryos were hybridized in buffer containing 1 μg/ml probe overnight at 70°C. Embryos were washed and incubated overnight with an anti-digoxigenin antibody (1:1000). Embryos were developed with BM purple (Boehringer Mannheim). Probe templates for ApoAV, Dkk1, Prss12 and Eya2 were generated by PCR

amplification from an E7.5 endoderm library (Harrison et al., 1995) using a gene-specific forward primer (mentioned above), and a vector specific (pSPORT) reverse primer. The resulting amplified product contained the 3′ end of the gene and an SP6 polymerase site from the pSPORT vector. The amplified products were verified by sequencing and used in an in vitro transcription reaction to generate antisense probes. To generate cRNA probes for *Foxa3*, galphao, osteopontin, *Myt1*, *Nkx2.2*, *Rbp*, *Spintz1*, and *Tm4sf3*, T7-reverse primers (has T7 promoter sequence at the 5′ end, see above) were used to amplify cDNA fragments with corresponding forward primers.

Results

Isolation of cells and generation of cRNA

Our approach focused on two questions. (1) Which transcripts are up- or down-regulated as undifferentiated endoderm adopts a pancreatic and then endocrine cell fate (temporal gene expression)? (2) Which transcripts distinguish developing endocrine cells from adjacent cells at each stage of development (spatial gene expression)? We isolated tissue samples from four stages of the developing endocrine pancreas (Fig. 1), and separated the developing endoderm, pancreatic, or endocrine cells from their neighboring cells using manual dissection and/or cell sorting. These highly enriched cell samples were used to make targets for hybridization to Affymetrix microarrays that contain over 12,000 genes and ESTs (Fig. 1 and Materials and methods).

The stages shown in Fig. 1 were chosen for the following reasons. (1) E7.5 endoderm. At E7.5, the endoderm is a sheet of cells on the outside of the embryo. At this stage, endoderm cells are plastic and are not yet determined to form the pancreas (Wells and Melton, 2000). Analysis of undifferentiated endoderm should provide a genetic baseline and highlight genes involved in endoderm plasticity and pancreas differentiation. (2) *Pdx1*-expressing cells of the E10.5 pancreatic rudiment. PDX1$^+$ cells will yield both the exocrine and endocrine components of the adult pancreas and are therefore considered pancreatic progenitor cells (Gannon et al., 2000; Gu et al., 2002). At this stage, PDX1$^+$ cells are also found in the stomach and duodenum (Offield et al., 1996). A transcriptional analysis of PDX1$^+$ cells from the pancreas versus PDX1$^+$ cells from the stomach and duodenum, or from PDX1$^-$ cells, should highlight genes that specify pre-pancreatic cells from their gastrointestinal neighbors. (3) Endocrine progenitor cells (NGN3$^+$) of the E13.5 pancreas. The cells that express *Ngn3* at this stage will form only endocrine tissue (Gu et al., 2002). A comparison of the transcriptional profile of NGN3$^+$ cells with NGN3$^-$ cells was aimed at distinguishing the endocrine and exocrine compartments of the embryonic pancreas. (4) Adult islets. Adult islets represent mature, differentiated endocrine cells and will highlight the genes that need to be up-regulated, as well as down-regulated, in order to form the endocrine pancreas. This experimental approach was designed to quantitatively identify genes that are temporally and spatially regulated during endocrine development.

The following methods were used to obtain tissue samples for transcriptional analysis. (1) The endoderm from 90 E7.5 embryos was manually separated from mesoderm/ectoderm. (2) The mouse *Pdx1* promoter, which recapitulates the endogenous *Pdx1* expression (Gu et al., 2002; Wu et al., 1997), was used to drive expression of eGFP, and eGFP expression

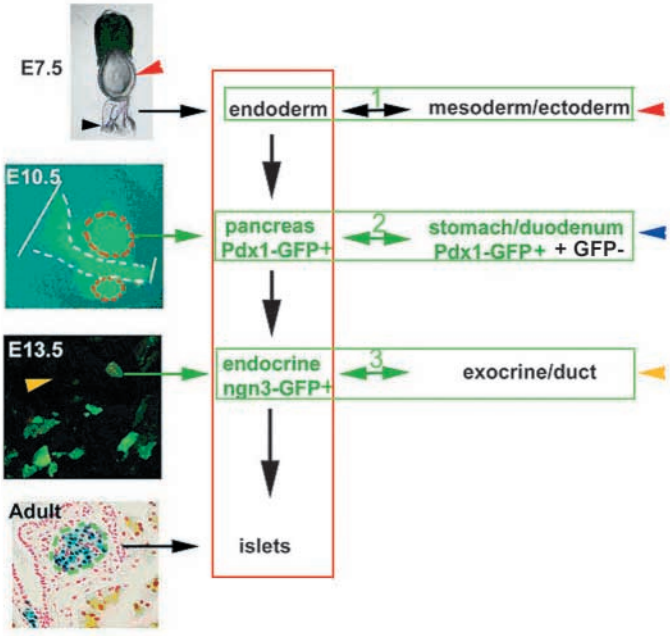

Fig. 1. Scheme of the temporal and spatial gene expression analysis for four stages of islet development. The red box highlights the temporal gene expression comparison and the green boxes highlight the spatial expression comparisons. (1) At E7.5, endoderm (black arrowhead) and mesectoderm (red arrowhead) were manually separated and analyzed. (2) At E10.5, the eGFP$^+$ cells from pancreatic buds (red dashed lines), stomach/duodenum (white dashed lines; two solid white lines indicate the approximate borders of dissection), and eGFP$^-$ cells from *Pdx1-eGFP* transgenic animals were separated by FACS and analyzed. (3) At E13.5, endocrine progenitor cells expressing Ngn3-eGFP were separated by FACS from eGFP$^-$ cells (yellow arrowhead, primarily exocrine and ductal cells) and both eGFP$^+$ and eGFP$^-$ cells were analyzed. Adult islets were hand picked and used for direct analysis (green dashed lines. The blue staining within the islet is from *Pdx1-lacZ*). cRNA probes for each sample were hybridized to Affymetrix microarrays Mu11K, Mu74Av1 or Mu74Av2 (Materials and methods). Temporal analysis (the red box) compared the gene expression patterns of endoderm, pancreatic progenitors (Pdx1-eGFP$^+$), endocrine precursors (Ngn3-eGFP$^+$), and adult islets. Spatial analysis (green boxes 1, 2 and 3) compared different samples from the same stages, e.g. genes expressed in E7.5 endoderm versus mesoderm + ectoderm.

was used to FACS sort PDX1$^+$ from PDX1$^-$ from dissected pancreas, stomach and duodenum. A total of 1.3×10^6 or 1.8×10^6 Pdx1-eGFP$^+$ cells (from pancreatic or stomach/duodenumal regions, respectively) were isolated from 350 E10.5 embryos. The trypsinization of tissue before cell sorting did not alter the ability of these cells to differentiate into insulin-producing cells in vitro (G.G. and D.A.M., unpublished data), nor did it dramatically alter the presence or absence of predicted gene expression in this analysis. However, we cannot rule out the possibility that the expression levels of some genes were altered by this isolation method. (3) The *Ngn3* promoter, which recapitulates endogenous *Ngn3* expression (Gu et al., 2002), was used to drive eGFP expression in endocrine progenitor cells. 1.3×10^6 Ngn3-eGFP$^+$ cells were collected from 1,300 E13.5 embryos. (4) Islets were isolated from 10 adult mice. All tissue or cell samples were separated into duplicates and used to generate labeled cRNA samples using

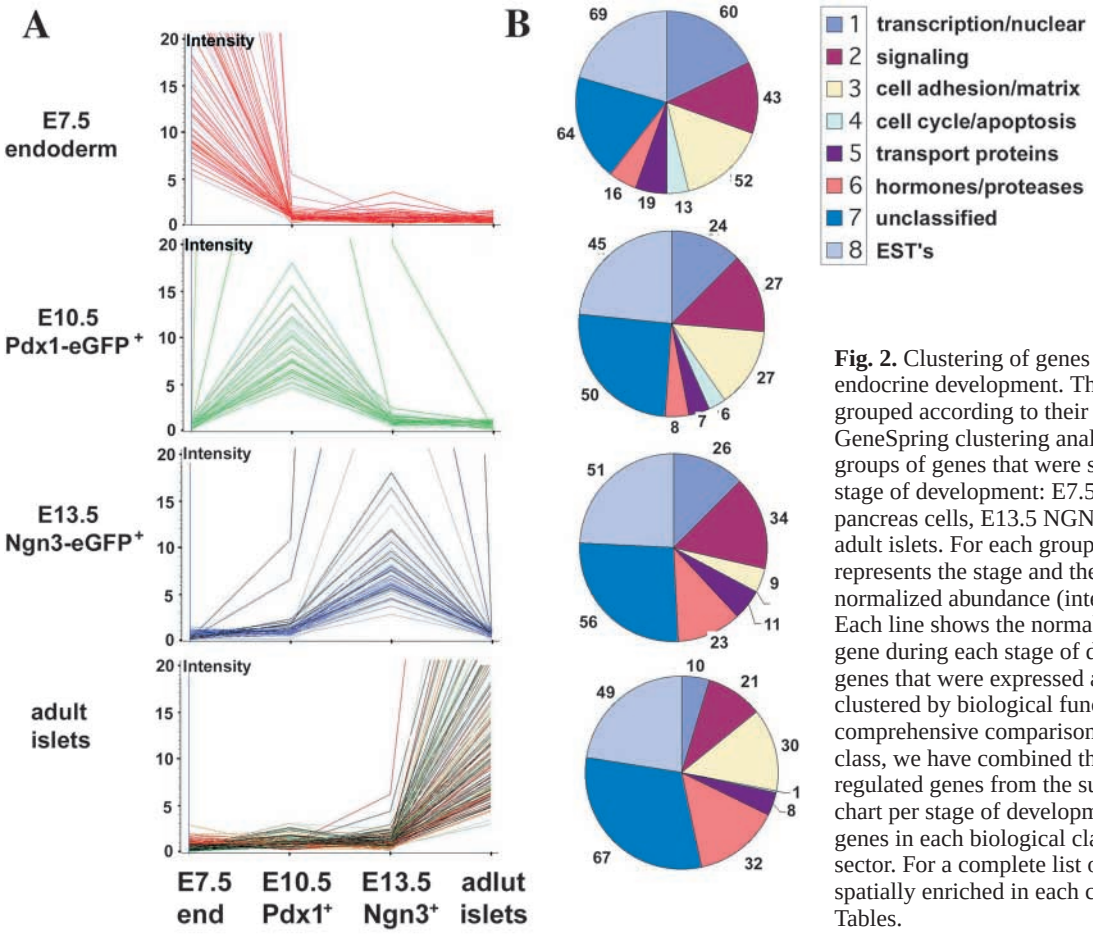

Fig. 2. Clustering of genes expressed at each stage of endocrine development. These clustered genes are also grouped according to their biological function. (A) A GeneSpring clustering analysis that identified four groups of genes that were specifically expressed in each stage of development: E7.5 endoderm, E10.5 PDX1$^+$ pancreas cells, E13.5 NGN3$^+$ endocrine progenitors, and adult islets. For each group of genes, the X-axis represents the stage and the Y-axis represents the normalized abundance (intensity) of each transcript. Each line shows the normalized expression level of one gene during each stage of development. (B) Pie charts of genes that were expressed at each stage of development, clustered by biological function. For a more comprehensive comparison of genes in each biological class, we have combined the temporally and spatially regulated genes from the supplemental tables into one pie chart per stage of development. The absolute number of genes in each biological class is shown next to each pie sector. For a complete list of genes either temporally or spatially enriched in each cell sample, see Supplementary Tables.

an in vitro transcription-based linear amplification protocol (Baugh et al., 2001). Amplified RNA samples were hybridized to the Affymetrix microarrays (Materials and methods), and the data were analyzed using GeneSpring and Resolver clustering analysis software. Genes expressed at each stage of development were grouped according to biological function (Fig. 2B and tables), and separated into classes that are temporally or spatially regulated during endocrine development. Genes that were expressed in the pancreas, but were not temporally or spatially regulated were generally not listed in the tables (see supplemental data for a complete listing of genes expressed in these samples: http://dev.biologists.org/supplemental).

Analysis of temporal gene expression during endocrine islet development

We used Affymetrix software (M.A.S.5) to identify genes expressed at significant levels within each sample. We found that 47, 38, 35 and 46% of the genes present on the microarrays are expressed in the E7.5 endoderm, E10.5 pancreatic progenitor cells, E13.5 endocrine precursors, and islets of Langerhans, respectively (data not shown). We used GeneSpring software (Silicon Genetics) to group genes whose expression is temporally restricted a specific stage of development. Three different statistical group comparisons were used (Student's *t*-test, Welch *t*-test and a nonparametric test). In order to have high confidence that selected genes are differentially expressed, we focused on genes that exhibit at least a threefold expression difference between samples. Raw data are available at www.genet.chmcc.org (contact G.G. for details). We identified 193, 60, 71 and 217 genes whose expression is enriched in E7.5 endoderm, E10.5 PDX1$^+$ pancreatic cells, E13.5 endocrine progenitors, and endocrine islets respectively (Fig. 2).

Endoderm cells express many transcripts involved in pattern-formation

E7.5 endoderm expresses 193 genes or ESTs (out of the ~12,000 on the microarray) at greater than threefold higher levels than cells at later stages of pancreas development. These include 25 growth factors or other signaling-related molecules and 44 transcription factors or other nuclear proteins (Fig. 2). Many of these factors were previously implicated in embryonic pattern formation. For example, endoderm expresses molecules involved in TGFβ signaling, including Nodal, cerberus 1, and follistatin and the *Wnt* antagonist dickkopf (Bouwmeester et al., 1996; Conlon et al., 1994; Mukhopadhyay et al., 2001). Endoderm-expressed transcription factors including *Cdx1*, *Hesx1*, *Irx3*, *Gata3*, *MespI* and *Sox17* (see Table S1, http://dev.biologists.org/supplemental). In addition, we have implicated several new signaling pathways in endoderm and pancreatic development by virtue of their abundant expression. Some examples include the *cKit* ligand, *Edg2* (G-protein coupled receptor) and *Epha2* (Eph receptor A2). The *cKit*

Table 1. A partial list of abundant genes both spatially and temporally restricted in each sample (for a complete list of these genes, see Table S1*)

Endoderm (E7.5)		PDX1+ cells (E10.5)		Endocrine progenitors (E13.5)		Mature islet	
Signaling molecules							
Cer1	Cerberus 1 homolog	*Dlk1*	Delta-like homolog 1	*Alk6*	Alk-6	*Notch4*	Notch homolog 4
Dab2	disabled homolog 1	*IGFbp5*	IGF binding protein 5	*Dlk2C*	Threonine/serine protein kinase Dlk	*Inha*	Inhibin alpha
Dkk1	Dickkopf homolog 1	*SfrP1*	Secreted frizzled-related protein 1	*Mfng*	Manic fringe	*Thra*	thyroid hormone receptor alpha
Edg2	Endothelial differentiation receptor 2			*Pim2*	Pim 2 kinase	*Prlr*	Prolactin receptor
EphA2	Eph receptor A2						
Kitl	c-Kit ligand						
Transcription factors							
Cdx1	Caudal type homeobox 1†	*Barx1*	BarH-like homeobox 1	*Brn4*	Brain pou-domain 4	*Atf5*	Activating transcription factor 5
Hesx1	Homeobox gene expressed in ES cells	*Nkx6.2*	NK related transcription factor 2	L-*Myc*	Murine L-myc	*Cpeb*	Cytoplasmic polyadenylation element binding protein
Irx3	Iroquouis related protein 3	*Ocut1*	One cut domain, member 1	*MafB*	v-Maf oncogene	*Myt1*	Myelin transcription factor 1
Gata3	GATA binding protein 3	*Sox11*	SRY-box containing gene 11	*Myt1L*	Myelin transcription factor 1-like	*Stat5B*	Stat5B
Msep1	Mesoderm posterior 1	*Zac1*	zinc finger protein regulator of apoptosis and cell cycle arrest	*NeuroD1*	Neural differentiation 1		
Sox17	SRY-box containing gene 17			*Pax4*	paired box-homeo-protein 4		
Cell adhesion/matrix protein							
Cubn	Cubilin	*Col1a1*	Collagen 1, alpha 1 subunit	*Anxa4*	Annexsin A4	*CD84*	CD84 antigen
EndoA	Cytokeratin endoA	*Col1a2*	Collagen 1, alpha 2 subunit	*Mtap1b*	Microtubule associated protein 1b	*Crpd*	Crp-ductin
		Col5a2	Collagen 5, alpha 2 subunit				
		Tnc	Tenascin C				
		Vnn1	Vanin 1				

*http://dev.biologists.org/supplemental/

pathway is known to function during hematopoiesis and germ cell migration and development (Ueda et al., 2002) and both of these processes involve interactions with endoderm. Thus, the role of endodermally expressed *cKit* may be restricted to hematopoietic and germ cell development.

Gene expression complexity decreases as cells become restricted to the pancreatic lineage

The PDX1$^+$ cells of the E10.5 pancreas (precursors to all components of the developing pancreas) expressed 60 genes at enriched levels (Fig. 2), a smaller number than the endoderm-specific genes. This is consistent with the PDX1$^+$ cells being a fate-restricted population while the endoderm cells contain progenitors for all endoderm-derived organs (Wells and Melton, 1999).

Examination of these genes suggested that Notch activity and Wnt signaling might play roles in promoting endoderm to adopt a pancreatic fate, since the genes for the Notch ligand *Dlk1* and Wnt signaling antagonist *Sfrp1* were highly expressed in these PDX1$^+$ cells (Table 1). In addition, genes for several transcription factors, including *Barx1, Nkx6.2, Onecut1*, *Sox11* and a few other zinc finger proteins were highly expressed in the PDX1$^+$ cells. Several ECM proteins, including collagens Iα1, Iα2, Vα2, tenascin and vinin 1 were also highly enriched in the PDX1$^+$ cells, suggesting that these molecules could be involved in the budding process of the early pancreatic epithelium (reviewed by Kim and Hebrok, 2001).

We identified 71 transcripts that are enriched in NGN3$^+$ endocrine progenitors (Fig. 2). *Manic fringe, IGFbp*, an activin-receptor-like kinase (*Alk6*), and two serine/threonine protein kinase transcripts are abundantly expressed (Tabe 1; Table S1, http://dev.biologists.org/supplemental). *Manic fringe* encodes a glycosyl transferase and is an important modifier of Notch signaling (Johnston et al., 1997; Shimizu et al., 2001). Its expression only in the endocrine progenitors suggested its involvement in endocrine development. Several transcription factors, including mouse brain-2 Pou domain protein and myelin transcription factor 1 are also expressed in the NGN3$^+$ progenitors, suggesting their involvement in development of the endocrine pancreas. Relative to the other stages of pancreatic development, the number of extracellular matrix/cell adhesion molecules is low in endocrine precursors. This finding is consistent with the idea that endocrine progenitor cells are not part of the epithelium, but rather have delaminated and remain apart from the branching exocrine pancreas (Kim and Hebrok, 2001).

Genes expressed in adult islets

The islet preparation contained the four major endocrine cell types, endothelial cells, some exocrine cells, and other cells that contaminated the islet preparations. We found that the expression of 217 genes (Fig. 2; Table S1, http://dev.biologists.org/supplemental) were enriched at this stage, and most of these are associated with the function of the adult organ. Among these, the transcripts for four endocrine hormones, hormone processing enzymes, secretory apparatus, prolactin receptor, REG1 and REG3, were found at very high levels. In addition, we identified the novel expression of numerous regulatory molecules in adult islets (Table S1, http://dev.biologists.org/supplemental). Genes for the transcription factors that were expressed include activating transcription factor 5 (*Atf5*), myelin transcription factor 1-like (*Myt1l*), putative homeodomain transcription factor (*Phtf*), and short stature homeobox 2 (*Shox2* also *Prx3*). Although the role of these transcription factors in islet function or maintenance is not known, *Atf5*, *Mytl1* and *Shox2* are all expressed in the

CNS, implicating them in neuroendocrine as well as pancreatic endocrine development and function (Angelastro et al., 2003; Kim et al., 1997a; van Schaick et al., 1997). There were also components of several signaling pathways expressed, including Notch 4, inhibin α, Wnt4, leukemia inhibitory factor receptor, and epidermal growth factor, to name a few. These molecules and pathways are possibly involved in regulation of islet size, function and perhaps maintenance.

The adult islets also expressed many of the same transcription factors that function in embryonic pancreatic development. One example is *Pdx1*, which is expressed in entire embryonic pancreas, but is restricted to β cells in the islets. PDX1 was shown to regulate expression of several genes in islets including insulin, glucagon, somatostatin, islet amyloid polypeptide (*Iapp*), glucokinase and *Glut2* (Brissova et al., 2002; Perfetti et al., 2001). PDX1 is also implicated in β cell maintenance in the adult (Sharma et al., 1999; Wells and Melton, 1999), suggesting that one additional role of some embryonic transcription factors might be maintain progenitor cells in the adult.

Gene expression levels as an indicator of differentiation, plasticity and transformation

As endocrine progenitor cells differentiate and form islets, the number of transcriptional and growth factor molecules expressed in endocrine cells decreased. These data suggest that maintenance of progenitor cell plasticity may depend on low-level expression of multiple regulatory genes. Alternatively, the fact that progenitor cells expressed numerous regulatory genes at low levels could reflect the heterogeneity of the progenitor pools. Analyses of genes expressed in single cells of the E10.5 pancreas suggested that *Pdx1*-expressing cells are a relatively heterogeneous population (Chiang and Melton, 2003). Another interpretation of that data is that only a subset of PDX1$^+$ cells are specified toward pancreatic lineages and the remainder are still plastic. This idea is supported by cell lineage studies which demonstrated that many of the cells of the embryonic pancreas, once thought to be pancreatic progenitor cells, never actually contribute to the adult organ (Herrera, 2000).

To identify additional genes that might regulate early cell plasticity, we performed a clustering analysis to identify genes that were down-regulated as a function of differentiation. This cluster of genes contains many known regulators of differentiation, proliferation and plasticity during development (Fig. S1; Table S2, http://dev.biologists.org/supplemental). Included in this cluster of 'down-regulated genes' were numerous tumor-associated genes such as *Tera* (teratocarcinoma expressed, serine rich), *Tacc3* (transforming acidic coiled coil containing protein 3), *Ptov1* (prostate tumor over expressed 1), *Tacstd2* (tumor-associated calcium signal transducer 2), *Trap1a* (tumor rejection antigen 1), *Frat1* (frequently arranged in advanced T-cell lymphomas), and *Lag* (leukemia associated gene). Although the function of these factors in pancreas development is unknown, they were all identified by their abundant expression in different types of tumors and are thus implicated in cellular transformation.

Analysis of genes that are spatially restricted during islet cell development

Our temporal analysis of gene expression identified genes that were known to regulate temporal cell differentiation during endocrine cell development. However, it is equally important to identify the genes that define developing pancreatic cells from their neighbors. For example, how are PDX1$^+$ cells of the pancreas different from the PDX1$^+$ cells of stomach or duodenum, and how do the NGN3$^+$ cells differ from NGN3$^-$ cells? To catalog the genes that control these cell fate decisions, we have generated a transcriptional profile from developing endocrine progenitor cells and from adjacent cells at each stage of development (Fig. 1, green boxes).

Genes differentially expressed in the early endoderm as compared to mesoderm and ectoderm

In order to identify the genes whose expression is spatially restricted to endoderm at E7.5, we compared gene expression profiles between E7.5 endoderm and mesoderm plus ectoderm (Fig. 1, green box 1). We identified 203 transcripts that are greater than threefold enriched in endoderm, while 262 were enriched in the mesoderm plus ectoderm (Fig. 3, Table 2; Table S3, http://dev.biologists.org/supplemental). We have verified endodermal expression of 25 genes by RT-PCR, and 17 of these were further analyzed by ISH analysis (Fig. 3; Table S6, http://dev.biologists.org/supplemental). The gene expression patterns shown in Fig. 3 (*ApoAIV, Dkk1, Prss12*, and *Eya2*) are representative examples of genes that were expressed in endoderm. The expression of these genes in endoderm validates that our approach was successful in identifying endodermally expressed genes.

Several transcription factor mRNAs such as *Foxa2* (*HNFf3β*) and *Sox17* were known to be expressed in endoderm and were detected using microarrays. We also found that several genes with homologs in *Drosophila* such as *Hes1* (hairy enhancer of split) and *Klf5* (kruppel-like factor 5) were enriched in endoderm (Table 2; Table S3a, Table S3b, http://dev.biologists.org/supplemental). Genes for two transcription factors, *EYA2* and *Six1*, that were characterized by their function during eye development, are expressed by E7.5 endoderm cells (Fig. 3), but their function here is unknown. There were several signaling molecule genes that were more abundantly expressed in endoderm, as compared to mesoderm + ectoderm, including *Wnt11, IgfII, chordin,* cerberus 1 (*Cer1*) and genes encoding proteins that enhance growth factor activity, such as, *Fgfbp1* and *Igfbp5*. The co-expression of factors with opposite activities in endoderm highlights the complex nature of signals involved in patterning the endoderm and the adjacent germ layers at this stage of development (Beddington and Robertson, 1999).

Genes differentially expressed in PDX1$^+$ cells of the pancreas, stomach and duodenum

Pdx1 expression marks all pancreatic progenitors of the E8.5-10.5 pancreas (Gannon et al., 2000; Gu et al., 2002). Yet, *Pdx1* is also expressed in cells in rostral stomach, and the mucosal cells of the duodenum (Offield et al., 1996), demonstrating that additional factors are necessary to specify pancreatic fate. We isolated PDX1$^+$ cells of the pancreas, stomach and duodenum to identify the genes that are specifically expressed in pancreatic progenitor cells (Fig. 1, green box 2). Cell lineage analyses have demonstrated that PDX1$^+$ cells in the pancreatic buds at E10.5 give rise to all pancreatic tissues whereas the PDX1$^+$ cells in the stomach/duodenum rudiment do not give rise to pancreatic tissues (Gu et al., 2002). We also analyzed PDX1$^-$ cells from the mesoderm surrounding the pancreas,

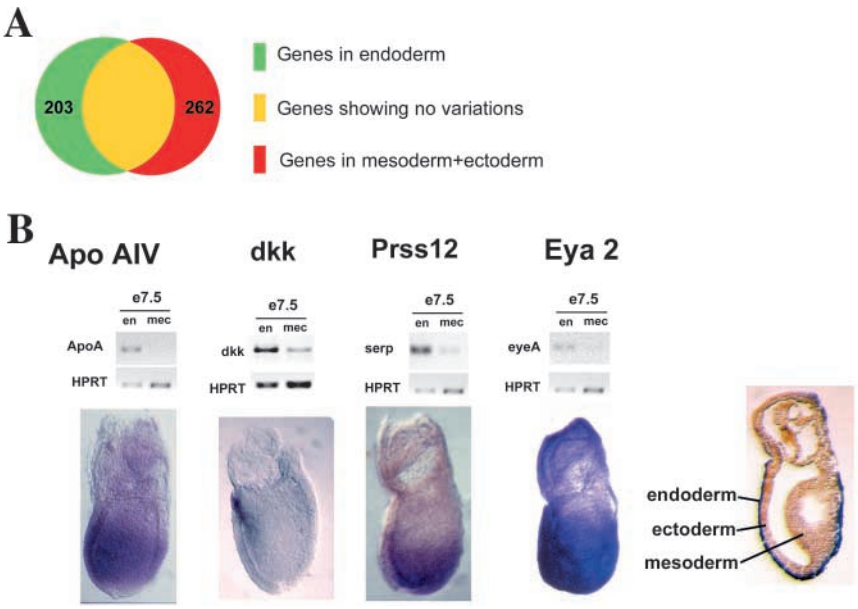

Fig. 3. Genes expressed in endoderm. (A) Venn diagram illustrating genes that were at least threefold enriched in unspecified endoderm as compared to adjacent mesoderm and ectoderm. 203 genes were enriched in the endoderm, whereas 262 genes were enriched in the mesectoderm. (B) The endodermal expression of four genes, *Apo AIV, Dkk1, Prss12* and *Eya2*, identified from our microarray analysis, was verified by RT-PCR (end, endoderm; mec, mesoderm + ectoderm; HPRT is used as a control for RT-PCR) and whole mount ISH. These genes were not previously shown to be expressed in endoderm. The endodermal expression of *Eya2* was verified by sectioning the stained embryo. ect, ectoderm; end, endoderm; mes, mesoderm.

stomach and duodenum. These include the mesenchymal cells surrounding the endoderm and PDX1$^-$ epithelial cells.

We identified the transcripts that are enriched in the pancreatic PDX1$^+$ cells by comparing the expression profile of these cells with that of the combined expression profile of the PDX1$^+$ cells of the stomach and duodenum, as well as that of the PDX1$^-$ cells. This clustering analysis identified 158 genes that are enriched in the PDX$^+$ pancreatic buds. 208 transcripts were enriched in the stomach, duodenum, and the PDX1$^-$ cells (Fig. 4A and Table 3; Table S4, http://dev.biologists.org/supplemental). We verified the expression pattern of 25 candidate genes whose transcripts were enriched in the PDX1$^+$ pancreatic cells by RT-PCR (25-30 cycles) and 12 by ISH. We determined that the transcripts of 21 of the 25 genes were highly enriched in the pancreatic epithelium, compared to that of the duodenum or stomach and surrounding mesenchymes. Expression of the remaining four candidates was not detectable in any tissue (Fig. 4; Table S6, http://dev.biologists.org/supplemental). Similarly, 15 of the 18 candidate genes whose transcripts were enriched in the non-pancreatic cells were found by RT-PCR and/or ISH to be enriched only in the mesenchyme, stomach or duodenum (Table S6, http://dev.biologists.org/supplemental). The remaining three transcripts were not detected in any tissues (Fig. 4B-E and data not shown). We increased the number of PCR cycles in our analysis to 45 and found that we could detect the seven low-abundance transcripts. Data from our Affymetrix analysis predicted these seven genes to be expressed at low levels.

Genes or signaling pathways known to function for pancreas development are expressed in the pancreatic PDX1$^+$ cells

Several genes that are known to be involved in pancreatic function were detected only in the PDX1$^+$ cells in the pancreatic buds. Examples include glucagon, *App* (amyloid precursor proteins), *Glut2* transporter, the vesicle forming proteins *Cop4*, and clathrin coating protein *AP47*. In addition, transcripts of different components for signaling pathways known to function for pancreas development were also detected at enriched level in the pancreatic PDX1$^+$ cells. These include *Notch1* and its ligand *Delta-like 1*, and several FGF receptors (Table 3). We also confirmed that genes that are known to play a role in stomach or duodenum development, including *Rab8, IGFBP2, Shh, Ihh* (Ramalho-Santos et al., 2000), and those of several transcription factors, including *Elf3, Eklf, KlF4, Pax1, Sox2* and *Sox11*, are substantially enriched in PDX1$^+$ cells in the stomach or duodenum and/or mesenchymal cells.

Identification of new pathways or factors that are expressed in pancreatic PDX$^+$ cells

Several genes that were not known to be involved in pancreatic development were found to be expressed by the pancreatic PDX1$^+$ cells. Examples include a G protein (*RhoB*), a related signaling member [diaphonos homolog 1 *(Dab1)*] and calmodulin (*Cldn*, Table 3). Because Rho plays an essential role in focal adhesion formation, another molecule, FAK (focal adhesion kinase), also detected in these cells, (data not shown) may interact with the four above-mentioned molecules to control the morphogenesis of the pancreatic rudiment.

Genes differentially expressed in early endocrine (NGN3$^+$) progenitors

Pancreata from E13.5 embryos were dissected from animals expressing eGFP from the *Ngn3* promoter, and cells were dissociated and separated into NGN3$^+$ and NGN3$^-$ cells based on their eGFP expression [The *Ngn3* promoter used in these experiments recapitulates endogenous *Ngn3* expression (Gu et al., 2002)]. We determined that 204 genes were enriched in endocrine progenitors, as compared to 256 genes that were enriched in non-*Ngn3*-expressing cells (Fig. 5A, Table 4; Table S5, http://dev.biologists.org/supplemental). All genes known to be important for islet development were detected at high levels only in the NGN3$^+$ cell samples (Table S5a). In addition, transcripts of many genes not previously identified as playing roles in endocrine development were also enriched in the Ngn3-eGFP$^+$ cells. In the Ngn3-eGFP$^-$ cells, *Ngn3* transcripts were not detected, demonstrating that our sorted Ngn3-eGFP$^-$ pool was devoid of *Ngn3*-expressing cells. We used ISH to analyze

Table 2. A partial list of genes enriched in early endoderm compared with mesoderm/ectoderm (for a complete list, see Table S3a,b*)

E7.5 endoderm		E7.5 mesoderm and ectoderm	
Signaling molecules			
Cer1	cerberus 1 homolog	*Crabp2*	Cellular RA bp2
Chrd	Chordin	*Fst*	Follistatin
Dkk1	Dickkopf homolog 1	*Tssc3*	Tumor surpressing
FGFbp1	FGF binding protein 1	*Gng3*	G-protein gamma 3
IGF11	Insulin growth factor like 1		
Igfbp5	IGF binding protein 5		
Wnt11	Wnt11 signaling molecule		
Transcription factors			
Eya2	eyes absent 2 homolog	*Gbx2*	gastrulation brain homeobox2
Fos	Fos oncogene		
Foxa2	Forkhead box a2	*Hoxa1*	homeobox a1
Hes1	Hairy enhancer of split 1	*Pou5f1*	POU class 5 TF 1
Klf5	Krupple factor 5	*Sox 2, 3*	SRY-box 2, 3
Six1	sine oculis homeobox 1	*TF1*	Transcrption factor 1
Sox17	SRY-box containing gene 17	*Zic2*	Zinc finger of the cerebellum 1
Cell adhesion/matrix/protease protein			
Prss12	protease, serine, 1	*Actc1*	Cardiac alpha actin
		LamR1	Laminin receptor
Unclassified			
Apoa4,e	Apolipo protein a4, e	*Bcat1*	Branched chain amino transfer
		UTF1	Undifferentiated in ES cells

*http://dev.biologists.org/supplemental/

the expression pattern of 18 candidate genes whose transcripts were only present in endocrine progenitor (NGN3$^+$) cells. 12/18 candidates analyzed were expressed in a scattered cell population in the E10.5, E12.5 and E15.5 pancreat ic rudiments (Fig. 5B-E; Table S6, http:// dev.biologists.org/supplemental), an expression pattern that is highly similar to that of *Ngn3* (Gradwohl et al., 2000). Six of the 18 candidates cannot be detected by ISH, possibly because they are expressed at low levels, which would be consistent with their low hybridization intensity on the microarray (data not shown).

Endocrine progenitors only transiently express *Ngn3* prior to differentiating into mature endocrine cells (Gu et al., 2002). Since eGFP protein is very stable, we anticipated that eGFP$^+$ cells isolated from the *Ngn3-eGFP* transgenic animals would contain some cells that had differentiated toward mature endocrine cells, yet still had eGFP. It is therefore not surprising that substantial levels of somatostatin, glucagon and insulin transcripts were detected in the Ngn3-eGFP$^+$ cell pool. The expression levels of these hormones in the Ngn3-eGFP$^+$ cells were less than 5% of the expression levels in adult islets (data not shown). This finding is consistent with the idea that Ngn3-eGFP$^+$ cells are the endocrine precursors that eventually give rise to mature endocrine islets.

Several G-protein signaling components were enriched in endocrine progenitors

Transcripts encoding several G protein-coupled receptors *(GPR27* and *GPR56*) and multiple guanine nucleotide-binding proteins, including *Gα0, Rab3D, Rab7* and a *GDP* dissociation inhibitor (Table 4), were highly enriched in the NGN3$^+$ cells.

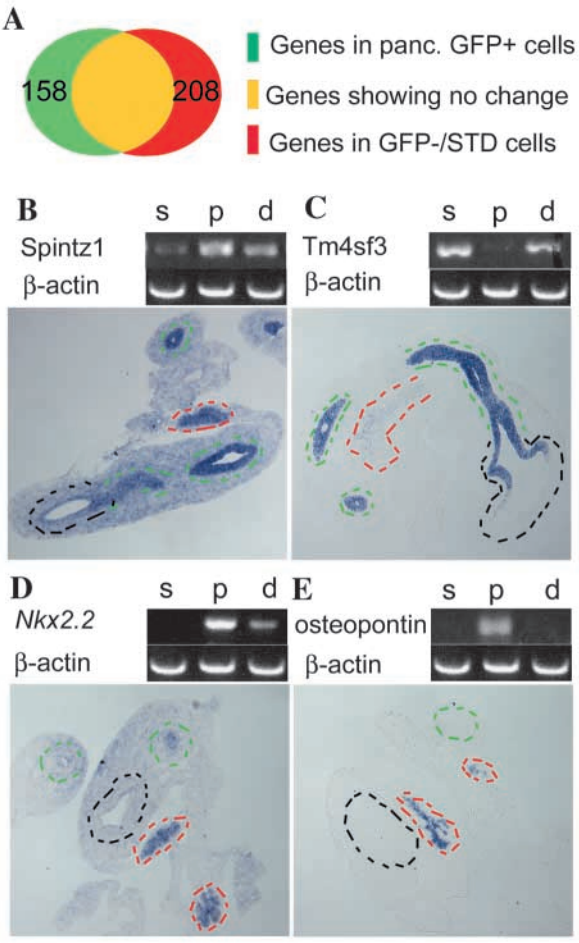

Fig. 4. A summary of genes expressed in the pancreatic Pdx1-eGFP$^+$ cells and non-pancreatic cells at E10.5 (including Pdx1-eGFP+ cells of the stomach + duodenum as well as PDX1-GFP$^-$ cells). (A) Venn diagram showing that 158 genes (green) were enriched in the eGFP$^+$ cells within the pancreatic region whereas 208 (red) genes were enriched in the PDX1$^+$ cells in the duodenum and stomach or eGFP- cells. (B-E) Expression analyses of four genes by RT-PCR and in situ hybridization. β actin expression was used as a control. (B) Spintz1(Af010499) was highly expressed in the pancreas and duodenum but not the stomach. (C) Tm4sf3 (AA571115) was expressed at a high level in the stomach and duodenum but not the pancreas. (D) *Nkx2.2* was expressed at a higher level in the pancreas. (E) Osteopointin was only expressed in the pancreas. Black dashed lines, stomach; green dashed lines, duodenum; red dashed lines, pancreas.

Transcripts for several calcium signaling-related molecules, a calcium-binding protein (*ALG2*), a calcium-dependent activator (*Cadps*), calcium-dependent kinase II (*Camk2b*), and a calcium-independent phospholipase A II (*Pla2g6*) were also highly enriched in endocrine progenitors. The presence of these molecules suggests that G-protein-mediated signaling, through receptor GPR27 or GPR56, and calcium mediated signaling might participate in endocrine development or function.

Components of the notch-signaling pathway are expressed by endocrine progenitor cells

Our screen not only revealed the presence of the transcripts for

Table 3. List of genes differentially expressed in PDX1$^+$ pancreatic (but not stomach) cells compared with those in PDX1$^+$ stomach and duodenum (but not pancreatic) and PDX1$^-$ cells (for detailed description of these genes and all candidates, see Table S4*)

PDX1$^+$ pancreatic cell		PDX1$^+$ stomach and duodenum cells	
Signaling molecules			
Cldn	Calmodulin	*Rab8*	Rab protein 8
Dl1	Delta like 1	*IGFBP2*	IGF binding protein 2
Dab1	Diaphonos homolog 1	*Ihh* and *Shh*	Indian and Sonic hedgehog
Notch1	Notch homolog 1		
RohB	Guanine nucleotide exchange factor		
FGFR2, 4	FGF receptor 2 or 4		
Transcription factors			
Hoxc5	Homeobox protein c5	*Cut1-L*	Cut1-like
C/EBP	CCAAT/Enhancer-binding protein	*EKLF*	Erythrocyte kluppel-like factor
		Elf3	E47-like factor 3
		KLF4	Kluppel factor 4
		Pax1	Paired box gene 1
		Sox2, 11	SRY-box containing gene 2 and 11
Cell adhesion/matrix protein			
AP47	Adaptor protein 47		
Cop4s	COP 9 homolog, 4		

*http://dev.biologists.org/supplemental/

Notch signaling members, but we also discovered that of a Notch modifier, manic fringe *(Mfng)* and a transcription factors, *Myt1* (Bellefroid et al., 1996) that participate in Notch signaling (Table 4). This finding suggests that *Mfng* and *Myt1* could be involved in endocrine cell development.

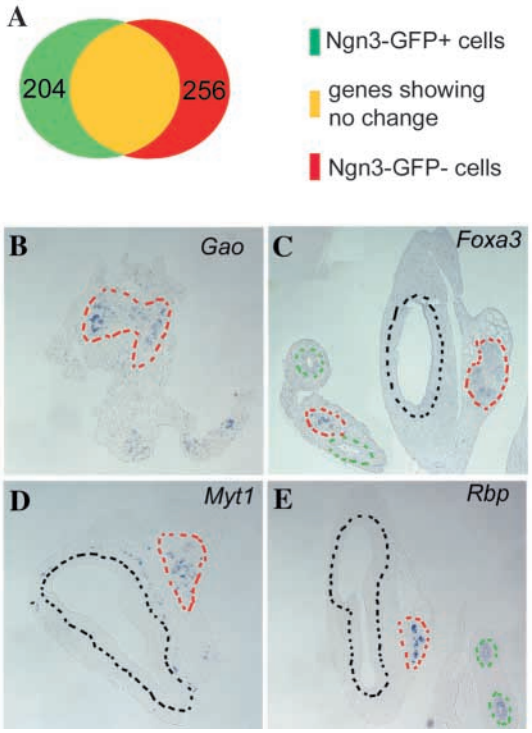

Table 4. A partial list of genes differentially expressed in endocrine compared with nonendocrine progenitor cells (for detailed description of these genes and all candidates, see Table S5*)

Endocrine ngn3+ progenitor cells (E13.5)		Nonendocrine ngn3– cells (E13.5)	
Signaling molecules			
Alg2	Calcium binding protein	*EdoR*	Endothelin R
Cadps	Ca dependent activator for secretion	*GPR26*	G-protein coupled receptor 26
Cmk2b	Ca/Calmodulin dependent kinase 2b	*IGF1*	Insulin like growth factor 1
G0α1	Guanine nucleotide binding, alpha O	*IGF2*	Insulin like growth factor 2
Gdpd1	GDP dissociation inhibitor1	*IGFBP5*	IGF binding protein 5
GPR27	G-protein coupled receptor 27	*Notch1*	Notch 1 homolog
GPR56	G-protein coupled receptor 56	*PDGFR*	PDGF receptor
Mfng	*Manic fringe*	*Sfrp1*	Secreted frizzled protein 1
Pla2g6	Phospholipase 2, group 6		
Rab3d	Member of ras oncoprotein 3d	*Sfrp2*	Secreted frizzled protein 2
Rab7	Member of ras oncoprotein 7	*Thrombin R*	Thrombin receptor
Transcription factors			
Myt1	Myelin transcription factor 1		

*http://dev.biologists.org/supplemental/

Signaling molecules expressed by NGN3$^-$ cells

The NGN3$^-$ cells included several tissue types, such as progenitor cells that had not been specified toward the endocrine cell fate (by virtue of its *Ngn3* expression), precursors that give rise to the exocrine pancreas, and mesodermally derived tissues within the pancreas. Consequently, diverse signaling pathways were found to be expressed by the NGN3$^-$ cells. Transcripts enriched in Ngn3-eGFP$^-$ cells included the endothelin receptor, PDGFR, thrombin receptor, which are known for hematopoietic development. However it is still possible that these genes are important for endocrine development.

Analysis of *Myt1* function during endocrine cell development

One goal of our gene expression analysis was to identify new genes that are functionally involved in endocrine islet development. Of the genes whose transcripts are enriched in the endocrine progenitors, one gene, *Myt1*, is a promising candidate regulator of endocrine development. In *Xenopus laevis*, xMyt1 has been shown to cooperate with xNgn1 to induce neurogenesis (Bellefroid et al., 1996). Because islet development has many similarities with that of neuronal development (Gu et al., 2003), we wanted to determine whether *Myt1* is involved in endocrine differentiation.

Fig. 5. Genes expressed in the Ngn3-eGFP$^+$ and Ngn3-eGFP$^-$ cells at E13.5. (A) Venn diagram showing that 204 genes (green) were enriched at least threefold in the Ngn3-eGFP$^+$ cells as compared to 256 (red) genes that were enriched in the Ngn3-eGFP$^-$ cells. (B-E) Expression analysis of four genes by ISH. *Gαo* (B), *Foxa3* (C), *Myt1* (D), and *Rbp* (E) were only detected at high levels in a set of scattered pancreatic cells, similar to *Ngn3*. Black dashed lines, stomach; green dashed lines, duodenum; red dashed lines, pancreas.

Table 5. The number of insulin$^+$ and glucagon$^+$ cells in wild-type and transgenic animals carrying dominant-negative *Myt1 (dnMyt1)* transgene

	Insulin$^+$ cells		Glucagon$^+$ cells	
	Number	% reduction	Number	% reduction
Wild type	2968, 2846, 2359, 2157	NA	1342, 981, 1192, 1280	NA
dnMyt1	2209, 1498, 1353, 1306, 1193	39%	647, 1286, 872, 587, 661	32%

Wild type, n=4; transgenic animals, n=5.

The *Myt1* locus produces two isoforms by utilizing alternative transcriptional starts, *Myt1* (noted as *Myt1a*) and *Nzf2b*, both containing C_2HC zinc fingers and a transcriptional activation domain. These two isoforms differ in their N-terminal 100 amino acid residues (Matsushita et al., 2002), such that NZF2b has an extra zinc finger (MYT1a has 6 zinc fingers and NZF2b has 7 zinc fingers). For simplicity, we refer to both RNA isoforms from the *Myt1* locus as *Myt1*, and we refer to the 6-zinc-finger *Myt1* cDNA as *Myt1a* (Kim et al., 1997a; Matsushita et al., 2002). Our semi-quatitative RT-PCR results showed that *Myt1a* and *Nzf2b* are both expressed in the developing pancreas, with *Nzf2b* being expressed at much higher levels (data not shown). In situ analysis using probes common to both isoforms demonstrated that *Myt1* is expressed in a few cells of the developing gut (E8.5) where the pancreatic buds will form [between the seventh and ninth somites, adjacent to the dorsal aorta (Fig. 6B and data not shown)], as well as in the nervous system (Fig. 6B). As embryogenesis proceeds, *Myt1* is expressed in the pancreas in a similar fashion to that of *Ngn3*, i.e. in a scattered subset of epithelial cells that are adjacent to or within characteristic duct-like structures (Fig. 6C). After E15.5, *Myt1* transcripts were considerably reduced (Fig. 6D), yet not abolished (longer exposure of these tissue sections yields positive *Myt1* mRNA hybridization signals, data not shown). The expression pattern of *Myt1* suggests that it functions, like *Ngn3*, during the early stages of endocrine cell specification. We utilized gain-of-function and loss-of-function approaches to determine if *Myt1* was involved in development of the endocrine pancreas, using both mouse and chicken embryos as model systems.

In mouse, we broadly misexpressed *Myt1a* in pancreatic buds during embryogenesis using a *Pdx1* promoter. We found that *Myt1a* ectopic expression did not affect exocrine or endocrine cell development at E14.5 or E16.5, in terms of pancreatic morphology or molecular marker expressions (data not shown). We next tested whether MYT1 is necessary for mouse endocrine differentiation. We constructed a dominant negative (dn) *Myt1* to inhibit MYT1 function during endocrine development by deleting the transcriptional activation domain from *Myt1a*. This strategy was previously used to inhibit *Myt1* function during neural development in *Xenopus* (Bellefroid et al., 1996). We used the *Ngn3* promoter to specifically over express *dnMyt1* in endocrine progenitor cells of first generation, transient transgenic E14.5 embryos, and characterized the pancreatic phenotype. We found that the total number of insulin- and glucagon-expressing cells were reduced on average by 39% and 32%, respectively, in five transient transgenic E14.5 embryos (Fig. 6E,F and Table 5). Although

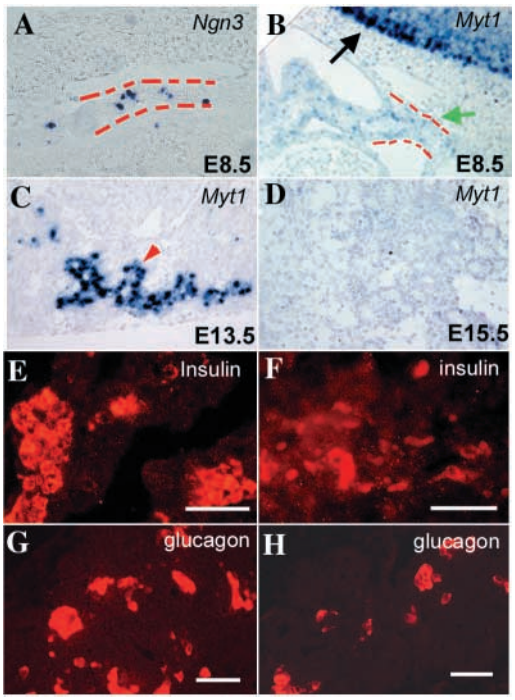

Fig. 6. Functional characterization of *Myt1* during mouse endocrine development. (A-D) In situ hybridization showing *Ngn3* and *Myt1* expression at several stages of embryogenesis. A is a control showing *Ngn3* expression in the developing pancreatic region (red dashed lines) at E8.5. (B) At E8.5 *Myt1* is expressed at a low level in the prospective pancreatic region in a manner similar to *Ngn3* (red dashed line). The prospective pancreatic region is recognized by its position below somite 8-10 (total of 15 somites), and the position of the dorsal aorta (green arrow). Strong *Myt1* expression was also detected in the neural tube (black arrow). (C) At E10.5, *Myt1* was detected in a scattered fashion in the pancreatic bud, within or close to duct-like structures (red arrowheads). (D) At E15.5, *Myt1* expression was much reduced (see text). (E-H) Transgenic expression of *dnMyt1* in endocrine progenitor cells inhibits endocrine differentiation. (E) Insulin$^+$ cells in a representative wild-type pancreas. (F) Insulin$^+$ cells in a representative pancreas section of a *Ngn3-dnMyt1* littermate. Note fewer insulin-expressing cells in F. Similarly, the number of glucagon$^+$ cells is reduced in the *Ngn3-dnMyt1* expressing embryo (H) compared with that of wild type embryo (G). In E-H, Cy3-conjugated antibodies were used. Scale bar: 25 μm.

the *Ngn3* promoter is cell specific to endocrine progenitor cells (Gu et al., 2002), the level of transgene expression may be too low for a dominant negative approach to totally abolish endocrine cell differentiation. Because somatostatin and pancreatic polypeptide are not yet expressed by E14.5, the effect of DnMYT1 on δ and PP cell development could not be determined.

As an alternative approach to test the function of *Myt1* in developing endocrine cells, we overexpressed *Myt1* (*Myt1a* and *Nzf2b*) and *dnMyt1* in chicken embryonic endoderm at approximately the 25-somite stage using electroporation (Materials and methods). This approach was previously used to demonstrate that misexpression of NGN3 in hindgut endoderm results in the differentiation of a large number of glucagon-expressing cells and a small number of somatostatin-expressing cells, but not insulin and pancreatic polypeptide-expressing

cells (Grapin-Botton et al., 2001). We found that misexpression of *Myt1a* in the chicken gut endoderm did not result in ectopic expression of any pancreatic markers (data not shown). However, misexpression of *Nzf2b* induced ectopic expression of glucagon and somatostatin, but not significant amount of insulin and pancreatic polypeptide, in the stomach and duodenum cells (Fig. 7). These results suggest that NZF2b is sufficient to partially initiate endocrine development in endoderm. Currently, it is not known why MYT1a and NZF2b have different activity in inducing endocrine marker expression. It is also not clear why *Ngn3* or *Myt1* fail to induce the formation of insulin and pancreatic polypeptide expressing cells.

Finally, we determined whether inhibition of *Myt1* activity suppresses the ability of NGN3 to induce ectopic endocrine differentiation in chick endoderm (Grapin-Botton et al., 2001). For this purpose, we co-expressed *dnMyt1* together with *Ngn3* in chicken embryonic hindgut endoderm and examined the expression of endocrine markers. Our results demonstrated that full-length *Myt1a* (Fig. 7E) or *Nzf2b* (data not shown) did not affect the ability of NGN3 to induce glucagon expression. Yet DnMYT1 significantly reduced the ability of NGN3 to induce ectopic glucagon expression (Fig. 7F). Since NGN3 has a very limited ability to induce formation of somatostatin-expressing cells, we were not able to use this dominant negative approach to determine if DnMYT1 could inhibit somatostatin expression. These results, combined with the transgenic mouse data, suggest that the ability of NGN3 to promote α and β cell differentiation depends, directly or indirectly, on *Myt1* function.

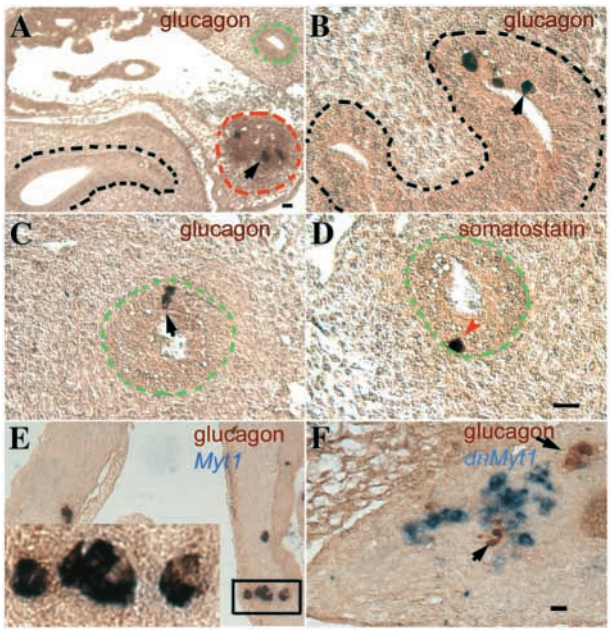

Fig. 7. *Myt1* is involved in generation of glucagon and somatostatin-expressing cells in chicken embryonic endoderm and may be in the same pathway as *Ngn3*. (A-C) In normal chicken embryos (A) glucagon (brown staining, black arrow) and somatostatin-expressing cells (data not shown) are absent in chicken stomach and duodenum. When *Nzf2b* was ectopically expressed in the chicken embryonic gut endoderm around stage 18, glucagon expression was induced in the stomach (B) and duodenum (C). A small, yet significant, number of somatostatin-expressing cells were also induced (D, red arrowhead). (E,F) dnMYT1 inhibits NGN3-mediated induction of glucagon$^+$ cell formation in chicken gut endoderm. *Ngn3* + *Myt1a* (E) or *Ngn3* + *dnMyt1* (F) were co-electroporated into chicken gut endoderm respectively. Glucagon expression (brown antibody staining) and *Myt1* or *dnMyt1* expression (blue, in situ hybridization) was measured. In E, co-expression of *Ngn3* and *Myt1a* (data not shown) (Grapin-Botton et al., 2001) resulted in induction of glucagon$^+$ cells (brown and blue color overlap resulting in dark brown/purple). In F, cells that co-express *Ngn3* and *dnMyt1* almost never express glucagon (brown and blue cells do not overlap) suggesting that the presence of dnMYT1 inhibited the ability of NGN3 to promote the generation of glucagon-expressing cells. Black dashed line, stomach; green dashed line, duodenum; red dashed line, pancreas.

Discussion

During organogenesis, specialized cell types are generated from progenitor cell populations and are precisely organized into the elaborate structure of the adult organ. This process involves numerous cell-cell communications and initiation of complex inter-regulating genetic networks to ensure fidelity of organogenesis. We have used a transcriptional profiling approach to begin to characterize the expression of regulatory or functional components during the development of early endoderm to pancreatic precursors, then to endocrine progenitors, and eventually to functional islet cells. The strength of this approach is that we started with enriched cell populations at each chosen stage of development, which we predict to greatly increase the sensitivity of the microarray analysis. The success of our analysis was immediately apparent since we detected the transcripts of most of the genes previously implicated in endocrine development in the developing endocrine islets. We subsequently have catalogued a large number of new candidate genes that may participate in islet cell development at different stage. These genes include cell-cell signaling molecules (receptors, growth factors), signal modifiers, transcription factors, transporters, ECM proteins, and many others. These analyses provide us with a global expression profile of genes that may interact to dictate the sequence of cellular development, from an unspecified progenitor to precursors whose fates are restricted to specific organs and finally to mature, functional cells.

New signaling molecules in endocrine development

In addition to those genetic pathways known to play a role in pancreatic development, our results have newly implicated several additional pathways (Table 6). In endoderm, we detected the expression of a genetic network that has been well studied in eye development. This network includes the genes *Eyes absent 2 (Eya2)* and *Sine oculis-related homeobox 1 (Six1)* that genetically interact during eye development in flies and mice (Heanue et al., 1999; Pignoni et al., 1997; Ridgeway and Skerjanc, 2001). In addition, we have identified multiple components of the *Wnt* pathway, including *Wnt* ligands, *Wnt* receptors (*Fzds*), *Wnt* receptor antagonists [secreted frizzled-related 1 and 3 (*SfrPs*)], and its downstream targets (*Dvls*) in early endoderm, general pancreatic progenitors and endocrine progenitors. In PDX1$^+$ cells in the pancreatic region (E10.5), *RhoB, Dab1, Cldn* and *FAK* are all expressed at enriched level. These genes have been shown to function in modifying cell cytoskeleton and they might be involved in pancreatic epithelia morphogenesis. In the endocrine progenitors, we found the

specific expression of G-protein cascade, *GPR14, GPR27, GPR56, Ga0, Rab3d, Rab7* and a *GDP dissociation factor* genes. These factors might interact with each other and participate in endocrine lineage differentiation. In addition, members of the calcium-activated signaling cascade may also participate in islet formation or function.

In addition to identifying new signaling pathways that possibly regulate islet formation, we identified new members of pathways that are known to function in islet differentiation. For example, we found that *Myt1* and *Mfng* are specifically expressed in the endocrine progenitors. These molecules have previously been linked to Notch signaling, either as a mediator or a modifier. Their presence in the endocrine progenitors suggests that these gene products may participate in islet formation. Our preliminary demonstration that DnMYT1 inhibits the generation of insulin and glucagon-expressing cells in mouse and/or chick supports this hypothesis.

The same signaling molecules regulate different cell fate decisions

We detected components of all common signaling pathways in each cell population representing different stages of islet generation (data not shown). Several specific growth factor receptors are expressed at each stage of development, yet probably direct the expression of different target genes, depending on the cell in which it is expressed. For example, cells at all stages analyzed expressed the activin receptor 2b. Activin/TGFβ signaling can be regulated by extracellular modifiers like Cer1, and receptors can transduce a signal via several different downstream *Smads*. It is therefore easy to speculate that the response of any given cell to activin signaling depends on many other cell-intrinsic and extrinsic factors according to the levels of signal strength and/or the competence factors present in the cells. This highlights the belief that a relatively small number of regulatory molecules can be used to determine the eventual cell type.

Global trends in gene expression to study complex biological processes

Our transcriptional profile of the developing endocrine pancreas has generated a quantitative gene expression database that can be used to analyze complex gene expression networks that would be impossible to study by other strategies. For example, our analysis suggests that the most plastic cell type, E7.5 endoderm, expressed many genes involved in cell fate specification, and the number of these genes becomes progressively fewer as endocrine cells begin to differentiate. Adult endocrine cells expressed the fewest number of cell fate regulatory genes but abundantly expressed genes associated with the adult function of the islets. The progressive decrease in the number of cell-fate regulators during endocrine development is consistent with the hypothesis that differentiation is a function of cells becoming progressively restricted toward one lineage. We identified another group of genes that were down regulated as a function of differentiation (Fig. S1 and Table S2, http://dev.biologists.org/supplemental). There is a significant number of tumor associated genes associated with this gene cluster, suggesting that the genetic machinery underlying cell plasticity in the embryo might overlap with the genes involved in the 'de-differentiation' that occurs during oncogenesis.

Table 6. A summary of new pathways detected by microarray analysis

Stage of development	Pathway members
E7.5 endoderm	*Eya2, Six1, Kit* receptor and ligand
E10.5 PDX1+ cells	*RhoB, Dab1, Caln, FAK*
E13.5 NGN3+ cells	*GpR27, GpR56, Ga0, Rabb3D, Rab7, GDPd1*
	Alg-2, Ca activator, CadKII, plpII
	Arx, Mfng, Myt1, Notch2

Expression data from these experiments will be available at www.genet.cchcc.org and can be directly compared to the expression profiles generated from other studies to look for informative biological trends between cell types and across organ systems. For example, a comparison of expression profiles between the developing and regenerating pancreas, or between two branching organs such as the pancreas and kidney could potentially uncover molecular trends associated with these processes.

Gene discovery: markers and regulators of developing pancreatic progenitor cells

Given the current research emphasis on deriving functional islets from stem or other cell types, the identification of new endocrine regulatory genes and markers is timely. There is ample evidence suggesting that many of the genes involved in endocrine pancreatic development also function in the homeostasis of the adult islet (Wells, 2003). It was our intention that a transcriptional profile of the developing endocrine pancreas would be an important resource for the diabetes research community. The genes identified in this study should facilitate analysis of the putative stem cells identified in the pancreatic ducts (Abraham et al., 2002; Cornelius et al., 1997; Ramiya et al., 2000; Zulewski et al., 2001). In addition, this catalog of signaling molecules and transcription factors expressed during endocrine development will expedite attempts to promote stem cell, embryonic or adult, differentiation into the islet cell lineage (Hori et al., 2002; Lumelsky et al., 2001).

Our temporal and spatial analysis of genes expressed in embryonic endocrine cells has generated a database of potential progenitor cell markers. For example, we have cross referenced our spatial and temporal analysis of genes expressed in E7.5 endoderm and identified 60 genes that were both spatially and temporally restricted to E7.5 endoderm (Table S1, http://dev.biologists.org/supplemental). These include genes of known endodermally expressed factors (*Sox17, Foxa2, Dkk1, Cer1*), and novel markers of endoderm (*Eya2*, cKit ligand, *Prss12*). Similar analyses revealed temporally and spatially restricted expression of genes in pancreatic and endocrine progenitor cells. We identified 16 genes that are highly enriched or only expressed in E10.5 $PDX1^+$ cells of the pancreatic rudiment (Table 4; Table S1, http://dev.biologists.org/supplemental), and 36 genes whose expression was temporally and spatially enriched in $NGN3^+$ endocrine precursors (Table 4; Table S1, http://dev.biologists.org/supplemental).

Thus far, we have not identified any genes that are exclusively restricted to developing endocrine cells. For example, *Eya2* and *Kit* ligand are expressed in E7.5 endoderm and in other tissues at later stages of development (Godin et

al., 1991; Motro et al., 1991; Xu et al., 1997). *Ngn3* and *Myt1* are both expressed in endocrine progenitor cells, as well as a set of neural progenitors in the developing nervous system (Apelqvist et al., 1999; Gradwohl et al., 2000; Kim et al., 1997b). It is possible that some of the ESTs in our database are truly expressed in a cell-specific manner. Alternatively, our results suggest that embryonic precursor cells seem to express many genes as a function of maintaining plasticity, where as adult islets expressed cell-type-specific genes. Nonetheless, the expression of a combination of several genes within each group may provide us with a diagnostic standard to determine whether cells are of endoderm, general pancreatic progenitor, endocrine precursors or mature islets.

Myt1 function might be necessary for endocrine islet development

Other than revealing general gene expression trends during islet development, our analysis also uncovered many candidate genes whose function could be required for islet development. One such example is *Myt1*. Our results suggested that the *NZF2b* isoform of *Myt1* promotes the formation of glucagon and somatostatin-expressing cells when ectopically expressed in chicken embryonic gut endoderm when it is expressed in developing endoderm (before endogenous pancreatic cells appear). MYT1a and NZF2b seemingly have different activity with regards to regulating the formation of hormone-expressing cells (insulin, glucagon and somatostatin). Differences in their activity could be due to differential stability of the proteins, or differences in post-translational modification or nuclear localization, or their different DNA binding activity.

We have also demonstrated that a dominant negative MYT1 partially inhibits endocrine cell differentiation in transgenic mouse embryos and efficiently inhibits NGN3 activity in chicken gut endoderm. Combine with the ectopic gene expression analysis, these results suggest that *Myt1* is involved in endocrine islet differentiation, and may function along the same pathway as NGN3. Although the β- and α-differentiation-inhibitory effect of dnMYT1 could result from the functional inhibition of other MYT1-like molecules, the other two *Myt1* homologues, *Myt1l* (Kim et al., 1997a) and *Myt3* (GeneBank acc. no.: BC032273) are not expressed in the developing pancreas (data not shown).

We are grateful to Kristin Stanley and Sara Rankin for assistance with Genespring and the Affymetrix software, and to Jennifer Kordich, Steve Potter, Maureen Gannon and Chris Wright for comments on the manuscript. This research was supported by grants from the Howard Hughes Medical Institute and the National Institutes of Health. J.M.W. was supported by a postdoctoral fellowship from Juvenile Diabetes Foundation.

References

Abraham, E. J., Leech, C. A., Lin, J. C., Zulewski, H. and Habener, J. F. (2002). Insulinotropic hormone glucagon-like peptide-1 differentiation of human pancreatic islet-derived progenitor cells into insulin-producing cells. *Endocrinology* **143**, 3152-3161.

Angelastro, J. M., Ignatova, T. N., Kukekov, V. G., Steindler, D. A., Stengren, G. B., Mendelsohn, C. and Greene, L. A. (2003). Regulated expression of ATF5 is required for the progression of neural progenitor cells to neurons. *J. Neurosci.* **23**, 4590-4600.

Apelqvist, A., Li, H., Sommer, L., Beatus, P., Anderson, D. J., Honjo, T., Hrabe de Angelis, M., Lendahl, U. and Edlund, H. (1999). Notch signalling controls pancreatic cell differentiation. *Nature* **400**, 877-881.

Baugh, L. R., Hill, A. A., Brown, E. L. and Hunter, C. P. (2001). Quantitative analysis of mRNA amplification by in vitro transcription. *Nucleic Acids Res.* **29**, E29.

Beddington, R. S. and Robertson, E. J. (1999). Axis development and early asymmetry in mammals. *Cell* **96**, 195-209.

Bellefroid, E. J., Bourguignon, C., Hollemann, T., Ma, Q., Anderson, D. J., Kintner, C. and Pieler, T. (1996). X-MyT1, a Xenopus C2HC-type zinc finger protein with a regulatory function in neuronal differentiation. *Cell* **87**, 1191-1202.

Bouwmeester, T., Kim, S., Sasai, Y., Lu, B. and De Robertis, E. M. (1996). Cerberus is a head-inducing secreted factor expressed in the anterior endoderm of Spemann's organizer. *Nature* **382**, 595-601.

Brissova, M., Shiota, M., Nicholson, W. E., Gannon, M., Knobel, S. M., Piston, D. W., Wright, C. V. and Powers, A. C. (2002). Reduction in pancreatic transcription factor PDX-1 impairs glucose-stimulated insulin secretion. *J. Biol. Chem.* **277**, 11225-11232.

Chiang, M. K. and Melton, D. A. (2003). Single-cell transcript analysis of pancreas development. *Dev. Cell* **4**, 383-393.

Conlon, F. L., Lyons, K. M., Takaesu, N., Barth, K. S., Kispert, A., Herrmann, B. and Robertson, E. J. (1994). A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. *Development* **120**, 1919-1928.

Cornelius, J. G., Tchernev, V., Kao, K. J. and Peck, A. B. (1997). In vitro-generation of islets in long-term cultures of pluripotent stem cells from adult mouse pancreas. *Horm. Metab. Res.* **29**, 271-277.

Edlund, H. (1998). Transcribing pancreas. *Diabetes* **47**, 1817-1823.

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.

Godin, I., Deed, R., Cooke, J., Zsebo, K., Dexter, M. and Wylie, C. C. (1991). Effects of the steel gene product on mouse primordial germ cells in culture. *Nature* **352**, 807-809.

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

Grapin-Botton, A., Majithia, A. R. and Melton, D. A. (2001). Key events of pancreas formation are triggered in gut endoderm by ectopic expression of pancreatic regulatory genes. *Genes Dev.* **15**, 444-454.

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., Brown, J. R. and Melton, D. A. (2003). Direct lineage tracing reveals the ontogeny of pancreatic cell fates during mouse embryogenesis. *Mech. Dev.* **120**, 35-43.

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.

Harrison, S. M., Dunwoodie, S. L., Arkell, R. M., Lehrach, H. and Beddington, R. S. (1995). Isolation of novel tissue-specific genes from cDNA libraries representing the individual tissue constituents of the gastrulating mouse embryo. *Development* **121**, 2479-2489.

Heanue, T. A., Reshef, R., Davis, R. J., Mardon, G., Oliver, G., Tomarev, S., Lassar, A. B. and Tabin, C. J. (1999). Synergistic regulation of vertebrate muscle development by Dach2, Eya2, and Six1, homologs of genes required for Drosophila eye formation. *Genes Dev.* **13**, 3231-3243.

Hebrok, M., Kim, S. K. and Melton, D. A. (1998). Notochord repression of endodermal Sonic hedgehog permits pancreas development. *Genes Dev.* **12**, 1705-1713.

Herrera, P. L. (2000). Adult insulin- and glucagon-producing cells differentiate from two independent cell lineages. *Development* **127**, 2317-2322.

Herrera, P. L., Orci, L. and Vassalli, J. D. (1998). Two transgenic approaches to define the cell lineages in endocrine pancreas development. *Mol. Cell Endocrinol*. **140**, 45-50.

Hori, Y., Rulifson, I. C., Tsai, B. C., Heit, J. J., Cahoy, J. D. and Kim, S. K. (2002). Growth inhibitors promote differentiation of insulin-producing tissue from embryonic stem cells. *Proc. Natl. Acad. Sci. USA* **99**, 16105-16110.

Johansson, B. M. and Wiles, M. V. (1995). Evidence for involvement of activin A and bone morphogenetic protein 4 in mammalian mesoderm and hematopoietic development. *Mol. Cell Biol.* **15**, 141-151.

Johnston, S. H., Rauskolb, C., Wilson, R., Prabhakaran, B., Irvine, K. D.

and Vogt, T. F. (1997). A family of mammalian Fringe genes implicated in boundary determination and the Notch pathway. *Development* **124**, 2245-2254.

Jonsson, J., Carlsson, L., Edlund, T. and Edlund, H. (1994). Insulin-promoter-factor 1 is required for pancreas development in mice. *Nature* **371**, 606-609.

Kim, J. G., Armstrong, R. C., v Agoston, D., Robinsky, A., Wiese, C., Nagle, J. and Hudson, L. D. (1997a). Myelin transcription factor 1 (Myt1) of the oligodendrocyte lineage, along with a closely related CCHC zinc finger, is expressed in developing neurons in the mammalian central nervous system. *J. Neurosci. Res.* **50**, 272-290.

Kim, S. K. and Hebrok, M. (2001). Intercellular signals regulating pancreas development and function. *Genes Dev.* **15**, 111-127.

Kim, S. K., Hebrok, M. and Melton, D. A. (1997b). Pancreas development in the chick embryo. *Cold Spring Harb. Symp. Quant. Biol.* **62**, 377-383.

Lammert, E., Cleaver, O. and Melton, D. (2001). Induction of pancreatic differentiation by signals from blood vessels. *Science* **294**, 564-567.

Lumelsky, N., Blondel, O., Laeng, P., Velasco, I., Ravin, R. and McKay, R. (2001). Differentiation of embryonic stem cells to insulin-secreting structures similar to pancreatic islets. *Science* **292**, 1389-1394.

Mathis, D., Vence, L. and Benoist, C. (2001). beta-Cell death during progression to diabetes. *Nature* **414**, 792-798.

Matsushita, F., Kameyama, T. and Marunouchi, T. (2002). NZF-2b is a novel predominant form of mouse NZF-2/MyT1, expressed in differentiated neurons especially at higher levels in newly generated ones. *Mech. Dev.* **118**, 209-213.

Miralles, F., Battelino, T., Czernichow, P. and Scharfmann, R. (1998a). TGF-beta plays a key role in morphogenesis of the pancreatic islets of Langerhans by controlling the activity of the matrix metalloproteinase MMP-2. *J. Cell Biol.* **143**, 827-836.

Miralles, F., Czernichow, P. and Scharfmann, R. (1998b). Follistatin regulates the relative proportions of endocrine versus exocrine tissue during pancreatic development. *Development* **125**, 1017-1024.

Motro, B., van der Kooy, D., Rossant, J., Reith, A. and Bernstein, A. (1991). Contiguous patterns of c-kit and steel expression: analysis of mutations at the W and Sl loci. *Development* **113**, 1207-1221.

Mukhopadhyay, M., Shtrom, S., Rodriguez-Esteban, C., Chen, L., Tsukui, T., Gomer, L., Dorward, D. W., Glinka, A., Grinberg, A., Huang, S. P. et al. (2001). Dickkopf1 is required for embryonic head induction and limb morphogenesis in the mouse. *Dev. Cell* **1**, 423-434.

Offield, M. F., Jetton, T. L., Labosky, P. A., Ray, M., Stein, R. W., Magnuson, M. A., Hogan, B. L. and Wright, C. V. (1996). PDX-1 is required for pancreatic outgrowth and differentiation of the rostral duodenum. *Development* **122**, 983-995.

Perfetti, V., Vignarelli, M. C., Casarini, S., Ascari, E. and Merlini, G. (2001). Biological features of the clone involved in primary amyloidosis (AL). *Leukemia* **15**, 195-202.

Pictet, R. L., Clark, W. R., Williams, R. H. and Rutter, W. J. (1972). An ultrastructural analysis of the developing embryonic pancreas. *Dev. Biol.* **29**, 436-467.

Pignoni, F., Hu, B., Zavitz, K. H., Xiao, J., Garrity, P. A. and Zipursky, S. L. (1997). The eye-specification proteins So and Eya form a complex and regulate multiple steps in Drosophila eye development. *Cell* **91**, 881-891.

Ramalho-Santos, M., Melton, D. A. and McMahon, A. P. (2000). Hedgehog signals regulate multiple aspects of gastrointestinal development. *Development* **127**, 2763-2772.

Ramiya, V. K., Maraist, M., Arfors, K. E., Schatz, D. A., Peck, A. B. and Cornelius, J. G. (2000). Reversal of insulin-dependent diabetes using islets generated in vitro from pancreatic stem cells. *Nat. Med.* **6**, 278-282.

Ridgeway, A. G. and Skerjanc, I. S. (2001). Pax3 is essential for skeletal myogenesis and the expression of Six1 and Eya2. *J. Biol. Chem.* **276**, 19033-19039.

Scearce, L. M., Brestelli, J. E., McWeeney, S. K., Lee, C. S., Mazzarelli, J., Pinney, D. F., Pizarro, A., Stoeckert, C. J., Jr, Clifton, S. W., Permutt, M. A. et al. (2002). Functional genomics of the endocrine pancreas: the pancreas clone set and PancChip, new resources for diabetes research. *Diabetes* **51**, 1997-2004.

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

Sharma, A., Zangen, D. H., Reitz, P., Taneja, M., Lissauer, M. E., Miller, C. P., Weir, G. C., Habener, J. F. and Bonner-Weir, S. (1999). The homeodomain protein IDX-1 increases after an early burst of proliferation during pancreatic regeneration. *Diabetes* **48**, 507-513.

Shimizu, K., Chiba, S., Saito, T., Kumano, K., Takahashi, T. and Hirai, H. (2001). Manic fringe and lunatic fringe modify different sites of the Notch2 extracellular region, resulting in different signaling modulation. *J. Biol. Chem.* **276**, 25753-25758.

Slack, J. M. (1995). Developmental biology of the pancreas. *Development* **121**, 1569-1580.

Stride, A. and Hattersley, A. T. (2002). Different genes, different diabetes: lessons from maturity-onset diabetes of the young. *Ann. Med.* **34**, 207-216.

Tisch, R. and McDevitt, H. (1996). Insulin-dependent diabetes mellitus. *Cell* **85**, 291-297.

Ueda, S., Mizuki, M., Ikeda, H., Tsujimura, T., Matsumura, I., Nakano, K., Daino, H., Honda Zi, Z., Sonoyama, J., Shibayama, H. et al. (2002). Critical roles of c-Kit tyrosine residues 567 and 719 in stem cell factor-induced chemotaxis: contribution of src family kinase and PI3-kinase on calcium mobilization and cell migration. *Blood* **99**, 3342-3349.

van Schaick, H. S., Smidt, M. P., Rovescalli, A. C., Luijten, M., van der Kleij, A. A., Asoh, S., Kozak, C. A., Nirenberg, M. and Burbach, J. P. (1997). Homeobox gene Prx3 expression in rodent brain and extraneural tissues. *Proc. Natl. Acad. Sci. USA* **94**, 12993-12998.

Warnock, G. L., Kneteman, N. M., Evans, M. G. and Rajotte, R. V. (1990). Isolation of purified large mammal and human islets of Langerhans. *Horm. Metab. Res.* **Supplement 25**, 37-44.

Wells, J. M. (2003). Genes expressed in the developing endocrine pancreas and their importance for stem cell and diabetes research. *Diabetes Metab. Res. Rev.* **19**, 191-201.

Wells, J. M. and Melton, D. A. (1999). Vertebrate endoderm development. *Annu. Rev. Cell. Dev. Biol.* **15**, 393-410.

Wells, J. M. and Melton, D. A. (2000). Early mouse endoderm is patterned by soluble factors from adjacent germ layers. *Development* **127**, 1563-1572.

Wilkinson, D. G. and Nieto, M. A. (1993). Detection of messenger RNA by ISH to tissue sections and whole mounts. *Methods Enzymol.* **225**, 361-373.

Wu, K. L., Gannon, M., Peshavaria, M., Offield, M. F., Henderson, E., Ray, M., Marks, A., Gamer, L. W., Wright, C. V. and Stein, R. (1997). Hepatocyte nuclear factor 3beta is involved in pancreatic beta-cell-specific transcription of the pdx-1 gene. *Mol. Cell Biol.* **17**, 6002-6013.

Xu, P. X., Woo, I., Her, H., Beier, D. R. and Maas, R. L. (1997). Mouse Eya homologues of the Drosophila eyes absent gene require Pax6 for expression in lens and nasal placode. *Development* **124**, 219-231.

Zulewski, H., Abraham, E. J., Gerlach, M. J., Daniel, P. B., Moritz, W., Muller, B., Vallejo, M., Thomas, M. K. and Habener, J. F. (2001). Multipotential nestin-positive stem cells isolated from adult pancreatic islets differentiate ex vivo into pancreatic endocrine, exocrine, and hepatic phenotypes. *Diabetes* **50**, 521-533.